

Universitat Autònoma de Barcelona
Facultat de Biociències
Departament de Bioquímica i Biologia Molecular

TESI DOCTORAL

Estudi del mecanisme d'autoregulació del gen *Trl* i caracterització funcional a nivell global del factor GAGA en cèl·lules S2 de *Drosophila melanogaster*

David Piñeyro Valerio

Abril de 2009

Programa de Bioquímica i Biologia Molecular (Opció A)

Estudi del mecanisme d'autoregulació del gen *Trl* i caracterització funcional a nivell global del factor GAGA en cèl·lules S2 de *Drosophila melanogaster*

Memòria presentada per David Piñeyro Valerio per aspirar al grau de Doctor en Biologia. Aquest treball ha estat realitzat sota la direcció del Dr. Jordi Bernués Martínez al Departament de Biologia Molecular i Cel·lular de l'Institut de Biologia Molecular de Barcelona (IBMB-CSIC).

Vist-i-plau del director

Vist-i-plau del tutor

Dr. Jordi Bernués Martínez

Dr. Jaume Farrés i Vicén

Autor

David Piñeyro Valerio

AGRAÏMENTS

Quan acabes una època de la teva vida com és el Doctorat, sorprèn mirar enrere i adonar-te de tot el que t'emportes. Ara que aquesta etapa s'acaba, crec que és just fer un petit recordatori, un mini tribut, a totes aquelles persones que d'alguna manera m'han ajudat a arribar fins aquí. El laboratori és gran, així que espero no fer-me molt pesat...

Primer, "al César lo que es del César". I és que el mèrit d'aquest treball recau també sobre el meu director de Tesi. Gràcies Jordi per fer-me de mestre en això de la ciència. Gràcies per ser un gran jefe, però a la vegada estar sempre a punt per a donar un cop de mà (què seria d'alguns clonatges sense tu!).

Gràcies també a la resta de jefes del laboratori, que amb el seu interès i bons consells, han deixat també la seva empremta en aquest treball. Gràcies Ferran, Marian, Josep, Dori i Àlex. Gràcies Lluïsa pels protocols, les teves idees i les discussions sobre GAGA. I especialment, moltes gràcies Elena, per donar-me la possibilitat de començar una nova etapa ja (per fi!) com a post-doc.

I què seria de qualsevol laboratori sense els tècnics? No sabeu com us trobo a faltar ara! Gràcies Gema i Ali (tornar a fer maxis és tan dur!).

També he d'agrar al David Rossell que fes fàcil el que per a mi semblava impossible, ets un crack! Tampoc podré oblidar a la Gema Tur i aquella pizza a les 12 de la nit fent *footprintings*. Y gracias Pablo, fue breve pero aún hoy se me escapa algún "la con... del loro!"

Gràcies també a la resta de gent del laboratori. Gràcies Mònica per l'ajuda amb els *footprintings*. Gràcies Olga (em mantindran informat, tinc contactes...), Carles, Antonio, Sylvia Mansilla, Olivera, Sergi (t'he dit ja que cada dia estàs més guapo?), Marta "Orxeta" (espero que puguem continuar xerrant quan vingui a buscar tubs!), Sònia (com trobo a faltar aquells *heat-shocks*...), Lorena (la primera fiesta en tu piso, que lejos queda ya!), Xavi "Dorito" (PC forever!), Joan (quins temps els del "Xavi-vine" i el 514!) i gràcies també Martí, per què tot i que jo ja no hi sigui, sé que amb tu al laboratori el "frikisme" no morirà. Gràcies a la Sara Sánchez i a l'Ester Valls, les ex-nenes de la Marian, i a les actuals, Naiara, Conchi i Elena.

Un lloc especial mereixen les persones del meu "grupillo" (a les que cal sumar l'Ali, és clar!), amb les que he compartit tants bons moments a l'hora de dinar (lluïtant per un lloc al bar) i també fora del Parc. Gràcies Marta Batlle, per la teva forma de ser, tan alegre i positiva. Gràcies Bet per suportar els meus comentaris biològics no científics, i per ensenyar-nos tants llocs genials per sopar (el Comendatore NO el Comodoro!). Gràcies Esther Fuentes, no només

per salvar-me la vida múltiples vegades amb el tema “mosquil”, també per ser una gran amiga i compartir tants bons moments (el “peluqueria niño-niña” no s’oblidarà mai!). Marc, sempre quedaran pel record els nostres moments íntimo-còmics fent “dientes” i gràcies per ser el millor públic d’un humorista. Xavi Aran (Guiuiuiuiu), gràcies per ser el millor company de minigrup, pel “cargadito-cargadito”, les abraçades i totes les bones estones (però encara penso que el meu humor és més intel·ligent). Noemí! Siempre seré la persona que te enseñó catalán, con esa mítica frase tan cierta ahora como entonces,... gracias por darme tu confianza y amistad casi desde el primer momento. Y, por supuesto, mi nueva compañera de laboratorio: gracias Rute por ser mi mejor amiga portuguesa, es más, la mejor portuguesa! Aunque no aceptes que “eu falo português melhor que tu”.

Però no només la gent del laboratori m’han ajudat en tots aquests anys. Els que més han aguantat queixes, moments de “baixon” i l’arribar tard sistemàticament per culpa de la feina, són els meus amics, els que considero ja de tota la vida. Sílvia, impossible agrair-ho tot, simplement, gràcies per ser la meva millor amiga. I tot contant que només et caic bé! Charlie, daaaadaadaada! Gracias por tu particular visión del humor y también por tus momentos de indignación míticos! Borja, sin tu selección musical esta Tesis hubiera sido muy dura de escribir (“larga vida a John Williams!”), pero sobretodo, ahora sí: yo antes y más joven! (como siempre). Per últim, però no menys importants, els meus quasi-germans: Isaac, Nachete, gràcies. Sou, en dues paraules, els millors. Per fi em podreu dir allò de Dr. Piñeyro amb propietat!

Y como no, un gracias como la copa de un pino a toda mi familia! Mamá, tu empanada es ya un mito que perdurará y papá, ahora ya te puedo responder: bien! La Tesis va bien! Infinitas gracias a mis hermanos Juan y Anna, y a mis respectivos cuñados Noelia y Rafa(el). A los pequeñajos: Yerik, Ruth, Iván y Derek, que me han alegrado estos últimos años. I no em puc deixar a la meva família adoptiva, que m’han fet sentir un més des del primer moment. Gràcies Isabel, Jaume, Albert, Joan, Lolita i Victor. A todos, sois la mejor familia que uno puede imaginar, y punto!

Pero sobretodo y por encima de todo, gracias Marta.

Aquesta Tesi us la dedico a tots!

ÍNDIX GENERAL

ÍNDIX GENERAL	i
ÍNDIX DE FIGURES	ix
ÍNDIX DE TAULES	xiii
ABREVIATURES	xv
A. INTRODUCCIÓ	1
A.1. L'expressió gènica	3
A.2. Transcripció a eucariotes per l'RNA Pol II	4
A.2.1. Estructura dels promotors tipus Pol II	4
A.2.2. El promotor <i>core</i> o basal dels gens Pol II	5
A.2.2.1. La caixa TATA (<i>TATA box</i>)	7
A.2.2.2. L'element Iniciador (<i>Inr</i>)	7
A.2.2.3. El DPE (<i>downstream promoter element</i>)	8
A.2.2.4. El BRE (<i>TFIIB recognition element</i>)	9
A.2.2.5. MTE (<i>Motif Ten Element</i>)	9
A.2.2.6. DCE (<i>Downstream Core Element</i>)	10
A.2.2.7. DRE (<i>DNA – replication – related Element</i>)	10
A.2.2.8. XCPE-1 (<i>X core promoter element 1</i>)	11
A.2.2.9. Característiques diferencials dels promotors basals de vertebrats i	
invertebrats	11
A.2.2.9.A. Illes CpG a vertebrats	11
A.2.2.9.B. Distribució de TSS	12
A.2.3. El Complex de Preiniciació (PIC)	13
A.2.3.1. RNA Polimerasa II	14
A.2.3.2. Els factors generals de la transcripció (GTFs)	16
	i

ÍNDIX GENERAL

A.2.3.2.A. TFIIA	16
A.2.3.2.B. TFIIB	17
A.2.3.2.C. TFIID	17
A.2.3.2.D. TFIIF	19
A.2.3.2.E. TFIIE	19
A.2.3.2.F. TFIH	20
A.2.3.3. Muntatge del PIC. Model seqüencial i model d'holoenzim	21
A.2.4. Fàbriques de transcripció	23
A.2.5. Etapes del procés de transcripció	24
A.2.5.1. Iniciació de la transcripció: <i>promoter clearance</i>	25
A.2.5.2. PrPP i elongació	26
A.2.5.3. Terminació	29
A.2.6. Paper de la cromatina en la regulació de la transcripció	29
A.2.7. Factors activadors/repressors de la transcripció	34
A.2.7.1. Factors de transcripció (FTs)	34
A.2.7.2. Cofactors o coreguladors	36
A.3. El factor GAGA de <i>Drosophila melanogaster</i>	38
A.3.1. El gen <i>Trithorax-like</i>	38
A.3.1.1. Regulació de l'expressió de <i>Trl</i>	41
A.3.2. Estructura de la proteïna GAGA	41
A.3.2.1. Domini d'unió al DNA (DBD)	42
A.3.2.2. Domini BTB/POZ	44
A.3.2.3. Domini ric en glutamines Q	45
A.3.2.4. Regió X	46
A.3.3. Funcions de GAGA	47
A.3.3.1. GAGA com a antirepressor i activador transcripcional	47
A.3.3.2. GAGA com a remodelador de la cromatina	48
A.3.3.3. GAGA com a proteïna TrxG i PcG	50
A.3.3.4. GAGA als <i>insulators</i>	52

B. OBJECTIUS	55
C. MATERIALS I MÈTODES	59
C.1. Materials	61
C.1.1. Reactius i enzims	61
C.1.2. Soques bacterianes	61
C.1.3. Plasmidis	61
C.1.3.1. Plasmidis de clonatge	61
C.1.3.2. Plasmidis reporter	63
C.1.3.3. Plasmidis per a l'expressió de proteïnes a bacteris	66
C.1.3.4. Plasmidis per a l'expressió de proteïnes a cèl·lules eucariotes	67
C.1.3.5. Plasmidis per a la producció de dsRNA	69
C.1.3.6. Plasmidis per a la producció de línies de mosques transgèniques	69
C.1.3.7. Plasmidis per a la producció de línies de cèl·lules S2 establiment transfectades	70
C.1.4. Oligonucleòtids	72
C.1.5. Anticossos específics	73
C.1.6. Cèl·lules eucariotes	74
C.1.7. <i>Stocks</i> de mosques	74
C.1.7.1. Cromosomes balancejadors i marcadors	74
C.1.7.2. Soques de <i>Drosophila melanogaster</i> utilitzades	75
C.2. Mètodes	76
C.2.1. Tècniques de manipulació de DNA plasmídic	76
C.2.1.1. Tècniques de DNA recombinant	76
C.2.1.2. Transformació de cèl·lules competents	76
C.2.1.3. Extracció de DNA plasmídic pel mètode de la lisi alcalina ràpida	76
C.2.1.4. Maxipreparació de DNA plasmídic	77
C.2.2. Mètodes de treball amb cèl·lules en cultiu	77

ÍNDEX GENERAL

C.2.2.1. Manteniment de cultius cel·lulars	77
C.2.2.2. Congelació de cèl·lules de <i>Drosophila</i>	78
C.2.2.3. Descongelació de cèl·lules de <i>Drosophila</i>	78
C.2.2.4. Comptatge de cèl·lules	79
C.2.2.5. Transfecció pel mètode del fosfat càlcic	79
C.2.2.6. Preparació d'extractes cel·lulars per a assajos d'activitat luciferasa i β -gal	80
C.2.2.7. Assaig d'activitat luciferasa	80
C.2.2.8. Assaig d'activitat β -galactosidasa	80
C.2.2.9. Tractament de cèl·lules S2 amb TSA i Nicotinamida	80
C.2.2.10. Producció de línies cel·lulars establement transfectades	81
C.2.2.11. Inducció de les cèl·lules establement transfectades amb CuSO_4	82
C.2.3. Tècniques d'anàlisi genètic de <i>Drosophila</i>	82
C.2.4. Tècniques d'anàlisi de DNA	83
C.2.4.1. Reacció en Cadena de la Polimerasa (PCR)	83
C.2.4.2. Extracció de DNA genòmic de cèl·lules	84
C.2.4.3. Extracció de DNA genòmic d'embrions	85
C.2.4.4. <i>Footprinting</i> genòmic <i>in vivo</i> amb DMS	85
C.2.4.4.A. Tractament amb DMS <i>in vivo</i> de cèl·lules S2	86
C.2.4.4.B. Obtenció i tractament <i>in vivo</i> amb DMS dels embrions transgènics de <i>Drosophila melanogaster</i>	87
C.2.4.4.C. Tractament <i>in vitro</i> amb DMS	89
C.2.4.4.D. Tractament amb àcid fòrmic <i>in vitro</i> (<i>ladder</i> G+A)	89
C.2.4.4.E. Tractament amb piperidina	90
C.2.4.4.F. <i>Linker-mediated</i> PCR (LM-PCR)	90
C.2.5. Tècniques d'anàlisi d'RNA	94
C.2.5.1. Extracció d'RNA total	94
C.2.5.2. RT-PCR	94
C.2.5.3. <i>Real Time</i> RT-PCR	96
C.2.5.4. Tractament de les cèl·lules amb RNA d'interferència	97
C.2.5.4.A. Preparació d'RNA de doble cadena (dsRNA)	97
C.2.5.4.B. Tractament de les cèl·lules amb dsRNA	99

C.2.5.5. Assajos de <i>microarray</i>	99
C.2.6. Tècniques d'anàlisi de proteïnes	100
C.2.6.1. Electroforesi en gel de poliacrilamida-SDS (SDS-PAGE)	100
C.2.6.2. <i>Western blot</i>	101
C.2.6.3. Expressió de proteïnes a bacteris	102
C.2.6.4. Assaig d'interacció proteïna-proteïna <i>in vitro</i> per GST-Pulldown	102
C.2.6.4.A. Preparació d'extractes proteics nuclears de cèl·lules S2 transfectades transitòriament per a la realització de GST-Pulldowns	103
C.2.6.4.B. Expressió i purificació de proteïnes fusionades a GST	104
C.2.6.4.C. Unió de les proteïnes fusionades a GST a la reïna <i>Glutathione Sepharose 4B</i>	105
C.2.6.4.D. Elució de les proteïnes fusionades a GST de la reïna <i>Glutathione Sepharose 4B</i>	105
C.2.6.4.E. Assaig de GST-Pulldown	106
C.2.6.5. Immunolocalitzacions a cèl·lules	107
C.2.6.6. Immunolocalitzacions a embrions	108
C.2.6.7. Immunolocalitzacions a discos imaginals de larves	109
D. RESULTATS	111
D.1. El factor GAGA de <i>Drosophila melanogaster</i> a sistemes heteròlegs	113
D.1.1. GAGA a cèl·lules de <i>Drosophila virilis</i> reprimeix <i>Trl</i>	113
D.1.2. GAGA a cèl·lules HeLa humanes activa <i>Trl</i>	116
D.2. Anàlisi detallada del promotor de <i>Trl</i> i la seva contribució a la repressió	118
D.2.1. Cerca bioinformàtica al promotor de <i>Trl</i> en busca de motius d'unió per a factors de transcripció	119
D.2.2. Anàlisi de l'ocupació del promotor de <i>Trl</i> <i>in vivo</i> per <i>footprinting</i> genòmic amb DMS	121
D.2.2.1. <i>Footprinting</i> genòmic a cèl·lules S2	121

ÍNDEX GENERAL

D.2.2.2. <i>Footprinting</i> genòmic a embrions	124
D.2.2.3. <i>Footprinting</i> genòmic a cèl·lules S2 establiment transfectades	126
D.2.3. Anàlisi del promotor de <i>Trl</i> per mitjà de delecions i mutacions	128
D.2.3.1. Delecions	129
D.2.3.2. Mutacions	130
D.3. Contribució de les diferents regions de GAGA a la repressió	132
D.4. Contribució, a la repressió de <i>Trl</i>, de diferents factors de transcripció que interaccionen amb GAGA: TTK69 i Psq no participen	135
D.5. Contribució de diferents corepressors a la repressió de <i>Trl</i>	137
D.5.1. Groucho no sembla participar en la repressió de <i>Trl</i>	137
D.5.2. Possible implicació de dCtBP en la repressió de <i>Trl</i>	138
D.6. Efectes de GAGA a mosques transgèniques	141
D.6.1. GAGA ₅₁₉ , GAGA ₅₈₁ i GAGA Δ Q reprimeixen <i>Trl</i> . GAGA ₅₁₉ BDKO no	142
D.6.2. La sobreexpressió de GAGA ₅₈₁ a mosques transgèniques produeix efectes fenotípics similars als de GAGA ₅₁₉ , però menys severs	145
D.7. GAGA reprimeix la còpia endògena de <i>Trl</i>	148
D.7.1. RT-PCRs i <i>Real Time</i> RT-PCRs de cèl·lules S2 establiment transfectades que sobreexpressen GAGA	149
D.7.2. RT-PCRs i <i>Real Time</i> RT-PCRs de discos imaginals de mosques transgèniques que sobreexpressen GAGA	150
D.8. Estudi funcional de la sobreexpressió i de la reducció de la quantitat de GAGA, en cèl·lules S2, per mitjà de <i>microarrays</i>	152
D.8.1. <i>Sickle</i> s'activa per la sobreexpressió de GAGA	160

E. DISCUSSIÓ	163
E.1. Estudi de la repressió de <i>Trl</i> per GAGA	165
E.1.1. GAGA a sistemes heteròlegs	165
E.1.2. Contribució del promotor de <i>Trl</i> a la repressió	168
E.1.3. Contribució de les diferents regions de GAGA a la repressió	173
E.1.4. Contribució de diversos factors al mecanisme de repressió de <i>Trl</i>	175
E.1.5. Repressió de <i>Trl</i> per GAGA, a la mosca	178
E.1.6. Efectes fenotípics de la sobreexpressió de GAGA a mosques transgèniques	179
E.1.7. Repressió de GAGA sobre la còpia endògena del gen <i>Trl</i>	181
E.1.8. Models per explicar el mecanisme de repressió de <i>Trl</i> per GAGA	183
E.2. Estudi funcional de GAGA a S2 per mitjà de <i>microarrays</i>	187
F. CONCLUSIONS	197
G. BIBLIOGRAFIA	203
H. ANNEXOS	221
H.1. Alineaments	223
H.1.1. Alineament de GAGA a 12 espècies de <i>Drosophila</i>	223
H.1.2. Alineament de Cid a 12 espècies de <i>Drosophila</i>	224
H.1.3. Alineament dels promotors de <i>Trl</i> de <i>D. melanogaster</i> i <i>D. virilis</i>	224
H.1.4. Estructura gènica de <i>Trl</i> a 12 espècies de <i>Drosophila</i>	226

ÍNDIX GENERAL

H.2. Comparacions de les dades dels experiments de *microarray*

i de *ChIP on chip* de GAGA 226

H.2.1. Comparació del *microarray* de sobreexpressió i el *ChIP on chip* de GAGA 226

H.2.2. Comparació del *microarray* d'RNAi i el *ChIP on chip* de GAGA 227

H.2.3. Comparació dels gens que solapen entre els dos experiments de *microarray* i les dades de *ChIP on chip* 228

ÍNDIX DE FIGURES

Figura A.1. Esquematització d'un promotor tipus Pol II.	5
Figura A.2. El promotor <i>core</i> és la diana última pels factors de transcripció i coreguladors de l'RNA Pol II.	6
Figura A.3. Resum esquemàtic dels diferents elements del promotor <i>core</i> caracteritzats avui dia, amb els corresponents GTFs que se sap que interaccionen.	6
Figura A.4. Representació dels dos tipus de promotors segons la seva distribució de TSS.	13
Figura A.5. Diferents models pel muntatge del PIC.	22
Figura A.6. Model per a una "fàbrica de transcripció" de l'RNA Pol II i la seva estructura.	24
Figura A.7. Resum del procés de transcripció per RNA Pol II i del fenomen de <i>promoter-proximal pausing</i> (PrPP).	28
Figura A.8. Models per a la terminació de la transcripció.	29
Figura A.9. Distribució de les diferents modificacions de les histones al llarg de la regió gènica i el seu efecte en transcripció.	32
Figura A.10. Estructura genòmica del gen <i>Trl</i> de <i>D. melanogaster</i> i els seus diferents transcrits.	40
Figura A.11. Esquema dels dominis de les dues isoformes de la proteïna GAGA.	42
Figura A.12. Representació esquemàtica del DBD de GAGA ₅₁₉ .	43
Figura A.13. Esquema de la regió reguladora del complex <i>Bithorax</i> .	52
Figura A.14. Propietats dels <i>insulators</i> .	53
Figura C.1. Esquema del plasmidi pMK33-CTAP.	81
Figura C.2. Mètode de preparació del medi Fauré.	83
Figura C.3. Esquema d'un experiment de <i>footprinting in vivo</i> .	86
Figura C.4. Esquema dels creuaments fets pels experiments de <i>footprinting</i> genòmic <i>in vivo</i> amb DMS.	88
Figura C.5. Esquematització de la LM-PCR.	91
Figura C.6. Condicions en les que s'han realitzat les RT-PCRs.	96
Figura C.7. Esquema dels plasmidis pBS-RNAi-dCtBP, pBS-RNAi-Psq i pBS-RNAi-GFP.	98
Figura D.1. Comparació de seqüència entre les isoformes de GAGA de <i>D. melanogaster</i> i les de <i>D. virilis</i> .	114
Figura D.2. La repressió de <i>Trl</i> per GAGA està conservada a <i>D. virilis</i> .	116
Figura D.3. Transfeccions de GAGA a HeLa.	117

ÍNDIX DE FIGURES

Figura D.4. Resum esquemàtic de la cerca bioinformàtica.	120
Figura D.5. Seqüència del promotor de <i>Trl</i> esquema dels diferents fragments.	122
Figura D.6. Anàlisi per <i>footprinting</i> genòmic amb DMS de les regions FII i FIII del promotor de <i>Trl</i> .	123
Figura D.7. <i>Footprinting</i> amb DMS a embrions.	125
Figura D.8. Caracterització de la línia de cèl·lules GAGA.	127
Figura D.9. Resum dels resultats obtinguts per la Dra. Ana Kosoy realitzant successives deleccions del promotor de <i>Trl</i> i assajant la seva activitat relativa a cèl·lules S2.	128
Figura D.10. Resum dels resultats de les transfeccions transitòries a S2 amb les deleccions del promotor de <i>Trl</i> .	129
Figura D.11. Esquema de les diferents mutacions del promotor de <i>Trl</i> realitzades.	130
Figura D.12. Transfeccions transitòries a S2 amb les construccions corresponents de cada mutació del promotor de <i>Trl</i> .	131
Figura D.13. Esquema resum de les diferents formes truncades de GAGA que es van realitzar i assajar.	133
Figura D.14. Transfecció transitòria de les formes truncades de GAGA.	134
Figura D.15. TTK69 i Psq no participen en la repressió de <i>Trl</i> per GAGA.	136
Figura D.16. Groucho sembla no estar implicat en el mecanisme de repressió de <i>Trl</i> per GAGA.	138
Figura D.17. Estudi de la contribució de dCtBP a l'activitat del promotor de <i>Trl</i> , així com a la seva repressió per GAGA.	140
Figura D.18. GAGA ₅₁₉ reprimeix <i>Trl</i> tant a embrió com a discos imaginals de larva.	143
Figura D.19. GAGA Δ Q reprimeix <i>Trl</i> tant a discos imaginals de larva com a embrió.	144
Figura D.20: GAGA ₅₁₉ DBDKO no reprimeix <i>Trl</i> mentre que GAGA ₅₈₁ sí.	145
Figura D.21. Comparació dels diferents fenotips d'ala, derivats de la sobreexpressió de GAGA ₅₁₉ i GAGA ₅₈₁ .	147
Figura D.22. GAGA reprimeix la còpia endògena de <i>Trl</i> a cèl·lules S2 establement transfectades.	150
Figura D.23. GAGA reprimeix la còpia endògena de <i>Trl</i> a disc imaginal d'ala de larves de 3r estadi.	151
Figura D.24. Comprovació per <i>western blot</i> de la correcta inducció de les cèl·lules emprades pels experiments de <i>microarray</i> .	153
Figura D.25. Resultats més estadísticament significatius dels dos experiments de <i>microarray</i> .	155
Figura D.26. Agrupació, segons la seva funció biològica, dels gens diferencialment expressats a l'experiment de sobreexpressió, utilitzant l'eina DAVID.	157
Figura D.27. Agrupació, segons la seva funció biològica, dels gens diferencialment expressats a	158

l'experiment d'RNAi, utilitzant l'eina DAVID.

Figura D.28. Creuament de les dades obtingudes amb els dos tipus d'experiment de <i>microarray</i> .	160
Figura D.29. La sobreexpressió de GAGA ₅₁₉ fa augmentar sickle a S2, al menys a un cert nombre de cèl·lules.	162
Figura D.30. La sobreexpressió de GAGA ₅₁₉ fa augmentar sickle a discos d'ala de larves de 3r estadi.	162
Figura E.1. Dues possibles hipòtesis per explicar l'activació de <i>Trl</i> per GAGA a HeLa.	168
Figura E.2. Resum esquemàtic del promotor de <i>Trl</i> (fragment p2b), mostrant les dades que hem obtingut de tots els experiments amb les formes delecionades i/o mutades d'aquest.	172
Figura E.3. Esquema de la proteïna GAGA ₅₁₉ i els seus dominis.	175
Figura E.4. Una de les hipòtesis possibles per explicar el mecanisme de repressió de <i>Trl</i> per GAGA.	176
Figura E.5. Representacions esquemàtiques dels diferents models hipotètics proposats per explicar la repressió de <i>Trl</i> per GAGA.	183
Figura E.6. Representació esquemàtica de la regió gènica de <i>Trl</i> , així com dels possibles transcrits.	189
Figura E.7. Possible explicació per a la sensiblement major correlació entre el <i>microarray</i> d'RNAi i les dades de <i>ChIP on chip</i> .	192
Figura E.8. Major esbiaix de GAGA cap a una funció activadora.	193
Figura E.9. Comparació de la relació de grups de gens, segons funció biològica, que s'activen per la sobreexpressió de GAGA i es reprimeixen al reduir-la.	194
Figura E.10. Llistat de gens que pertanyen al grup d'adhesió cel·lular, que apareixen activats al <i>microarray</i> de sobreexpressió i reprimits al <i>microarray</i> d'RNAi.	194

ÍNDIX DE TAULES

Taula A.1. Paral·lelismes entre subunitats de les diferents RNA Pol presents a bacteris, arquea i eucariotes, basats en la seva homologia estructural i funcional.	3
Taula A.2. Comparació de les subunitats de les diferents RNA Polimerases d'eucariotes.	15
Taula A.3. Nomenclatura dels TAFs implicats en la transcripció per l'RNA Pol II.	18
Taula A.4. Modificacions de les histones relacionades amb transcripció.	31
Taula C.1. Relació dels oligonucleòtids emprats en aquest treball.	72
Taula C.2. Relació dels anticossos emprats en aquest treball.	73
Taula C.3. Balancejadors i marcadors genètics de <i>Drosophila</i> utilitzats en aquest treball.	75
Taula C.4. <i>Stocks</i> de mosques emprats per a la realització d'aquest treball.	75
Taula C.5. Reacció de PCR en unes condicions estàndard.	84
Taula C.6. Relació d'encebadors emprats als <i>footprintings</i> , i la seva temperatura d' <i>annealing</i> .	94
Taula D.1. Letalitat observada per la sobreexpressió de GAGA ₅₁₉ i GAGA ₅₈₁ a mosques transgèniques.	146

ABREVIATURES

<i>Abd-A / Abd-B</i>	<i>Abdominal-A / Abdominal-B</i>
ACF	<i>ATP-utilizing chromatin assembly and remodeling factor</i>
Act	<i>actina</i>
ATP	adenosín trifosfat
BIP2	<i>Bric à brac Interacting Protein 2</i> , també conegut com a dTAF _{II} 155 o dTAF3
BR	<i>Basic Region</i> , regió bàsica
BRE	<i>TFIIB recognition element</i>
BRF	<i>TFIIB related factor</i>
BSA	albúmina sèrica bovina
BTB/POZ	<i>broad complex tramtrack bric-a-brac/poxvirus and Zinc finger</i>
BX-C	complex <i>Bithorax</i>
CDK	<i>cycline-dependent kinase</i>
cDNA	DNA complementari
ChIP	<i>Chromatin immunoprecipitation</i>
CHRAC	<i>chromatin remodeling and assembly complex</i>
Cid	<i>Centromere identifier</i>
CMV	citomegalovirus
CtBP	<i>C-terminal Binding Protein</i> , dCtBP a <i>Drosophila</i>
CTD	domini C-terminal, de la subunitat major de l'RNA Pol II
DAPI	4',6-diamidino-2-phenylindole
DAVID	<i>Database for Annotation, Visualization and Integrated Discovery</i>
DBD	<i>DNA binding domain</i> , domini d'unió a DNA
DCE	<i>Downstream Core Element</i>
DMS	Dimetilsulfat
DMSO	Dimetilsulfòxid
DNA	àcid desoxiribonucleic
dNTPs	desoxiribonucleòtids trifosfat
DPE	<i>downstream promoter element</i>
<i>dpp</i>	<i>decapentaplegic</i>
DRE	<i>DNA-replication-related Element</i>
DREF	<i>DRE binding Factor</i>
DSIF	<i>DRB (5,6-dichloro-1-h-D-ribofuranosylbenzimidazole) sensitivity-inducing factor</i>
DSP1	<i>Dorsal switch protein 1</i>
dsRNA	<i>double strand RNA</i> , RNA de doble cadena
E(var)	<i>Enhancers of variegation</i>
ECFP	<i>Enhanced Cyan Fluorescent Protein</i>
EGFP	<i>Enhanced Green Fluorescent Protein</i>
<i>en</i>	<i>engrailed</i>
<i>eve</i>	<i>even skipped</i>
FACT	<i>facilitates chromatin transcription</i>
FBS	sèrum boví fetal
FTs	factors de transcripció
<i>ftz</i>	<i>fushi tarazu</i>

ABREVIATURES

g	força relativa de centrifugació, també RCF
GBP	complex GAGA-Batman-Psq
GEFs	<i>General Elongation Factors</i>
GFP	<i>Green Fluorescent Protein</i>
Gro	groucho
GST	Glutatió S-Transferasa
GTFs	<i>General Transcription Factors</i> , factors generals de transcripció
H	hipersensibilitat
HAT	<i>Histone Acetyltransferase</i>
HDACs	<i>histone deacetylases complexes</i> , complexos desacetilases d'histones
HSE	<i>Heat shock element</i>
HSF	<i>Heat shock factor</i>
<i>hsp</i>	<i>heat shock protein</i>
Inr	Iniciador
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ISWI	<i>imitation switch</i>
KDa	KiloDaltons
Kpb	Kiloparells de bases
Kr	<i>Krüppel</i>
LB	Luria-Bertani
LM-PCR	<i>linked-mediated PCR</i>
Mod(mdg4)	<i>modifier of mdg4</i>
mRNA	RNA missatger
mRNP	Ribonucleoproteïna missatgera
MTE	<i>Motif Ten Element</i>
NAD	nicotin-adenosín dinucleòtid
NCBI	<i>National Center for Biotechnology Information</i>
ncRNAs	RNAs no codificants de proteïnes
NELF	<i>Negative Elongation Factor</i>
NER	<i>Nucleotide Excision Repair</i>
NLS	<i>Nuclear Localization Signal</i> , senyal de localització nuclear
NPC	<i>Nuclear pore complex</i> , complex del porus nuclear
NURF	<i>nucleosome remodeling factor</i>
ORF	<i>Open Reading Frame</i> , pauta oberta de lectura
P	protecció
PAGE	Electroforesi en gel de poliacrilamida
pb	parells de bases
PC	<i>Polycomb</i>
PcG	<i>Polycomb group</i>
PCNA	<i>Proliferating Cell Nuclear Antigen</i>
PCR	<i>Polymerase Chain Reaction</i> , Reacció en Cadena de la Polimerasa
PEV	<i>Position Effect Variegation</i>
PGRP-SD	<i>peptidoglycan-recognition protein-sd precursor</i>
PH	<i>Polyhomeotic</i>
PIC	<i>pre-initiation complex</i> , complex de preiniciació

ABREVIATURES

PLB	<i>Protein Loading Buffer</i>
Pol	Polimerasa
<i>prd</i>	<i>paired</i>
PREs	<i>Polycomb Response Elements</i>
PrPP	<i>promoter-proximal pausing</i>
Psq	Pipsqueak
<i>ptc</i>	<i>patched</i>
P-TEFb	<i>Positive Transcription Elongation Factor b</i>
RAP30 i 74	<i>RNA Polymerase II - associated protein 30 i 74</i>
RMN	ressonància magnètica nuclear
RNA	àcid ribonucleic
RNAi	RNA d'interferència
RNPs	ribonucleoproteïnes
Rpb1-12	RNA Polimerasa B 1-12
rpm	revolucions per minut
rRNA	RNA ribosòmic
RT	retrotranscripció
SAP18	<i>Sin3-associated polypeptide 18KDa</i>
SDS	sodium dodecil sulfat
<i>skl</i>	<i>sickle</i>
snRNA	<i>small nuclear RNA, RNA nuclear petit</i>
SRBs	<i>supressors of RNA Polymerase B mutations</i>
ssDNA	<i>single strand DNA, DNA de cadena senzilla</i>
Su(var)	<i>Suppressors of variegation</i>
SWI/SNF	<i>switch mating type/sucrose non-fermenting</i>
TAFs	<i>TBP-Associated factors</i>
TBP	<i>TATA box binding protein</i>
TFII	<i>Transcription Factor II, factor de transcripció de l'RNA Pol II</i>
TLF	<i>TBP-like factor</i>
TLP	<i>TBP-like protein</i>
TREs	<i>Trithorax Response Elements</i>
TRFs	<i>TBP-related factors</i>
<i>Trl</i>	<i>Trithorax-like</i>
tRNA	RNA de transferència
TRP	<i>TBP-related protein</i>
TrxG	<i>Trithorax group</i>
TSA	Tricostatina-A
TSS	<i>Transcription Start Site, lloc d'inici de la transcripció</i>
TTK / TTK69	Tramtrack / Tramtrack 69
UAS	<i>Upstream Activating Sequence</i>
<i>Ubx</i>	<i>Ultrabithorax</i>
UTR	<i>untranslated region, regió no traduïda</i>
XCPE-1	<i>X core promoter element 1</i>

A. INTRODUCCIÓ

A.1. L'expressió gènica

La regulació de l'expressió gènica és de vital importància per a qualsevol organisme ja que és el procés que permet que, a partir d'un conjunt determinat de gens, es pugui respondre de manera adequada a diferents estímuls, ja provinquin del medi extern o intern del propi organisme.

L'expressió gènica implica tot un conjunt de processos, cadascun finament regulat. La metodologia clàssica emprada per a l'estudi d'aquests processos sovint ha portat a una visió reduccionista, en la qual cada una de les fases estava separada de la resta. El panorama actual ha permès una comprensió més global de l'expressió gènica. Així, ara s'accepta que les etapes consecutives estan íntimament relacionades i són interdependents, estant física i funcionalment lligades (Orphanides et al. 2002). Una de les etapes clau d'aquest procés és aquella en la que la informació continguda a la seqüència de DNA (àcid desoxiribonucleic) que compon els gens, és copiada a una cadena d'RNA (àcid ribonucleic). Aquest fenomen, conegut com a transcripció, és objecte d'una fina i complexa regulació.

Els responsables directes de la síntesi d'RNA són uns complexos proteics formats per múltiples subunitats anomenats RNA Polimerases (RNA Pol). Existeixen diferents tipus d'RNA Pol segons el tipus d'organisme. Així, podem distingir entre l'RNA Pol de bacteris, l'RNA Pol dels arqueas i les tres RNA Pol (I, II i III) d'eucariotes. En conjunt es podria considerar que formen una família de proteïnes altament conservada. Tot i les evidents diferències, és sorprenent l'alt grau d'homologia estructural i/o funcional entre subunitats concretes, tret que permet establir paral·lelismes entre diverses subunitats dels diferents tipus d'organismes (veure taula A.1). El fet que organismes tan diferents i evolutivament separats presentin subunitats de les seves RNA Pol molt conservades evidencia la vital importància d'aquestes proteïnes i de la funció que realitzen (Ebright 2000). Així doncs, no és estrany que també es trobin semblances en la regulació dels gens que transcriuen.

RNA Pol de bacteris	β'	β	α'	α''	ω	
RNA Pol d'arquea	A'/A''	B	D	L	K	+ altres 6
RNA Pol I d'eucariotes	RPA1	RPA2	RPC5	RPC9	RPB6	+ altres 9
RNA Pol II d'eucariotes	RPB1	RPB2	RPB3	RPB11	RPB6	+ altres 6
RNA Pol III d'eucariotes	RPC1	RPC2	RPC5	RPC9	RPB6	+ altres 11

Taula A.1. Paral·lelismes entre subunitats de les diferents RNA Pol presents a bacteris, arquea i eucariotes, basats en la seva homologia estructural i funcional (Ebright 2000).

INTRODUCCIÓ

Centrant l'atenció en els organismes eucariotes, l'RNA Pol I típicament s'associa a la síntesi d'RNA ribosòmic (rRNA); l'RNA Pol II és la responsable de la síntesi dels RNA missatgers (mRNA) i d'alguns tipus d'RNA petits nuclears (snRNA, de l'anglès *small nuclear RNAs*); per últim, l'RNA Pol III sintetitza els RNA de transferència (tRNA), la resta de snRNA i molts RNAs de mida petita. Recentment s'ha descobert un altre tipus d'RNA Pol (la IV) només present a plantes, que sembla involucrada en mecanismes de silenciament de la cromatina via RNA (Herr et al. 2005; Onodera et al. 2005; Vaughn et al. 2005).

Tot i que les diferents polimerases comparteixen trets comuns, com una alta homologia al centre actiu (Cramer et al. 2008) o la incapacitat de totes tres per reconèixer específicament els promotors dels gens (el que fa necessari l'actuació d'altres factors), difereixen substancialment en detalls concrets dels seus mecanismes d'acció (Novina et al. 1996).

A.2. Transcripció a eucariotes per l'RNA Pol II

La transcripció gènica a organismes eucariotes és un procés complex precedit per múltiples esdeveniments: descondensació del *locus* a transcriure, remodelació dels nucleosomes, modificació de la cua de les histones d'aquests, unió de factors activadors i coactivadors als *enhancers* (amplificadors) i promotors, i finalment reclutament de la maquinària basal de la transcripció al promotor *core* o basal. És important comentar que aquests esdeveniments no formen part d'un procés fix, ni tenen un ordre d'execució establert. Per a cada gen a transcriure ocorren de la manera específica que necessita. Un cop la maquinària basal ha sigut ancorada al promotor basal, formant el que s'anomena PIC (*pre-initiation complex* o complex de preiniciació), és quan tenen lloc les tres fases clàssiques de la transcripció: iniciació, elongació i terminació. En contra del que es va pensar durant molt de temps, ara està clar que aquest procés es dona a la vegada que el processament de la cadena d'RNA que l'RNA Pol II va formant (pre-mRNA) (Orphanides et al. 2002).

A.2.1. Estructura dels promotors tipus Pol II

En general, cada una de les tres RNA Pol tenen uns requeriments especials. Així, pels gens transcrits per l'RNA Pol II parlarem de que tenen promotors tipus Pol II.

A grans trets, s'ha considerat que l'estructura típica per a un promotor està composta per dos tipus de regions (veure figura A.1). Una adjacent a l'inici de transcripció (TSS, per *Transcription Start Site*), anomenada promotor *core* o basal, que conté els elements per a captar l'RNA Pol II i els GTFs (Factors Generals de Transcripció), que formen el PIC. L'altra regió estaria formada

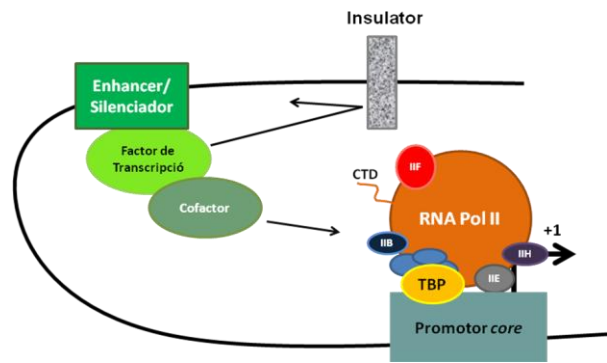


Figura A.1. Esquematzació d'un promotor tipus Pol II.

per seqüències de reconeixement per a diferents factors de transcripció activadors (regions *enhancer* o amplificadores) i/o repressors (regions silenciadores), que modularien l'activitat del promotor i que acostumen a estar situades distants al promotor basal. Aquestes dues regions es comunicarien per mitjà de cofactors, que bé poden ser activadors (coactivadors) o repressors (corepressors). A més, existeixen regions *insulator* que impedeixen la propagació dels senyals activadors dels *enhancers* o repressors dels silenciadors i d'altres efectes repressors del tipus heterocromatització (Kuhn et al. 2003).

A.2.2. El promotor *core* o basal dels gens Pol II

Clàssicament, el promotor *core* o basal ha sigut definit com el mínim fragment continu de DNA capaç de dirigir correctament nivells basals d'iniciació de la transcripció (Butler et al. 2002). Així, el promotor basal és el lloc d'acció de la maquinària transcripcional de l'RNA Pol II. D'acord amb aquesta definició, el promotor basal usualment s'estén uns 40-50 pb (parells de bases) *upstream* ("riu amunt" o anteriors) i *downstream* ("riu avall" o posteriors) del lloc d'inici de la transcripció i pot contenir varis elements diferents del promotor *core*.

Encara que es podria pensar que l'estructura dels promotors *core* ha de ser invariable, per facilitar la funció de l'RNA Pol, en realitat s'ha vist que és d'una variabilitat estructural i funcional enorme. A més, sembla que aquesta variabilitat, lluny de presentar un problema, contribueix de manera significativa a la combinatòria de l'expressió gènica (Smale et al. 2003).

El promotor *core*, és de vital importància per a la regulació de l'expressió de qualsevol gen i de fet, és l'objectiu últim de l'acció dels factors de transcripció i els coreguladors, els quals regulen l'activitat de cada gen (veure figura A.2) (Kadonaga 2002).

INTRODUCCIÓ

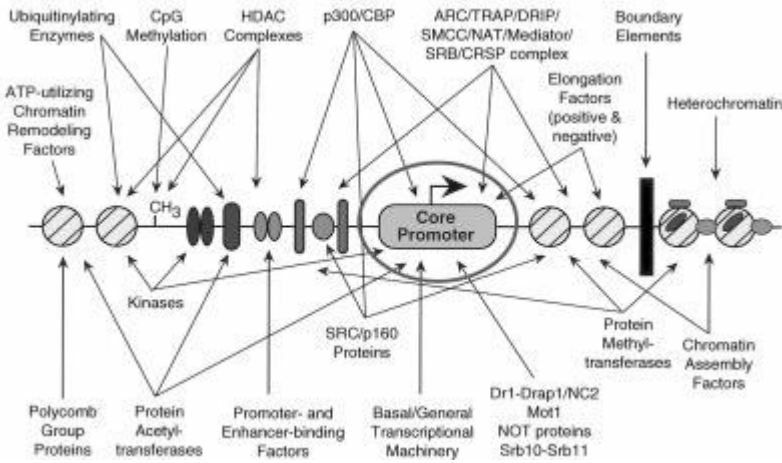


Figura A.2. El promotor *core* és la diana última pels factors de transcripció i coreguladors de l'RNA Pol II (Kadonaga 2002).

Tot i que aquesta gran variabilitat fa difícil la caracterització dels seus elements, sí que es poden trobar certes similituds que han ajudat a establir uns quants elements que són comuns a alguns promotors. Alguns d'aquests elements, donada la seva naturalesa

modular, poden coexistir en un mateix promotor *core*, però algunes combinacions són més freqüents que d'altres i alguns elements complementen d'altres. Les noves tècniques d'anàlisi informàtic de seqüències, així com el constant augment de genomes seqüenciats, ha permès que els últims anys s'hagin definit multitud de nous elements del promotor *core* i ha fet paleses les grans diferències en l'estructura del promotor *core* entre diferents espècies. A continuació es recullen els elements més ben caracteritzats actualment (veure figura A.3 per a resum esquemàtic).

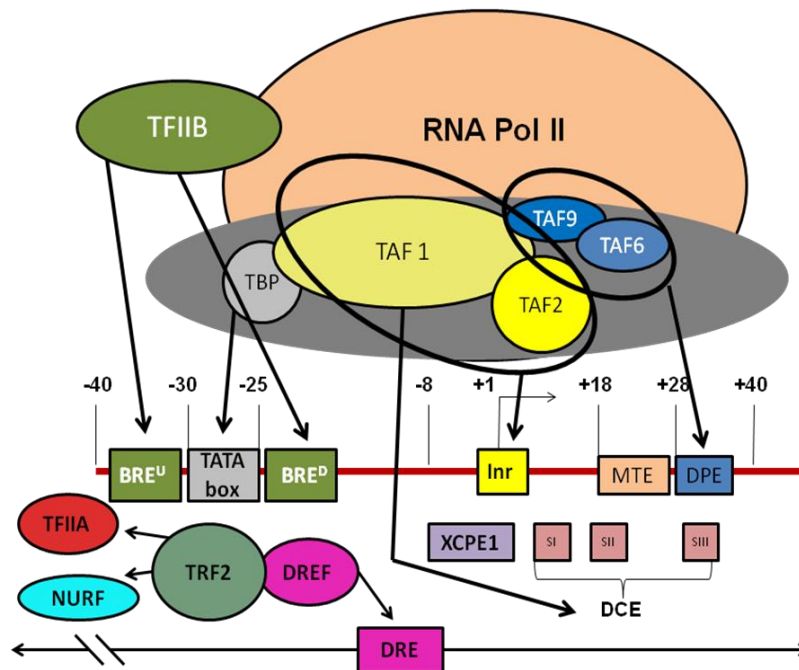


Figura A.3. Resum esquemàtic dels diferents elements del promotor *core* caracteritzats avui dia, amb els corresponents GTFs que se sap que interaccionen. No existeixen elements universals per als promotors *core*, per tant, cada un dels elements aquí mostrats està present únicament en una fracció del total de promotors. És probable que encara quedin molts elements per identificar.

A.2.2.1. La caixa TATA (TATA *box*)

Aquest va ser el primer element del promotor *core* identificat a eucariotes. Primerament es va creure que seria un element present a la majoria de promotors Pol II, però s'ha vist que en realitat és un element poc freqüent que, per exemple, actualment s'estima que només està present en menys del 10% dels promotors humans (FitzGerald et al. 2004; Bajic et al. 2006) i segons certs estudis, entre gairebé un 20% i menys d'un 10% als promotors de *Drosophila* (FitzGerald et al. 2006; Gershenzon et al. 2006). Acostuma a estar situat entre -25 i -30 pb de l'inici de transcripció, tot i que hi ha notables excepcions, com el cas de *Saccharomyces cerevisiae*, la TATA *box* del qual pot situar-se entre -40 i -200 pb de l'inici de transcripció. La seqüència consens és TATAAA, però s'ha observat que un gran rang de seqüències poden actuar com a TATA *box in vivo* (Singer et al. 1990). La funció primordial d'aquesta seqüència és unir la TATA *box binding protein* (TBP), la qual és part del complex TFIID (*Transcription Factor IID*), un factor general de transcripció de l'RNA Pol II. Les TATA *box* s'han associat a promotors forts i usualment poden coexistir amb l'Iniciador (Inr, comentat més endavant) i actuar sinèrgicament. La unió de TBP a la TATA *box*, als promotors on és present, força al PIC a seleccionar un lloc d'inici de la transcripció en una regió delimitada de l'espai genòmic (Sandelin et al. 2007). Tot i que TBP té una lleugera tendència per unir la TATA *box* amb una certa orientació, aquesta no sembla determinant per a l'orientació *in vivo* de l'inici de la transcripció. Actualment s'atribueix aquesta funció a la posició relativa entre els llocs d'unió d'activadors, la TATA *box* i altres elements del promotor *core* (Smale et al. 2003).

Sembla que TBP és la proteïna predominant d'unió a la TATA *box*, però existeixen altres factors estretament lligats a la TBP, els *TBP-related factors* (TRFs), alguns dels quals també poden unir-se a la caixa TATA (veure més endavant, apartat A.2.3.2.C).

A.2.2.2. L'element Iniciador (Inr)

A mamífers ve definit per la seqüència YYANWYY, on Y és qualsevol pirimidina (citosina o timina), A és l'adenina corresponent a la posició +1 de l'inici de transcripció, N és qualsevol nucleòtid i W correspon a una adenina o una timina. A *Drosophila* aquest motiu és similar, però no idèntic, i ve definit pel consens TCA(G/T)TY, on l'adenina també correspon al +1 de l'inici de transcripció. Així doncs, aquest element comprèn l'inici de transcripció als promotors on és

INTRODUCCIÓ

present. És un motiu independent de la TATA *box*, tot i que els dos poden estar presents al mateix promotor *core* i actuar sinèrgicament. La TATA *box* i l'Inr són els dos únics elements que poden per si sols reclutar el PIC i iniciar la transcripció en experiments *in vitro* (Sandelin et al. 2007). Diversos estudis a *Drosophila* situen aquest element entre un 14% i més del 20% dels promotors analitzats (Ohler et al. 2002; FitzGerald et al. 2006) o inclús a més del 60% (Gershenson et al. 2006). El que sí està clar és que sembla tenir una prevalença més important que la TATA *box*. Aquest element no és simètric i la seva contribució en dictar la direccionalitat del promotor sembla ser mínima, al igual que la TATA *box*.

Sembla que existeixen diferents factors capaços d'interaccionar amb l'Inr. El més conegut és el complex multiproteic TFIID (*Transcription Factor II D*), per mitjà de dos dels seus TAFs (*TBP-Associated factors*): TAF_{II}250 i TAF_{II}150 (TAF1 i TAF2 en la nova nomenclatura) (Chalkley et al. 1999). Però també altres factors semblen poder interaccionar amb l'Inr, com és el cas de la pròpia RNA Pol II, que tot i semblar molt ineficient en reconèixer el DNA específicament per si sola, sí se l'ha observat una feble preferència intrínseca per elements Inr (Carcamo et al. 1991). TFII-I, identificat inicialment com un GTF d'estructura complicada amb un domini d'unió a DNA de sis hèlix-volta-hèlix, podria funcionar com un GTF específic de reconeixement dels elements Inr (Roy et al. 1991; Roy et al. 1993), tot i que segons certs estudis *in vitro* no sembla essencial per a la funció dels elements Inr (Martinez et al. 1998). Per últim YY-1, una proteïna amb un dit de Zinc, descrita com a un factor de transcripció per a alguns promotors vírics, també sembla tenir un paper en la funció de l'Inr d'aquests promotors (Seto et al. 1991).

A.2.2.3. El DPE (*downstream promoter element*)

Situat entre +28 i +32 de l'inici de transcripció, acostuma a estar present en promotors sense caixa TATA, però no de forma exclusiva. El seu consens definit a *Drosophila* és (A/G)G(A/T)(C/T)(G/A/T), però s'ha vist que a mamífers també existeixen seqüències similars amb la mateixa funció. A *Drosophila* podria estar present en un 22% dels promotors (Gershenson et al. 2006), tot i que altres estimacions no són, ni de lluny, tan optimistes i el situen en prop del 8% dels promotors (Ohler 2006). Pot anar associat freqüentment a un Inr i es pensa que funciona cooperativament amb aquest motiu. Es creu que la seva funció és dirigir la nucleació del PIC de la mateixa manera que ho fa la caixa TATA, i de fet els factors TAF 6 i TAF 9

de TFIID el poden reconèixer per mitjà dels seus *histone fold domains* (dominis de plegament d'histones) (Burke et al. 1997; Kutach et al. 2000; Shao et al. 2005).

A.2.2.4. El BRE (*TFIIB recognition element*)

Originàriament es va identificar a humans com un element que interaccionava amb TFIIB i que es trobava immediatament *upstream* de la caixa TATA d'alguns promotors. Estudis recents també el situen a *Drosophila* (Gershenzon et al. 2006). La seva seqüència consens es va definir com (G/C)(G/C)(A/G)CGCC (Lagrange et al. 1998). Posteriorment es va veure que un altre element situat immediatament *downstream* de la caixa TATA també unia TFIIB, per una altra regió de la proteïna. Així, finalment es varen definir el BRE^U (BRE *Upstream*), que és l'originàriament identificat, i el BRE^D (BRE *downstream*) que reconeix la seqüència consens G/A-T-T/G/A-T/G-G/T-T/G-T/G (Deng et al. 2005). És l'únic element del promotor *core* de gens que codifiquen per a proteïnes pel qual està ben caracteritzat que interacciona amb un GTF diferent de TFIID. S'ha associat tant a elements activadors com repressors de la transcripció, tot i que els mecanismes pels quals funciona encara no es coneixen prou (Deng et al. 2007). La seva abundància als promotors no es coneix a fons. Un estudi bioinformàtic a partir de seqüències de promotors humans apunta a prop d'un 25% de promotors amb BRE^U, i semblaria que aquest seria més abundant en els promotors sense caixa TATA (28,1% d'aquests tindrien BRE^U) que en els que en tenen (11,8% d'aquests contenen també un BRE^U). A més BRE^U és l'únic element que és més probable que aparegui en promotors amb illes CpG que en promotors sense (Gershenzon et al. 2005). Estudis fets per BRE^D en promotors amb caixa TATA indiquen una freqüència similar de BRE^D en aquests promotors i independent de la presència de BRE^U (Deng et al. 2005).

A.2.2.5. MTE (*Motif Ten Element*)

Identificat a *Drosophila*, amb la seqüència consens "C(G/C)A(A/G)C(G/C)(G/C)AACG(G/C)", sembla que també podria estar conservat a humans. Situat entre les posicions +18 i +27 relatives a l'adenina a +1 de l'Inr. Està aparentment associat a l'Inr i funciona independentment de la caixa TATA i el DPE, tot i que mostra un fort sinergisme amb aquests motius. Aquest motiu interacciona amb certes subunitats del complex TFIID (Lim et al. 2004). A

INTRODUCCIÓ

Drosophila podria estar present en un 10% dels promotors (Gershenzon et al. 2006). Els estudis de comparació de seqüències fets a humans no semblen detectar-lo de manera significativa (FitzGerald et al. 2006), tot i que hi ha certes evidències de que també podria estar-hi present (Lim et al. 2004).

A.2.2.6. DCE (*Downstream Core Element*)

Identificat al promotor del gen de la β -globina humana, s'ha vist que també està àmpliament representat a la resta del genoma humà, així com al de certs virus (Lee et al. 2005) i a espècies tan allunyades com *Drosophila* (Gershenzon et al. 2006). Es localitza al llarg dels 30 pb *downstream* el TSS i està format per tres subelements (Lewis et al. 2000). En un mateix promotor es poden donar tots tres subelements, cap o qualsevol combinació d'ells. El subelement I (S_I) està situat entre +6 i +11 i té una seqüència consens CTTC, el subelement II (S_{II}) situat entre +16 i +21 té una seqüència consens CTGT, per últim el subelement III (S_{III}) situat entre +30 i +34 té una seqüència consens AGC. Tot i que la posició del S_{III} coincideix espacialment amb la del DPE, s'ha vist que tant el S_I com el S_{III} no apareixen en promotors amb DPE. El DCE és reconegut per la subunitat TAF1 de TFIID (Lee et al. 2005).

A.2.2.7. DRE (*DNA – replication – related Element*)

Diversos estudis fets a *Drosophila* adreçats a trobar elements consens als seus promotors destaquen especialment, apart dels elements ja coneguts, la identificació del motiu DRE en un alt número de promotors (Ohler et al. 2002). Aquest motiu va ser inicialment descrit a *Drosophila* com un element dels promotors dels gens de la DNA Polimerasa α i PCNA (*Proliferating Cell Nuclear Antigen*). És capaç d'unir un factor anomenat DREF (*DRE binding Factor*) i consta d'una seqüència palindròmica de 8 pb (5'-TATCGATA-3') (Hirose et al. 1993). Estudis posteriors han identificat DREF a altres espècies com per exemple humans, on s'ha vist que també reconeix en els promotors seqüències similars al DRE de *Drosophila* (Ohshima et al. 2003). Els estudis més recents situen la freqüència del motiu DRE als promotors de *Drosophila* en més d'un 20% (Ohler 2006). Comparant els diferents promotors amb DRE, aquest motiu no sembla tenir unes posicions establertes dins el promotor, i tot i mostrar una lleugera tendència per situar-se *upstream* el TSS, es pot trobar a ambdós costats del TSS i inclús situar-se més lluny

del TSS que altres elements del promotor *core* (Ohler et al. 2002). A part del seu paper com a element regulador de gens de proliferació i diferenciació cel·lular, el DRE s'ha trobat a multitud d'altres tipus de gens (Matsukage et al. 2008). S'ha vist que el complex DRE/DREF està implicat en portar TRF2 a certs promotors per mitjà d'interaccions TRF2-DREF (Hochheimer et al. 2002). De fet, en estudis duts a terme per Isogai i col·laboradors es pot veure com aproximadament un 20% de promotors que responen a TRF2 contenen el motiu DRE, sent el motiu més freqüent en aquests tipus de promotors (Isogai et al. 2007a).

A.2.2.8. XCPE-1 (*X core promoter element 1*)

Identificat al promotor del gen X del virus de la hepatitis B, és present al voltant de l'1% dels promotors humans, la majoria sense caixa TATA. De moment no s'ha trobat a altres eucariotes. Situat entre la posició -8 i +2 relatiu al TSS, la seva seqüència consens és G/A/T-G/C-G-T/C-G-G-G/A-A-G/C₊₁-A/C. Actua juntament amb diversos activadors de la transcripció com són NF-1, NRF1 i Sp1. S'ha postulat que XCPE-1 podria ser membre d'una família de motius que actuen, juntament amb activadors específics de seqüència, dirigint l'inici de la transcripció a promotors amb illes CpG (Tokusumi et al. 2007).

A.2.2.9. Característiques diferencials dels promotors basals de vertebrats i invertebrats

A.2.2.9.A. Illes CpG a vertebrats

Una característica diferencial dels promotors de vertebrats respecte als d'invertebrats és la presència d'illes CpG. Les illes CpG són seqüències genòmiques en les que el dinucleòtid CG està més representat del que és normal, sense presentar una seqüència consens com a tal. Centrant-nos en mamífers, diversos estudis estadístics prediuen que estan presents en prop del 72% dels promotors (Saxonov et al. 2006). Estan freqüentment associades a promotors de gens *House-keeping* i de gens d'expressió ubíqua. Només una fracció petita dels gens que contenen illes CpG s'ha vist que també continguin caixes TATA (Schug et al. 2005). Es pensa que la presència d'illes CpG pot estar relacionada amb el control epigenètic de la transcripció, per mitjà de la metilació del DNA (Siegfried et al. 1999). Generalment es considera que els gens actius contenen illes CpG hipometilades i que el correcte establiment i manteniment del patró de

INTRODUCCIÓ

metilació a cada llinatge cel·lular és essencial per a la seva correcta funció, tot i que el patró de metilació de certs gens s'ha vist que pot canviar en resposta a certs estímuls (Costello et al. 2001; Kangaspeska et al. 2008).

Als promotors dels gens d'invertebrats, com *Drosophila*, no es troben illes CpG. Durant un temps aquesta carència es va relacionar amb el fet que a *Drosophila* no s'havia trobat metilació al DNA. Als darrers anys s'ha vist que el DNA de *Drosophila* sí es troba metilat (Gowher et al. 2000) tot i que predominantment a llocs CpT/A, no CpG (Lyko et al. 2000; Kunert et al. 2003). A altres espècies d'insectes, com els del gènere *Apis*, sí s'han trobat metilases funcionals *in vivo* específiques de CpG. Tot i amb això els nivells de metilació a invertebrats semblen més baixos que a vertebrats (Wang et al. 2006) i encara no s'han relacionat directament amb regulació de la transcripció.

A.2.2.9.B. Distribució de TSS

Parant atenció en l'ús dels llocs d'inici de la transcripció, l'anàlisi (sobretot a mamífers) d'un ampli rang de promotors ha portat a identificar diverses classes de promotors *core*. Així, als extrems d'un ampli espectre trobem dos tipus bàsics de promotors (veure figura A.4). Uns on la transcripció s'inicia normalment en un únic lloc definit i constant, i uns altres on la transcripció pot iniciar-se a diversos punts en una zona àmplia del promotor, indicant múltiples TSS possibles (Carninci et al. 2006). Aquests últims acostumen a mostrar una distribució aleatòria dels TSS en una àmplia finestra d'uns 100 pb. No s'ha de confondre aquest últim tipus de promotors amb els gens amb diversos promotors alternatius, on cada un d'aquests promotors sí acostuma a tenir un TSS definit i està regulat de manera particular (Kimura et al. 2006).

Els promotors amb TSS definit estan presents en un ampli rang d'organismes, tant invertebrats com vertebrats, mentre que els promotors amb una distribució àmplia de TSS estan més relacionats amb illes CpG, presents a vertebrats, sobretot vertebrats superiors com mamífers. Així, es podria considerar que els promotors amb TSS definit són més antics evolutivament parlant que els de distribució àmplia (Juven-Gershon et al. 2008). A més, comparant per exemple humà i ratolí, es veu que els promotors amb un TSS definit evolucionen més lentament que els de distribució àmplia (Carninci et al. 2006), suggerint que el fet de tenir un únic lloc d'inici de la transcripció fa a aquests promotors més sensibles a qualsevol canvi. Per contra, els promotors amb distribució àmplia, amb els seus múltiples TSS redundants,

respondrien a la majoria de mutacions amb petits canvis en l'activitat del promotor, el que facilita l'adaptació evolutiva i la compensació de mutacions desfavorables (Sandelin et al. 2007). S'ha observat que els promotors amb TSS definit s'associen més a promotors amb els elements de seqüència descrits a dalt, sobretot la caixa TATA. L'altre tipus de promotor s'associa més a illes CpG, acostumen a no tenir caixa TATA, i tenen menys freqüència d'elements consens (Muller et al. 2007).

A diferència dels promotors de vertebrats, els promotors de *Drosophila* acostumen a tenir un TSS definit i la prevalença de motius de seqüència definida és major. S'especula que a mesura que un genoma va creixent, el més important per a la seva regulació és mantenir silenciats el que no s'ha d'expressar. Així, la diferència de mida entre el genoma de *Drosophila* i l'humà (uns deu cops més gran) podria en part explicar les diferents estratègies utilitzades en la regulació gènica. Segons això, el pas crític en la regulació gènica d'humans seria desfer la repressió basal del promotor i en el cas de *Drosophila* seria muntar correctament la maquinària basal. Això podria explicar l'evolució de les illes CpG a vertebrats, ja que poden ser fàcilment reprimides per metilació del DNA. A la vegada també explicaria l'escassíssim grau de metilació del DNA a *Drosophila*, ja que el silenciament gènec no seria tan crític (FitzGerald et al. 2006).

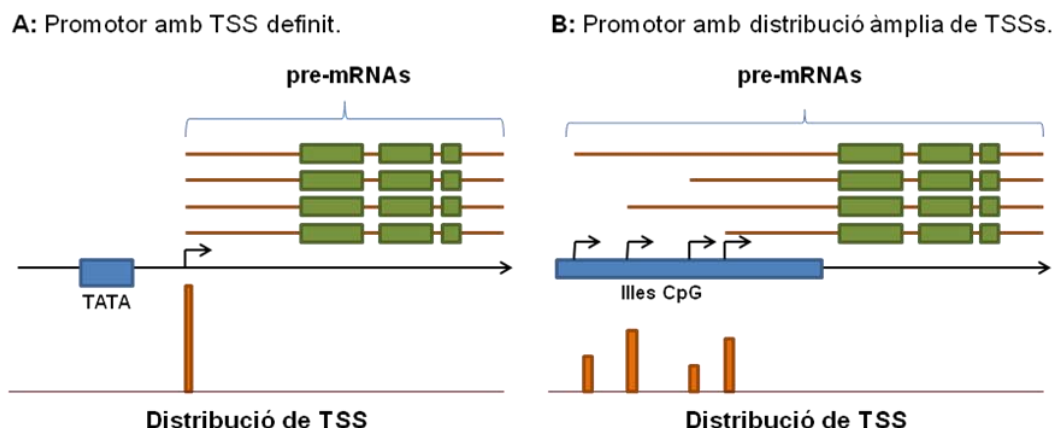


Figura A.4. Representació dels dos tipus de promotors segons la seva distribució de TSS. Els promotors amb TSS definit (A) estan presents a tots els eucariotes, però són més freqüents als genomes d'invertebrats i acostumen a anar acompanyats d'elements del promotor com la caixa TATA. Els promotors amb distribució àmplia de TSS (B) són els més freqüents a vertebrats i normalment porten associades illes CpG (Sandelin et al. 2007).

A.2.3. El Complex de Preiniciació (PIC)

El pas previ a l'inici de la transcripció és la formació del complex de preiniciació (PIC) al promotor *core* del gen a transcriure. Aquest és un pas clau en la transcripció i objecte d'una

INTRODUCCIÓ

acurada regulació per part de la cèl·lula. Per aquesta raó ha sigut també objecte d'un intens estudi. Clàssicament s'ha considerat que el PIC està format pel complex responsable directe de la síntesi de la cadena d'RNA, l'RNA Pol II en el cas dels mRNA, i els factors necessaris per que aquesta reconegui específicament el promotor i pugui iniciar correctament la transcripció, els GTFs (Factors Generals de la Transcripció).

A.2.3.1. RNA Polimerasa II

És el complex multiproteic encarregat de la síntesi dels mRNAs dels gens que codifiquen per a proteïnes, així com de certes espècies de ncRNAs (RNAs no codificants de proteïnes). Encara que la funció primària de les RNA Pol I i III és transcriure ncRNAs, l'RNA Pol II s'ha vist que també en produeix una certa quantitat. Entre aquests cal destacar diferents snRNAs, i d'altres amb funcions diverses en la regulació gènica com funcionar com "antisentit" (*antisense*) per a altres RNAs, o associar-se a RNPs (ribonucleoproteïnes) implicades en la síntesi proteica. D'entre aquests ncRNAs produïts per l'RNA Pol II alguns poden ser poliadenilats, tot i que la seva funció és desconeguda (Koch et al. 2008).

L'RNA Pol II està formada per un total de 12 subunitats, amb una massa conjunta superior a 500 KDa, codificades per diferents gens, del Rpb1 al Rpb12 (RNA Polimerasa B). El nucli de l'enzim està constituït per 10 subunitats al qual s'associen perifèricament les dues subunitats Rpb4 i Rpb7 formant un heterodímer. Existeix una conservació extensiva entre les diferents subunitats de l'RNA Pol II, a més d'existir també una alta homologia entre algunes subunitats de les tres RNA Pol d'eucariotes (veure taula A.2). A banda d'això, com a mostra de la gran conservació que existeix en aquest complex multiproteic, també s'han trobat homòlegs de diferents subunitats de l'RNA Pol II a l'RNA Polimerasa de bacteris i arquea (Cramer et al. 2008).

La més gran de les subunitats (Rpb 1) conté una llarga repetició d'un heptapèptid altament conservat que conforma el seu domini C-terminal (CTD). Aquest heptapèptid, la seqüència consens del qual és Tyr-Ser-Pro-Thr-Ser-Pro-Ser, és essencial per a la regulació de la transcripció. Tot i que la seqüència està conservada, el número de repeticions és variable entre espècies. Així, trobem que a llevat està compost per 26 o 27 repeticions, a *C. elegans* 34, a *Drosophila* 45, i a mamífers, humans inclosos, 52 (Hampsey 1998; Oelgeschlager 2002).

RNA Polimerasa	Pol I	Pol II	Pol III
Nucli de deu subunitats	A190	Rpb1	C160
	A135	Rpb2	C128
	AC40	Rpb3	AC40
	AC19	Rpb11	AC19
	A12.2	Rpb9	C11
	Rpb5 (ABC27)	Rpb5	Rpb5
	Rpb6 (ABC23)	Rpb6	Rpb6
	Rpb8 (ABC14.5)	Rpb8	Rpb8
	Rpb10 (ABC10 α)	Rpb10	Rpb10
	Rpb12 (ABC10 β)	Rpb12	Rpb12
Subcomplex Rpb4/7	A14	Rpb4	C17
	A43	Rpb7	C25
Subcomplex tipus TFIIIF (a)	A49	(Tfg1/Rap74)	C37
	A34.5	(Tfg2/Rap30)	C53
Subcomplex específic Pol III	–	–	C82
	–	–	C34
	–	–	C31
Número de subunitats	14	12	17

Taula A.2. Comparació de les subunitats de les diferents RNA Polimerases d'eucariotes. Es situen alineades les subunitats que presenten una gran homologia. (a) Les dues subunitats de l'RNA Pol I i Pol III es prediu que formen heterodímers que s'assemblen en part al factor TFIIIF d'iniciació/elongació de la transcripció, compost per Tfg1 i Tfg2 a *Saccharomyces cerevisiae* i per Rap74 i Rap30 a humans (Cramer et al. 2008).

In vivo hi ha dues formes de l'RNA Pol II. La forma IIO, hiperfosforilada al CTD i la forma IIA, hipofosforilada. La forma IIA preferencialment es troba al PIC, mentre que la forma hiperfosforilada IIO es troba al complex d'elongació (Dahmus 1996). Cinc dels set residus de l'heptapèptid són potencialment fosforilables (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) i la fosforilació del CTD s'ha vist que juga un important paper en la regulació de l'activitat transcripcional de l'RNA Pol II, la seva estabilitat, l'acoblament de la transcripció amb el processament del pre-mRNA i la reparació del DNA associada a transcripció. A més, s'han observat canvis globals en la fosforilació del CTD durant el decurs del cicle cel·lular, en resposta a factors de creixement, a mitògens, a les infeccions víriques i en resposta al dany sobre el DNA (Oelgeschlager 2002). La fosforilació de l'heptapèptid ocorre majoritàriament a les Serines (Ser) (Koch et al. 2008). Més endavant s'explicarà amb més detall la implicació de les fosforilacions a les Ser 5 i 2 en les diverses etapes de la transcripció (veure apartat A.2.5). La fosforilació de la Ser 7 ha sigut recentment descoberta i encara no s'ha trobat la quinasa responsable. Se sap que hi ha un pic de fosforilació de la Ser 7 cap el final de la transcripció tant dels mRNAs com dels snRNAs codificats per l'RNA Pol II. La mutació d'aquest residu de l'heptapèptid no afecta a la producció correcta dels mRNAs, però sí té efectes dràstics en el processament dels extrems 3' dels snRNAs,

INTRODUCCIÓ

on es pensa que, juntament amb la fosforilació de la Ser 2, podria ser responsable del reclutament del complex *Integrator*, necessari per a la correcta transcripció dels snRNAs (Egloff et al. 2008).

A.2.3.2. Els factors generals de la transcripció (GTFs)

Com ja s'ha dit abans, l'RNA Pol II per si mateixa és incapaç d'iniciar específicament la transcripció *in vitro* (Manley et al. 1980). En experiments inicials es va veure que aquesta només es donava si s'afegeixen una sèrie de components proteics purificats d'extractes cel·lulars crus (Matsui et al. 1980). Posteriorment es van poder aïllar i identificar cada un dels components necessaris per a la transcripció *in vitro*. Aquests factors s'anomenen "factors generals de la transcripció" (GTFs) i sorprenentment formen part d'una maquinària fortament conservada a eucariotes. Aquests factors s'anomenen TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH. TF per *Transcription Factor*, II per l'RNA Pol II, i la lletra fa referència a la fracció de la que van ser eluïts i l'ordre de descobriment.

A.2.3.2.A. TFIIA

El paper que juga aquest GTF ha sigut particularment elusiu. Inicialment va ser classificat com a GTF pel fet que fraccions d'extracte cel·lular que contenien TFIIA eren necessàries per a la transcripció *in vitro* (Reinberg et al. 1987a). Tot i això, estudis no molt posteriors demostraren que això no és cert per a tots els casos (Van Dyke et al. 1989). Aquestes inconsistències inicials venien donades pel grau de purificació dels extractes utilitzats per a les reaccions *in vitro*. S'ha vist que TFIIA estimula la transcripció estabilitzant la unió de TBP a la caixa TATA (Buratowski et al. 1989; Meisterernst et al. 1991). S'ha suggerit que TFIIA funciona com un antirepressor de les activitats inhibidores presents a les altres fraccions necessàries per a la transcripció, probablement incrementant l'afinitat de TBP o TFIID pel DNA i potenciant així la formació del PIC (Lee et al. 1992a; Imbalzano et al. 1994). Aquesta funció la fa interaccionant directament amb TBP, pel cantó oposat a TFIIB, i també amb el DNA per la regió immediatament *upstream* de la caixa TATA (Lagrange et al. 1996; Bleichenbacher et al. 2003).

A *Drosophila* està format per tres subunitats: α , β i γ . Les dues primeres estan codificades pel mateix gen el producte del qual es separa per proteòlisi. Un fet característic de TFIIA $\alpha\beta$ és que a les cèl·lules somàtiques sembla existir majoritàriament en una forma

proteolitzada, donant lloc a un TFIIA heterotrimèric, mentre que a cèl·lules embrionàries semblen coexistir la forma sencera i la proteolitzada de TFIIA $\alpha\beta$, donant un TFIIA heterodimèric (DeJong et al. 1993; Yokomori et al. 1993; Ozer et al. 1994; Mitsiou et al. 2000).

A.2.3.2.B. TFIIB

TFIIB és absolutament necessari per a la transcripció *in vitro* (Reinberg et al. 1987b). És l'únic GTF format per un únic polipèptid que existeix com a monòmer en solució. Està molt conservat a eucariotes, presenta homòlegs a arqueas i té certa homologia amb el factor σ de procariotes (Reinberg et al. 1987b; Hisatake et al. 1991; Wampler et al. 1992; Yamashita et al. 1992; Creti et al. 1993; Palenchar et al. 2006). El paper de TFIIB en la formació del PIC va ser evident des d'un principi i de fet posteriorment s'ha vist que TFIIB és capaç d'interaccionar amb gran part de la maquinària basal així com el propi DNA (sobretot pel BRE, veure apartat A.2.2.4). A més d'aquestes funcions, TFIIB és capaç d'interaccionar amb multitud de factors activadors de la transcripció que podrien modular la seva capacitat per unir-se al promotor *core*, estimulants així la transcripció (Deng et al. 2007). A part del seu paper en la formació del PIC, TFIIB està també implicat en la selecció del TSS (Pinto et al. 1992; Hawkes et al. 1999).

A.2.3.2.C. TFIID

És un dels primers GTFs que s'uneix al promotor *core* i dirigeix la formació del PIC. Està format per la TBP i un conjunt de Factors Associats a TBP (TAFs), de mida molt diversa (veure taula A.3). El conjunt de TAFs, tot i que és característic de cada espècie, també està altament conservat als eucariotes. Primerament es va anomenar als TAFs per la seva massa molecular, la qual s'ha vist que difereix entre espècies. Per aquesta raó es va proposar una nomenclatura unificada pels TAFs, que facilita la seva comparació entre espècies i que és la que avui dia s'utilitza (Tora 2002). El fet que tant la TBP com alguns TAFs reconeguin elements del promotor *core* ha portat a classificar TFIID com a un GTF d'unió al promotor *core*. Així, per mitjà de TBP pot unir-se als promotors amb caixa TATA i per mitjà d'alguns TAFs pot unir-se a altres elements del promotor *core* com l'Inr, el DCE i el DPE, conferint a TFIID la capacitat d'unir-se a promotors sense caixa TATA (veure figura A.3).

A més dels components descrits de TFIID, existeixen proteïnes paràlogues (codificades per altres gens però de la mateixa família) tant per a TBP com pels TAFs. Els paràlogs dels TAFs

INTRODUCCIÓ

s'associen a formes específiques de certs llinatges cel·lulars i semblen implicats en diferenciació cel·lular (Thomas et al. 2006). Els paràlogs de TBP són els *TBP-related factors* (TRFs). S'han identificat 4 TRFs a eucariotes. De tots només TRF2, també anomenat *TBP-related protein* (TRP), *TBP-like factor* (TLF) i *TBP-like protein* (TLP), sembla present a tots els metazous (Ohbayashi et al. 1999; Kaltenbach et al. 2000; Veenstra et al. 2000). Aquest no conserva la mateixa especificitat de seqüència que TBP però sí que conserva inalterats els dominis d'unió a TFIIA i TFIIB, amb els quals s'ha vist que interacciona (Dantonel et al. 1999; Moore et al. 1999).

New name	<i>H. sapiens</i>	<i>D. melanogaster</i>	<i>S. cerevisiae</i>	<i>S. pombe</i>	<i>C. elegans</i>	
					New name	Previous name
TAF1	TAF ₂₅₀	TAF ₂₃₀	Taf145/130	TAF ₁₁₁	taf-1	taf-1
TAF2	TAF ₁₅₀	TAF ₁₅₀	Taf150	T38673*	taf-2	taf-2
TAF3 [†]	TAF ₁₄₀	TAF ₁₅₅	Taf47		taf-3	C11G6.1*
TAF4 [†]	TAF _{130/135}	TAF ₁₁₀	Taf48	T50183*	taf-4	taf-5
TAF4b [†]	TAF ₁₀₅					
TAF5	TAF ₁₀₀	TAF ₈₀	Taf90	TAF ₇₂	taf-5	taf-4
TAF5b				TAF ₇₃		
TAF5L	PAF65 α	Cannonball				
TAF6 [†]	TAF ₈₀	TAF ₆₀	Taf60	CAA20756*	taf-6.1	taf-3.1
TAF6L	PAF65 α	AAF52013*			taf-6.2	taf-3.2
TAF7	TAF ₅₅	AAF54162*	Taf67	TAF _{62/PTR6}	taf-7.1	taf-8.1
TAF7L	TAF2Q				taf-7.2	taf-8.2
TAF8 [†]	BAB71460*	Prodos	Taf65	T40895*	taf-8	ZK1320.12*
TAF9 [†]	TAF _{32/31}	TAF ₄₀	Taf17	562536*	taf-9	taf-10
TAF9L	TAF _{31L}					
TAF10 [†]	TAF ₃₀	TAF ₂₄	Taf25	T39928*	taf-10	taf-11
TAF10b		TAF ₁₆				
TAF11 [†]	TAF ₂₈	TAF _{30α}	Taf40	CAA93543*	taf-11.1	taf-7.1
TAF11L					taf-11.2	taf-7.2
TAF12 [†]	TAF _{20/15}	TAF _{30α}	Taf61/68	T37702*	taf-12	taf-9
TAF13 [†]	TAF ₁₈	AAF53875*	Taf19	CAA19300*	taf-13	taf-6
TAF14			Taf30/ANC1*			
TAF15	TAF ₆₈ *					
BTAf1	TAF ₁₇₀	HelB98	Mot1	T40642*	btaf-1	F15D4.1*

Taula A.3. Nomenclatura dels TAFs implicats en la transcripció per l'RNA Pol II (Thomas et al. 2006).

Pel que fa a la resta de TRFs, TRF1 s'ha trobat només a *Drosophila* i *Anopheles* (Crowley et al. 1993; Isogai et al. 2007b) i, al igual que TBP, pot unir-se a seqüències TATA, a TFIIA i a TFIIB (Hansen et al. 1997). Tot i així, la majoria de TRF1 a *Drosophila* està associat a BRF (TFIIB related factor), un factor de transcripció de l'RNA Pol III molt relacionat amb TFIIB, i es considera necessari per a la transcripció per RNA Pol III, a aquests organismes (Takada et al. 2000). TRF3 sembla restringit a vertebrats, és el TRF que més s'assembla a TBP i, a diferència de TRF1 i TRF2, s'expressa de forma ubiqua a tots els teixits examinats (Persengiev et al. 2003). També s'uneix a la caixa TATA i a TFIIA i TFIIB (Bartfai et al. 2004) i forma complexos diferents de TFIID que podrien estar implicats en diferenciació cel·lular (Deato et al. 2007). Per últim, TRF4 s'ha descrit només per a Tripanosomes (Ruan et al. 2004) on es requereix per la transcripció de gens de les

tres RNA Pol nuclears, el que fa plausible la idea de que TRF4 sigui el substitut de TBP a aquests paràsites.

A.2.3.2.D. TFIIF

Està format pels factors RAP30 i RAP74 (*RNA Polymerase II - associated protein 30 i 74*, respectivament). Des d'un principi es va pensar que a humans existia formant un heterotetràmer compost per dues subunitats de cada (Flores et al. 1990), però experiments recents indiquen que les seves dues subunitats realment formen un heterodímer en solució i probablement també al PIC (Akashi et al. 2008). És de destacar que RAP30 conté regions d'homologia amb el factor σ de bacteris (McCracken et al. 1991).

TFIIF juga diversos papers durant la formació del PIC. En una fase inicial, TFIIF s'associa estretament a l'RNA Pol II i aquesta associació facilita la unió a un complex TFIID-TFIIB-TFIIA prèviament unit al promotor (Flores et al. 1991). A més, TFIIF serveix com a factor estabilitzador que augmenta l'afinitat de l'RNA Pol II pels complexos promotor-TBP-TFIIB, aportant diversos llocs d'interacció tant amb el DNA com amb els altres factors i, probablement, provocant canvis en la topologia del DNA que fan que es disposi embolicant l'RNA Pol II (Robert et al. 1998). Després de que l'RNA Pol II entri al promotor, TFIIF també és necessari per a reclutar TFIIE i TFIIH al PIC (Orphanides et al. 1996). Això s'havia suposat per interaccions directes amb TFIIE (Maxon et al. 1994), però aquesta suposada interacció TFIIF-TFIIE s'ha vist recentment que probablement no es dona (Akashi et al. 2008). TFIIF també juga un paper en la selecció del TSS, juntament amb l'RNA Pol II i TFIIB (Ghazy et al. 2004) gràcies a la proximitat tant de TFIIB com de TFIIF, al centre actiu de l'RNA Pol II (Bushnell et al. 2004; Cramer 2004; Freire-Picos et al. 2005). Així mateix també facilita el desacoblament de l'RNA Pol II del promotor (*promoter clearance* o *promoter escape*) (Yan et al. 1999) i augmenta l'eficiència d'elongació de l'RNA Pol II (Price et al. 1989; Shilatifard et al. 2003). Per últim, també s'ha vist que, de forma similar al factor σ de bacteris, TFIIF augmenta l'especificitat i eficiència de l'RNA Pol II evitant iniciacions inespecífiques inhibint o desfent unions de l'RNA Pol II a zones no promotores (Hampsey 1998).

A.2.3.2.E. TFIIE

Com ja hem vist, el pas següent a la formació del complex TFIID-TFIIB-RNA Pol II/TFIIF-promotor és el reclutament dels GTFs TFIIE i TFIIH. Al igual que per a TFIIF, es pensava que TFIIE

INTRODUCCIÓ

existia formant un heterotetràmer compost per dues subunitats α i dues β (Ohkuma et al. 1990b). De la mateixa manera que per TFIIF, recentment s'ha vist que en realitat està format en solució per un heterodímer de les seves dues subunitats, tot i que encara no està clar si podrien ser dos d'aquests heterodímers els que es trobessin formant part del PIC (Itoh et al. 2005; Akashi et al. 2008). Les seves dues subunitats tenen regions d'homologia amb certs factors σ de bacteris (Ohkuma et al. 1995; Okamoto et al. 1998) i també s'ha trobat homòleg de TFIIE a arqueas (anomenat TFE), amb certa homologia amb la meitat N-terminal de TFIIE α (Bell et al. 2001).

Una de les funcions principals de TFIIE és formar, juntament amb altres GTFs i l'RNA Pol II, un PIC capaç de reclutar TFIIH. De fet, es coneix que TFIIE pot regular TFIIH (Maxon et al. 1994; Watanabe et al. 2003). Així, s'ha vist que TFIIE estimula les activitats ATPasa, CTD quinasa, i DNA helicasa de TFIIH. Sembla que tant TFIIE com TFIIH són necessaris per obrir el DNA, ja que tots dos elements són prescindibles per a la transcripció d'un DNA prèviament obert. A favor d'això està la capacitat de TFIIE per unir ssDNA (*single strand DNA*) (Thomas et al. 2006).

A.2.3.2.F. TFIIH

És un complex multiproteic compost per deu subunitats, que a humans s'anomenen: p80/XPD (gen defectiu en pacients amb *xeroderma pygmentosum*, grup de complementació D), p89/XPB (gen defectiu en pacients amb *xeroderma pygmentosum*, grup de complementació B), p62, p52, p44, p40/CDK7 (*cycline-dependent kinase 7*), p38/Ciclina H, p34, p32/MAT1 (Svejstrup et al. 1996) i la recentment descoberta a humans i llevats TFB5 (Giglia-Mari et al. 2004; Ranish et al. 2004). Té tres activitats enzimàtiques requerides per a la transcripció: ATPasa DNA-dependent (Conaway et al. 1989; Roy et al. 1994), helicasa ATP-dependent (Schaeffer et al. 1993; Drapkin et al. 1994) i CTD quinasa (Feaver et al. 1991; Lu et al. 1992; Serizawa et al. 1992). A més d'aquestes activitats, alguns components de TFIIH estan implicats en la resposta NER (*Nucleotide Excision Repair*) a dany al DNA (Schaeffer et al. 1993; Drapkin et al. 1994; Wang et al. 1995; Giglia-Mari et al. 2004; Ranish et al. 2004; Coin et al. 2006). Aquest fet, juntament amb l'observació de que els gens que es transcriuen són els que preferencialment es reparen (Bohr et al. 1985; Mellon et al. 1989) ha fet sorgir la idea de que la reparació del DNA podria estar acoblada a la transcripció. A més, també s'ha vist que podria jugar un paper en la transcripció per l'RNA Pol I (Hoogstraten et al. 2002; Iben et al. 2002).

L'activitat ATPasa DNA-dependent de TFIIH és necessària per a l'obertura estable del promotor i per a la formació del primer enllaç fosfodièster (Dvir et al. 1996; Holstege et al. 1996). Sense TFIIH, l'RNA Pol II tendeix a parar-se pròxima al promotor donant transcrits abortius. L'addició de TFIIH i ATP dóna com a resultat una baixada dràstica dels transcrits abortius, el que demostra el paper de TFIIH en el desacoblament de l'RNA Pol II del promotor (*promoter clearance*) (Dvir et al. 1997; Kumar et al. 1998). Així, mentre que l'activitat ATPasa és necessària per obrir el promotor, l'activitat helicasa de TFIIH és crucial pel *promoter clearance* (Kim et al. 2000; Lin et al. 2005) i està implicada en la reparació del DNA (Schaeffer et al. 1994).

L'activitat CTD quinasa de la subunitat CDK7 de TFIIH és l'encarregada de fosforilar el CTD de l'RNA Pol II a la Serina 5 (Trigon et al. 1998). Aquesta fosforilació possibilita el reclutament de la maquinària de *capping* del 5' de l'RNA (Cho et al. 1997; Komarnitsky et al. 2000) i està implicada en el *promoter clearance* (Schroeder et al. 2000). A més de fosforilar el CTD, TFIIH s'ha vist que pot fosforilar p53 (Lu et al. 1997) i PC4 (Kershner et al. 1998), entre d'altres factors.

A.2.3.3. Muntatge del PIC. Model seqüencial i model d'holoenzim

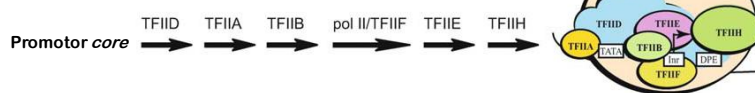
En base als estudis de transcripció *in vitro*, des d'un principi es va proposar que cada un dels components del PIC entra de forma ordenada i seqüencial (veure figura A.5.A): primer es dóna el reconeixement d'elements del promotor basal per TFIID, seguidament el reconeixement del TFIID-promotor per part de TFIIB, un cop format el complex TFIID-TFIIB-promotor, l'RNA Pol II unida ja a TFIIF pot ser reclutada al promotor. Finalment, la unió de TFIIE i TFIIH completa el PIC. Un cop el PIC està format, es separen les cadenes del DNA a la zona (*promoter melting*) i té lloc la síntesi del primer enllaç fosfodièster de la cadena d'RNA naixent. Finalment, l'RNA Pol II s'allibera del promotor (*promoter clearance*) i es dóna l'elongació de l'RNA. L'únic GTF no inclòs en aquest esquema és TFIIA, que podria unir-se al complex a qualsevol etapa posterior a l'entrada de TFIID, estabilitzant el complex (Orphanides et al. 1996).

Un model alternatiu per a la formació del PIC va sorgir quan certs laboratoris van aïllar components de la maquinària basal preensablats, el que va portar a pensar que l'RNA Pol II en realitat podia existir formant un gran holoenzim (Kim et al. 1994; Koleske et al. 1994). Aquest macrocomplex estaria format per l'RNA Pol II, certs GTFs, un gran número de factors que es van anomenar SRBs (per *supressors of RNA Polymerase B mutations*), molts dels quals es troben

INTRODUCCIÓ

també al complex *Mediator* (descriu com a necessari per a l'activació i també com a part de l'holoenzim (Kim et al. 1994)), a més d'altres proteïnes implicades en remodelació de la cromatina, processament de l'RNA i reparació del DNA (Ossipow et al. 1995; Maldonado et al. 1996; Yuryev et al. 1996; Cho et al. 1998). Tot i que la composició exacta de factors de cada holoenzim varia segons el mètode de purificació emprat, una característica que sembla comuna a tots és el fet que manquen de TFIID (Wu et al. 1998), cosa que situa a aquest GTF com element clau en el reconeixement del promotor i assimila el model d'holoenzim a la transcripció bacteriana, on una subunitat dissociable s'encarrega de reclutar al promotor el *core* de l'RNA Pol (veure figura A.5.B).

(A) Muntatge seqüencial del PIC



(B) Muntatge segons el model d'holoenzim

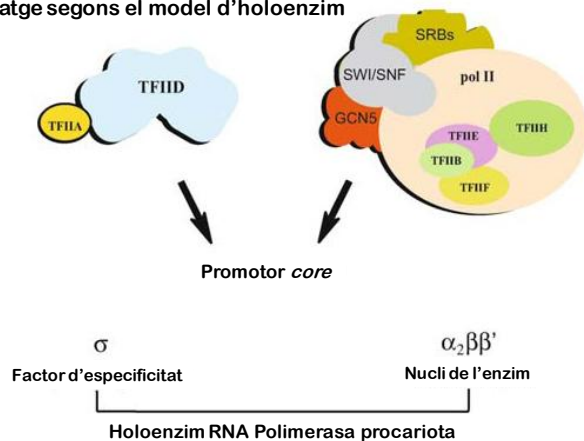


Figura A.5. Diferents models pel muntatge del PIC. La formació del PIC es pot donar per reclutament seqüencial, pas a pas, de la maquinària general de transcripció **(A)** o per reclutament d'un holoenzim de l'RNA Pol II preformat i el complex TFIID, com descriu el model de dos components de l'holoenzim **(B)**. El model d'holoenzim és similar al sistema de bacteris, en el qual una subunitat dissociable, el factor σ , dirigeix l'entrada del nucli de l'RNA Pol ($\alpha_2\beta\beta'$) al promotor (Thomas et al. 2006).

Si *in vivo* l'RNA Pol II existeix com a holoenzim o si es recluta de manera seqüencial com a entitat separada dels altres factors encara no ha estat definitivament aclarit. El fet que pugui existir com a holoenzim comporta certs avantatges i desavantatges. Un holoenzim seria, en teoria, més ràpid en respondre a certs estímuls i el fet que aportés múltiples superfícies de contacte pels activadors podria possibilitar que no fos tan dependent de la concentració cel·lular de cada un dels seus factors. D'altra banda, aquest macrocomplex preformat no s'enquadra bé amb l'existència d'una gran varietat de coreguladors diferents, ja que és més favorable que actuïn en múltiples etapes de la formació del PIC. Com a procés vital i complex que és, la formació del PIC hauria d'imposar múltiples punts de regulació per tenir la flexibilitat i capacitat de resposta que la cèl·lula requereix. Així, un model com el d'holoenzim és menys

eficaç responnent a les petites variacions de factors reguladors que tenen lloc a la cèl·lula. A més, el fet de que els complexos elongants siguin diferents dels PICs (Zawel et al. 1995) suggereix que haurien d'existir holoenzims addicionals a cada promotor per a les subsegüents reiniciacions, el que sembla inconsistent amb la significativa major quantitat d'RNA Pol II respecte dels altres factors (Kimura et al. 1999). Aquestes i altres evidències han portat a l'opinió actual de que és probable que tots dos mecanismes existeixin *in vivo*. Depenent de les molècules senyalitzadores involucrades en el context del promotor, un o altre mecanisme podria utilitzar-se. Així, mentre alguns elements sí podrien existir com a complexos fortament lligats, com poden ser per exemple les subunitats de la pròpia RNA Pol II, altres actuarien com a mòduls, podent-se associar segons els requeriments de cada cas.

A.2.4. Fàbriques de transcripció

Certs estudis citològics suggereixen que alguns coreguladors i components de la maquinària basal podrien estar localitzats al nucli a llocs discrets (Reyes et al. 1997). Tenint en compte la concentració limitant de certs factors de transcripció així com l'important paper que juga l'estructura de la cromatina en la regulació gènica, seria possible que l'organització subnuclear de la maquinària de transcripció tingués un paper fonamental en la regulació gènica.

És ben conegut que l'RNA Pol I es concentra majoritàriament al nuclèol per realitzar la seva funció, transcriure rRNAs (RNAs ribosomals), i dins d'aquest es disposa d'una manera determinada (Hozak et al. 1994). Així, el fet que l'RNA Pol I tingués una disposició determinada dins el nuclèol va fer sorgir la qüestió de si les formes actives de les altres dues RNA Pol, presents al nucleoplasma, es disposen també en llocs definits dins el nucli. Anàlisis a cèl·lules HeLa han revelat que els aproximadament 75000 transcrits d'RNA Pol II i III que aquest tipus cel·lular té en un moment donat es troben formant uns 10000 clústers de distribució no homogènia al nucli. Uns 8000 corresponen a agrupacions d'RNA Pol II i uns 2000 serien agrupacions d'RNA Pol III. Es calcula que per aquestes cèl·lules l'RNA Pol II es responsable d'uns 65000 transcrits i l'RNA Pol III d'uns 10000. Així, per a cada un d'aquests clústers d'RNA Pol II s'estarien produint uns 8 transcrits de mitjana, i uns 5 pels d'RNA Pol III. A més s'ha vist que les RNA Pol II i III actives estan espacialment separades (Pombo et al. 1999). Aquests clústers de transcripció activa s'han anomenat "fàbriques de transcripció" i pel cas d'RNA Pol II mesurarien uns 80 nm de diàmetre (Jackson et al. 1998).

INTRODUCCIÓ

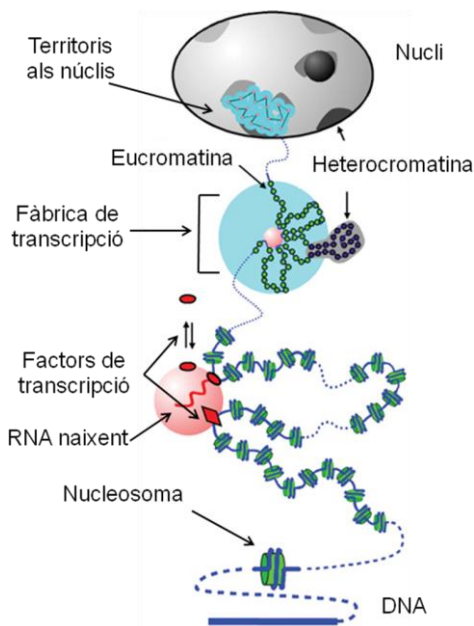


Figura A.6. Model per a una “fàbrica de transcripció” de l’RNA Pol II i la seva estructura (Faro-Trindade et al. 2006).

Al contrari del que s’ha pensat durant molt de temps, sembla que no és el promotor qui recluta a la maquinària basal, sinó més aviat al contrari. Per mitjà d’experiments de 3D-FISH (*three-dimensional Fluorescent in situ hybridization*), immuno-fluorescència i 3C (*Chromosome Conformation Capture*) es va veure que els gens eren dinàmicament reclutats als llocs actius en transcripció, no a la inversa, i que molts gens poden moure’s dins i fora d’aquests llocs, resultant en la seva activació i inactivació (Osborne et al. 2004).

D’acord amb aquest model de “fàbriques de transcripció”, diverses RNA Pol II s’associarien amb una certa quantitat de factors de transcripció que

quedarien embolcallats per una mena de “núvol” format per fins a 20 *loops* (voltes) de DNA/cromatina (veure figura A.6) (Faro-Trindade et al. 2006).

L’RNA Pol II seria un element immòbil d’aquestes fàbriques i el DNA es mouria a mesura que avancés la transcripció, fent que els *loops* apareguessin i desapareguessin. Cada fàbrica contindria un únic tipus d’RNA Pol i podrien estar enriquides en alguns factors de transcripció implicats en la transcripció específica de grups de gens, augmentant així l’eficiència de transcripció, ja que la concentració local de factors seria major. Els gens propers a una fàbrica és més probable que es transcrivin. Al acabar una transcripció, els gens estarien propers a una fàbrica i amb les modificacions a la cromatina necessàries per ser transcrits, de manera que es poden donar reiniciacions eficients (Schneider et al. 2007). Tot i que aquest model encaixa amb el fet que els gens actius s’agrupin transientment (Simonis et al. 2006), encara no hi ha evidències bioquímiques de l’existència d’aquestes “fàbriques de transcripció” com podrien ser el seu aïllament i caracterització funcional.

A.2.5. Etapes del procés de transcripció

De la mateixa manera que el cicle cel·lular es pot dividir en múltiples etapes, la transcripció també és un procés que consta de diverses fases. La transcripció, com és obvi,

ocorre en el context de la cromatina, i la primera etapa d'aquest procés consisteix precisament en el reclutament per part de la maquinària de transcripció de factors remodeladors de la cromatina, tals com els enzims que modifiquen les histones i la maquinària de remodelació de la cromatina dependent d'ATP (Workman et al. 1998). Com pot la maquinària de transcripció reconèixer els gens a transcriure dins el context de la cromatina, abans de que l'estructura d'aquesta hagi estat relaxada? Aquesta qüestió no està encara del tot resolta, però certs estudis demostren que alguns factors activadors de la transcripció poden reconèixer els seus llocs diana, inclús a zones de cromatina compactada, promovent la remodelació de la cromatina a aquelles zones (Eissenberg 2001; Cuesta et al. 2007). Així, podria ser necessària la presència d'aquests factors de transcripció anomenats "pioners", que possibilitarien l'obertura inicial de la zona a transcriure, per a que la maquinària transcripcional pogués posteriorment dur a terme la seva funció. Seguidament a la remodelació de la cromatina en el context del promotor, es pot muntar el complex de preiniciació i, tot seguit, la Polimerasa realitza la síntesi del primer enllaç fosfodièster de la cadena d'RNA naixent (iniciació). Després d'això, el procés de transcripció ja es pot considerar començat, però encara queden algunes tasques per a l'RNA Pol II i els seus cofactors. Primer la Polimerasa necessita escapar a les unions que l'ancoren al promotor, per a poder ser processiva en la producció del pre-mRNA (*promoter clearance*). A això segueix la fase d'elongació, que necessita salvar diversos obstacles, com la pròpia estructura de la cromatina de la regió codificant o proteïnes associades a ella, per no restar aturada períodes llargs. A la vegada que es va produint el transcrit, aquest va madurant i es dona el "*capping*" de l'extrem 5', que té lloc a l'inici de la transcripció, i l'*splicing*. Aquests processos, juntament amb l'addició de la cua de poli-A (poli-adenines), ocorren cotranscripcionalment (Bentley 2002; Proudfoot et al. 2002). Finalment termina la transcripció, i la polimerasa es recicla per a següents iniciacions.

A.2.5.1. Iniciació de la transcripció: *promoter clearance*

Com s'acaba de comentar, un cop el PIC està muntat i s'ha donat la síntesi dels primers enllaços fosfodièster, la polimerasa necessita escapar de les unions que la lliguen al promotor per a poder continuar elongant. Aquest procés s'anomena *promoter clearance* ("neteja" del promotor) i l'estat de fosforilació del CTD de la subunitat major de l'RNA Pol II hi juga un paper important. Com ja s'ha vist anteriorment (veure apartat A.2.3.1), l'RNA Pol II hipofosforilada és la que entra al promotor, mentre que una forma hiperfosforilada és la que es troba elongant

INTRODUCCIÓ

(Dahmus 1996). És poc conegut el mecanisme pel qual el CTD regula el *promoter clearance* i per tant defineix la transició d'iniciació a elongació. Com ja s'ha comentat, *in vivo* l'RNA Pol II als promotors es troba associada al complex *Mediator* (Kim et al. 1994). Almenys en part, aquesta associació és a través del CTD, de qui estimula varis centenars de vegades la fosforilació duta a terme per TFIIH (Kim et al. 1994; Naar et al. 2002). Encara que el *Mediator* forma un complex molt estable amb l'RNA Pol II, aquest és també un complex molt dinàmic, permetent al *Mediator* intercanviar-se d'una molècula d'RNA Pol II a una altra ràpidament *in vitro* (Svejstrup et al. 1997; Yudkovsky et al. 2000). També s'ha vist que el *Mediator* no està present en els complexos d'elongació on l'RNA Pol II està hiperfosforilada (Svejstrup et al. 1997; Pokholok et al. 2002). Així, és possible que un dels majors efectes de la fosforilació del CTD sigui la disrupció de les interaccions RNA Pol II-*Mediator*, permetent que aquest sigui reciclat per a la iniciació amb noves RNA Pol II (Svejstrup et al. 1997). De la mateixa manera, la fosforilació al CTD pot desfer també les interaccions de l'RNA Pol II amb alguns GTFs no necessaris en elongació (Usheva et al. 1992), possibilitant així que restin al promotor per a noves reiniciacions. S'ha vist que la fosforilació de la Ser 5, efectuada per TFIIH, es dona majoritàriament a les zones pròximes al promotor i decreix amb l'avanç de l'elongació, mentre que la fosforilació a la Ser 2, duta a terme per P-TEFb (*Positive Transcription Elongation Factor b*), és baixa vora el promotor i augmenta a mesura que la transcripció s'acosta a l'extrem 3' (Cho et al. 2001). Per tant, sembla que la fosforilació a la Ser 5 seria la implicada en el *promoter clearance*. De tota manera, mentre l'RNA Pol II desfà algunes interaccions, també en fa de noves amb factors necessaris pel processament i maduració de l'RNA naixent. Així, per a la interacció de l'RNA Pol II amb factors maduradors de l'RNA com els enzims responsables del "capping" a 5' (Cho et al. 1997; Yue et al. 1997; Rodriguez et al. 2000) o altres que actuen més enllà del context de la iniciació de la transcripció (Proudfoot et al. 2002), es requereix de la fosforilació del CTD. A més, la fosforilació del CTD és important també per a la interacció amb altres factors que actuen durant l'elongació tals com Set 1 i Set 2, relacionats amb l'estat de metilació de les histones (Hampsey et al. 2003) i el complex *Elongator* (Elongador)(Otero et al. 1999).

A.2.5.2. PrPP i elongació

Durant molt de temps s'ha pensat que la formació del PIC era el pas més important en la regulació de la transcripció. Aquest dogma s'ha vist alterat recentment, sobretot per l'estudi

dels promotors *Heat Shock* de *Drosophila* i altres vírics i de mamífers, que han fet evident que la regulació posterior a la iniciació és molt més freqüent del que es creia. Un cop iniciada l'elongació, en certs casos, l'RNA Pol II és ràpidament subjecte d'una demora al procés d'elongació, que es coneix com a PrPP (*promoter-proximal pausing*) (veure figura A.7). Aquest fenomen va ser primerament observat pels gens *Heat Shock* de *Drosophila* (Rougvie et al. 1988), que en un estat de no inducció contenen l'RNA Pol II aturada pròxima al promotor i havent transcrit uns 25 nucleòtids. La transcripció només continua en condicions d'inducció. Actualment, gràcies als assajos massius, sabem que aquest és un mecanisme de control de l'expressió gènica molt estès, present a multitud més d'organismes, com *Drosophila*, mamífers i llevat, i molt més utilitzat del que es pensava en un principi (Radonjic et al. 2005; Guenther et al. 2007; Zeitlinger et al. 2007). Les funcions del fenomen de PrPP encara no estan del tot clares, però sembla que podria ser un mecanisme que “preparés” gens que poden ser requerits amb molta urgència per a una ràpida inducció en resposta a canvis ambientals sobtats. També hi ha indicacions de que podria ser requerit pel correcte “capping” de l'RNA (Pei et al. 2003), sent així un sistema per evitar cicles inútils de transcrits sense el *cap* a 5'. El PrPP està mediat per certs factors com DSIF (*DRB (5,6-dichloro-1-h-D-ribofuranosylbenzimidazole) sensitivity-inducing factor*) i NELF (*Negative Elongation Factor*). DSIF consta dels factors d'elongació Spt4 i Spt5, presents des de llevats fins a humans. NELF comprèn quatre subunitats (A, B, C/D i E) i està conservat entre humans i *Drosophila*, però no està present a *C. elegans*, *S. cerevisiae* o *A. thaliana* (Saunders et al. 2006). Concretament NELF sembla estretament implicat amb el PrPP a *Drosophila*. L'eliminació de NELF per mitjà d'RNA d'interferència redueix el número de gens pausats a menys de la meitat (Muse et al. 2007). A més, un recent estudi massiu, també a *Drosophila*, situa NELF a la majoria dels gens amb l'RNA Pol II pausada (Lee et al. 2008). NELF, com el seu nom indica, s'ha vist que inhibeix l'elongació i per tant s'ha assumit que és un factor repressor de la transcripció. Sorprenentment, en l'estudi de Lee i col·laboradors es veu com la majoria dels gens on es situa NELF són gens d'alta expressió, indicant que NELF podria no ser obligatòriament repressor i el PrPP mediat per NELF podria ser una etapa necessària més, sovint transitòria, del cicle de transcripció. En aquest mateix treball també troben GAGA (objecte d'estudi del nostre grup) freqüentment relacionat amb NELF i el PrPP de l'RNA Pol II. Entre els factors que aboleixen el PrPP provocat per NELF i DSIF es troben TFIIS i P-TEFb. P-TEFb consisteix en CDK9 (*Cyclin-dependent-Kinase 9*) i la Ciclina T. La CDK9 és la responsable de la fosforilació de

INTRODUCCIÓ

la Serina 2 del CTD de la subunitat major de l'RNA Pol II, implicada en el control de l'elongació, el processament del pre-mRNA i la terminació de la transcripció (Phatnani et al. 2006). A la vegada, CDK9 fosforila NELF i DSIF. Així, aquesta fosforilació possibilita la dissociació de NELF i converteix a DSIF en un factor positiu de l'elongació (Peterlin et al. 2006).

L'elongació de la transcripció dista molt de ser un procés continu. L'eficiència d'elongació es veu incrementada per una sèrie de factors anomenats GEFs (*General Elongation Factors*), tals com TFIIIS, Elongina, el complex Elongador o el propi TFIIIF. S'ha vist que l'RNA Pol II, fins i tot en condicions *in vitro*, durant l'elongació pot arribar a detenir-se nombroses vegades i inclús pot tornar enrere alguns nucleòtids. Els GEFs funcionarien evitant aquestes fenòmens, fet que augmenta l'eficiència d'elongació de l'RNA Pol II (Svejstrup 2007).

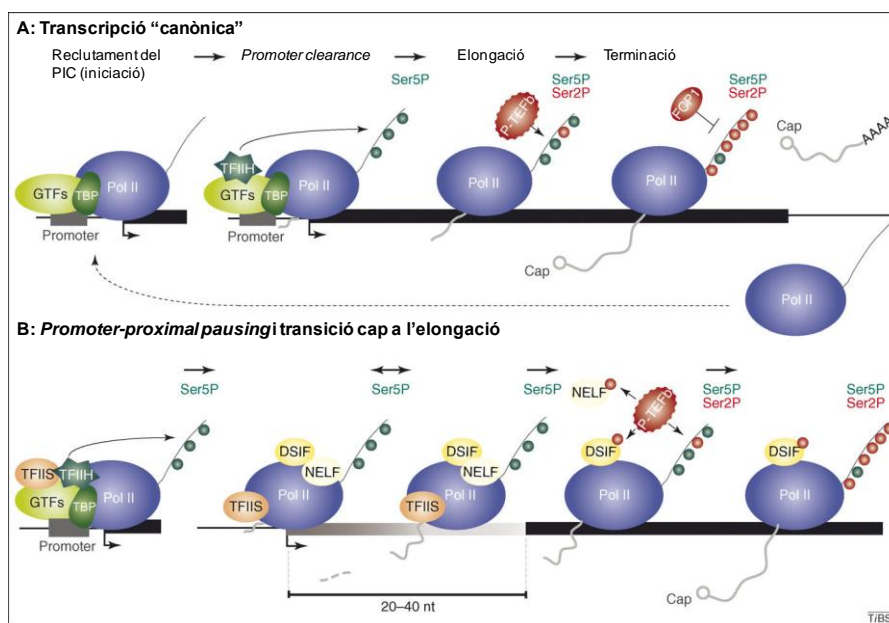


Figura A.7. Resum del procés de transcripció per RNA Pol II i del fenomen de promoter-proximal pausing (PrPP). Mentre que les etapes que comprenen el procés de transcripció "canònic" estan relativament ben estudiades, els fets que es donen durant el PrPP són menys coneguts i aparentment involucren diversos factors. **(A)** La transcripció s'inicia pel reclutament del PIC, el qual conté GTFs (en verd clar) i l'RNA Pol II (en blau), al promotor per la TBP (en verd fosc). Diverses quinases incloent les de TFIIH (verd fosc amb forma d'estrella) poden fosforilar la Ser 5 (cercles verds) del CTD per facilitar la sortida de l'RNA Pol II del promotor (*promoter clearance*). L'elongació es caracteritza per la fosforilació a la Ser 2 (cercles vermells) del CTD per P-TEFb (cercle vermell de forma rugosa). Un cop s'arriba al senyal de terminació, el CTD es desfosforila completament per la fosfatasa Fcp1 (en vermell). El mRNA, amb "cap" i poliadenilació s'allibera i l'RNA Pol II pot ser reciclada. Com s'indica, el nivell de fosforilació a la Ser 5 del CTD disminueix mentre que el de fosforilació a la Ser 2 augmenta, a mesura que la transcripció arriba a l'extrem 3'. **(B)** Durant el PrPP l'RNA Pol II resta aturada havent portat a terme la transcripció d'uns 20-40 nucleòtids de l'RNA naixent, esperant els senyals que permetin la fosforilació de la Ser 2 del CTD. Aquest procés està regulat per DSIF (en groc) i NELF (en color crema). TFIIIS ajuda evitant altres pauses, retraccions i inclús detencions de l'RNA Pol II, que es poden donar durant tota l'elongació. Durant la sortida de l'estat de PrPP, P-TEFb fosforila DSIF (que es torna activador de l'elongació), la Ser 2 del CTD i segurament NELF, que a resultes de la fosforilació es dissocia. Després d'això pot continuar l'elongació (Koch et al. 2008).

A.2.5.3. Terminació

La terminació de la transcripció encara avui és un fenomen poc conegut. Existeixen dos models per explicar la terminació de la transcripció (veure figura A.8) i la opinió actual és que no existeix una única manera de terminar la transcripció, depenent del gen en concret es pot utilitzar una o altra (Buratowski 2005). El que sí sembla clar és la relació entre la fosforilació de la Ser 2 per P-TEFb i el reclutament de la maquinària d'addició de la cua Poli-A (Ahn et al. 2004; Ni et al. 2004).

Finalment, un cop alliberada del DNA motlle, l'RNA Pol II necessita ser desfosforilada per poder ser utilitzada en un altre cicle de transcripció. Probablement en això tinguin un paper les fosfatases específiques del CTD: Ssu72 per a la Ser 5 (Krishnamurthy et al. 2004) i Fcp1 per a la Ser 2 (Cho et al. 2001). A més, de la mateixa manera que necessita desfer-se de certs GTFs a la fase de *promoter clearance*, probablement també necessiti desfer les unions amb factors d'elongació. Tot i això, els mecanismes pels quals ocorre encara estan per descriure.

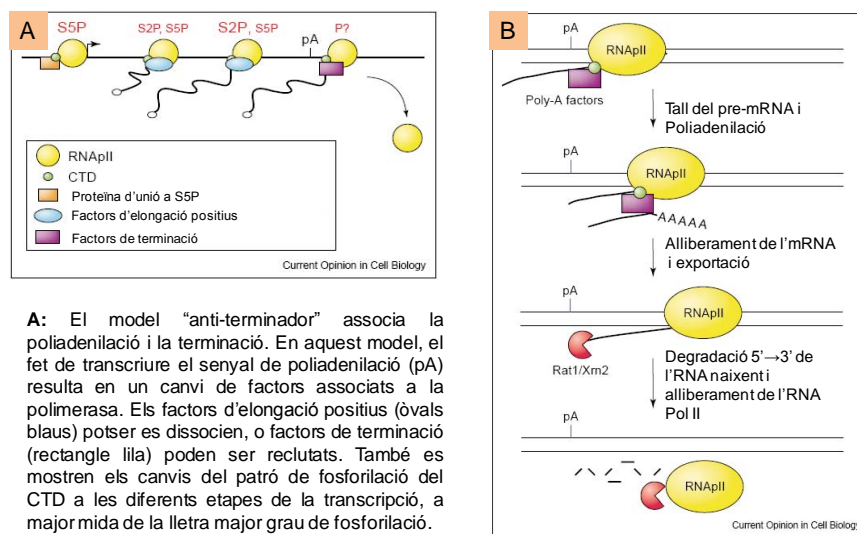


Figura A.8. Models per a la terminació de la transcripció (Buratowski 2005).

A.2.6. Paper de la cromatina en la regulació de la transcripció

La cromatina, definida com a DNA cromosòmic associat a proteïnes, juga un paper integral en la regulació de l'expressió gènica. L'empaquetament del DNA als nucleosomes

INTRODUCCIÓ

sembla afectar totes les etapes de la transcripció, des de la unió de factors activadors i la formació del PIC fins a l'elongació. El nucli o *core* del nucleosoma està format per aproximadament 147 pb de DNA embolicat formant 1,65 voltes al voltant de l'octàmer d'histones. Així organitzat, el DNA té 14 punts de contacte amb les histones (Luger et al. 1997). Això fa que el nucleosoma sigui un dels complexos DNA-proteïna més estables que hi ha, el que el fa ideal per les seves funcions de compactació. Però el nucleosoma no és una unitat estàtica. Posseeix propietats dinàmiques altament regulades per diferents complexos proteïcs.

La posició dels nucleosomes respecte les seqüències de reconeixement de factors de transcripció és crítica per a determinar si un gen s'expressa. Per exemple, el posicionament de nucleosomes sobre seqüències TATA *box* o al TSS inhibeix la transcripció (Martinez-Campa et al. 2004). Així, la disposició dels nucleosomes regula la disponibilitat del DNA per a factors que actuen en *trans*.

Tradicionalment es diu que els genomes eucariotes estan organitzats en dos tipus de cromatina: eucromatina i heterocromatina. L'heterocromatina constitueix una part important del genoma de la majoria dels organismes i acostuma a situar-se a les regions centrals i telomèriques dels cromosomes, on el DNA de còpia única està escassament representat i abunden les seqüències de DNA repetitiu. Posseeix una mínima o nul·la capacitat de recombinació meiótica, és de replicació tardana i transcripcionalment repressiva. Les seqüències heterocromàtiques estan empaquetades en nucleosomes regularment disposats i són relativament resistents a les digestions amb nucleases. Generalment, els nucleosomes de regions heterocromàtiques contenen histones hipoacetilades i metilades a la lisina 9 de la histona H3 (H3K9me) i a la lisina 20 de la histona H4 (H4K20me), característiques que generalment s'associen a silenciament gènic (Schulze et al. 2007). Tot i que l'heterocromatina conté una densitat gènica molt inferior a l'eucromatina, a *Drosophila* hi ha gens actius que resideixen a zones heterocromàtiques, alguns essencials pel desenvolupament (Sinclair et al. 2000). Per contra, l'eucromatina es condensa i descondensa amb el cicle cel·lular. Comprèn la majoria de les regions entre centròmers i telòmers, conté la gran majoria de gens de còpia única i realitza recombinació meiótica. Les zones d'eucromatina habitualment contenen un patró irregular de nucleosomes. Generalment abunden les marques d'activació transcripcional com l'acetilació de les cues N-terminals de les histones H3 i H4 i la metilació de la lisina 4 de la histona H3 (H3K4me) (Schulze et al. 2007).

Taula A.4. Modificacions de les histones relacionades amb transcripció (Li et al. 2007).

Modificacions	Posició	Enzims				Mòduls que reconeixen les modificacions *	Funcions en transcripció
		<i>S. cerevisiae</i>	<i>S. pombe</i>	<i>Drosophila</i>	Mamífers		
Metilació a lisines	H3K4	Set1	Set1	Trx, Ash1	MLL, ALL-1, Set9/7, ALR-1/2, ALR, Set1	PHD, Chromo, WD-40	Activació
	H3K9	n/a	Clr4	Su(var)3-9, Ash1	Suv39h, G9a, Eu-HMTase I, ESET, SETBD1	Chromo (HP1)	Repressió, activació
	H3K27				E(Z)	Ezh2, G9a	Repressió
	H3K36	Set2			HYPB, Smyd2, NSD1	Chromo(Eaf3), JMJD	Reclutament de Rpd3S que reprimeix la iniciació interna
	H3K79	Dot1			Dot1L	Tudor	Activació
	H4K20		Set9	PR-Set7, Ash1	PR-Set7, SET8	Tudor	Silenciament
	H3R2				CARM1		Activació
Metilació a arginines	H3R17				CARM1		Activació
	H3R26				CARM1		Activació
	H4R3				PRMT1	(p300)	Activació
	H3S10	Snf1				(Gcn5)	Activació
Ubiquitinació	H2BK120/123	Rad6, Bre1	Rad6		UbcH6, RNF20/40	(COMPASS)	Activació
	H2AK119				hPRC1L		Repressió
Acetilació	H3K56					(Swi/Snf)	Activació
	H4K16	Sas2, NuA4		dMOF	hMOF	Bromodomini	Activació
	Htz1K14	NuA4, SAGA					Activació

* Les proteïnes indicades entre parèntesis es sap que reconeixen les modificacions corresponents, però els dominis específics responsables encara estan per determinar.

Les modificacions postraduccionalment de les histones, tant a les cues com als dominis globulars, estan estretament lligades a la regulació gènica. Aquestes modificacions inclouen metilacions en residus arginina (R); metilació, acetilació, ubiquitinació, ADP-ribosilació i sumoilació de residus lisina (K); i fosforilació de serines (S) i treonines (T) (veure taula A.4). La majoria de les modificacions es distribueixen en patrons concrets a diverses regions gèniques tals com les seqüències *upstream*, el promotor basal, el 5' de l'ORF (de l'anglès *Open Reading Frame*, pauta oberta de lectura) i el 3' de l'ORF (veure figura A.9). De fet, la localització de les modificacions és quelcom molt regulat i és crucial per a la seva funció en transcripció (Li et al. 2007). Típicament, l'acetilació de les histones ocorre en múltiples residus i és portada a terme per una varietat de complexos acetiltransferasa d'histones (HAT, de l'anglès *Histone Acetyltransferase*) (Brown et al. 2000). S'ha proposat que els diferents patrons d'acetilació de les lisines de les histones podrien estar relacionats amb diverses funcions, com la regulació de

INTRODUCCIÓ

gens biològicament relacionats i coexpressats (Kurdistani et al. 2004). Tot i això, altres observacions indiquen que la funció de l'acetilació de les histones depèn del número de lisines

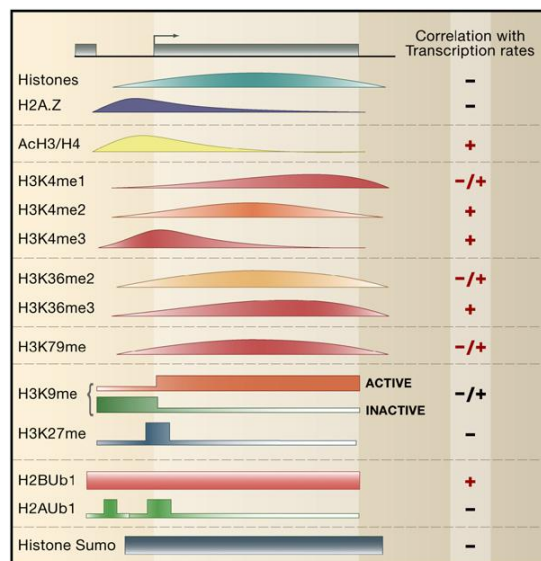


Figura A.9. Distribució de les diferents modificacions de les histones al llarg de la regió gènica i el seu efecte en transcripció. Els patrons en corba indiquen que les dades han estat obtingudes per estudis genòmics globals. Els patrons en rectangle indiquen que les dades s'han obtingut d'uns quants casos estudiats. Amb l'excepció de la metilació de la K9 i K27, la majoria de la informació està basada en gens de llevat (Li et al. 2007).

modificades, per efecte acumulatiu, amb la coneguda excepció de l'acetilació de la lisina 16 de la histona 4 (H4K16Ac) (Dion et al. 2005). A diferència de l'acetilació, la resta de modificacions normalment estan catalitzades per un enzim concret, a un lloc específic i amb funcions úniques. La diferència entre l'acetilació i la resta de modificacions encara no està clara, però els canvis biofísics que confereix l'acetilació podrien oferir una explicació parcial (Li et al. 2007). Les modificacions de les histones tenen múltiples papers en la regulació gènica, que inclouen distorsionar l'estructura dels nucleosomes reduint les interaccions DNA-histones (Reinke et al. 2003), ser diana de reconeixement per diferents factors (Strahl et al. 2000) i afectar a les estructures d'ordre superior de la cromatina (Shogren-Knaak et al. 2006).

A més dels factors que modifiquen postraduccionalment les histones existeixen altres que utilitzen l'energia de l'hidròlisi de l'ATP per alterar les interaccions DNA-histones. Es coneixen com a factors remodeladors de la cromatina. A conseqüència de l'acció d'aquests factors els nucleosomes es poden moure per la cadena de DNA, transitòriament el DNA pot desenrotllar-se de l'octàmer d'histones o inclús les histones poden ser desacobrades de l'estructura del nucleosoma. Tots aquests efectes canvien l'accessibilitat del DNA per a factors de transcripció (Saha et al. 2006). Els factors de transcripció reconeixen seqüències del DNA on s'uneixen per fer la seva funció. Però quan aquestes seqüències estan immerses dins l'estructura de la cromatina pot ser necessària l'acció dels factors remodeladors de cromatina per a que aquesta unió tingui lloc. D'acord amb aquest fet, estudis realitzats a nivell genòmic indiquen que a la regió promotora dels gens s'estableix una zona lliure de nucleosomes, flanquejada a ambdós cantons per nucleosomes posicionats (Yuan et al. 2005).

A banda dels factors que poden alterar les histones “canòniques” tradicionalment considerades com a integrants dels nucleosomes, existeixen altres variants de les histones que també juguen un paper en la regulació gènica. Una característica que les distingeix de les histones “canòniques” és que són sintetitzades fora de la fase S i s’incorporen al DNA de manera independent de la replicació, ja sigui per acció de complexos remodeladors de cromatina dependents d’ATP que les intercanvien amb les histones “canòniques” (Mizuguchi et al. 2004), ja sigui per acció de chaperones (Park et al. 2005) o de manera dependent de la transcripció (Schwartz et al. 2005).

Els nucleosomes posicionats a la regió codificant formen una forta barrera per a l’avanç de l’RNA Pol II. L’estructura completa de l’octàmer d’histones impedeix la progressió de l’RNA Pol II i s’ha vist que aquesta necessita com a mínim que un dímer H2A/H2B es perdi (Kireeva et al. 2002). Com ja s’ha comentat anteriorment, TFIIIS estimula l’elongació i ho fa precisament facilitant la transcripció a través dels nucleosomes (Kireeva et al. 2005). Un altre factor implicat en ajudar la polimerasa a avançar a través dels nucleosomes és el complex amb funció de chaperona d’histones FACT (*facilitates chromatin transcription*) que actua en el desacoblament i acoblament dels dímers H2A/H2B, necessaris per a la progressió de l’RNA Pol II (Belotserkovskaya et al. 2003).

En relació amb la transcripció i l’estat de la cromatina cal esmentar també el fenomen conegut com a *position effect variegation* (PEV), descobert a *Drosophila*, i que també hi ha evidències de que podria existir a humans (Norwood et al. 2004). Consisteix en diferències, entre cèl·lules d’un mateix teixit, en els nivells d’expressió de gens situats molt propers a zones heterocromàtiques o dins d’elles. Aquest fenomen aporta una sòlida prova de que les fronteres entre eucromatina i heterocromatina no són quelcom fix. Els descobriments a llevats sobre com es formen les zones heterocromàtiques han aportat una valuosa llum sobre aquest tema (Volpe et al. 2002). Mitjançant la maquinària relacionada amb l’RNA d’interferència es recluten a les zones de DNA repetitiu factors com la metiltransferasa d’histones Clr4 (SU(VAR)3-9 a *Drosophila* per *Suppressor of PEV 3-9*) que metila la histona H3 a la lisina 9, fet que constitueix una marca d’heterocromatina. Aquesta marca genera llocs d’unió per a Swi6 (HP1 a *Drosophila*, per *Heterochromatin Protein 1*) que a la vegada recluta més molècules de Clr4, estenent la formació de l’heterocromatina (*spreading* de l’heterocromatina). Aquesta extensió de la cromatina no és igual entre cèl·lules i és el que explica el PEV. En base a aquest fenomen s’han caracteritzat

INTRODUCCIÓ

multitud de factors com a Su(var) (*Suppressors of variegation*), que generalment actuen produint un estat silenciats de la cromatina del tipus existent a l'heterocromatina centromèrica, i E(var) (*Enhancers of variegation*) que funcionen generant cromatina activa que permet l'expressió gènica (Schulze et al. 2007).

A.2.7. Factors activadors/repressors de la transcripció

Com ja s'ha comentat, la mínima maquinària necessària per a una correcta transcripció *in vitro* consisteix en l'RNA Pol II i els diferents GTFs. A la cèl·lula però, la transcripció és un fenomen molt més complex i altament regulat que requereix de l'acció col·lectiva de factors específics de seqüència, anomenats factors de transcripció (FTs), juntament amb la maquinària basal (RNA Pol II i GTFs) i tota una sèrie de coreguladors que els connecten, a més de tot un seguit de factors remodeladors de la cromatina que mobilitzen els nucleosomes de la zona del promotor i d'una varietat d'enzims que catalitzen la modificació covalent de les histones (tals com acetilació, metilació, fosforilació, etc.) i d'altres proteïnes. L'activació i la repressió de la transcripció esdevé clau a la transcripció *in vivo*, ja que l'estructura de la cromatina estableix un impediment a la transcripció que ha de ser superat, alhora que la repressió activa de certs gens és crucial per a la regulació fina de les funcions cel·lulars. A més, si tenim en compte que la regulació acurada de la transcripció capacita a la cèl·lula per respondre a estímuls i a la vegada és responsable de la gran diversitat cel·lular d'un organisme, podem entendre el gran nombre de proteïnes destinades a la regulació d'aquest procés. Per exemple, a humans s'estima que hi ha uns 3000 FTs, representant per tant el 10% dels gens totals (Babu et al. 2004).

A.2.7.1. Factors de transcripció (FTs)

Els FTs són proteïnes d'unió a DNA específiques de seqüència que tant es poden unir a la zona del promotor pròxima al TSS com a zones reguladores més distals, tals com les anomenades "*enhancers*" (augmentadors) i "*silenciadors*". La seva funció és interpretar i transmetre la informació continguda a la seqüència del DNA als factors i cofactors responsables de la transcripció. Són per tant l'interfície que comunica la informació reguladora codificada al genoma amb la maquinària de transcripció (Kadonaga 2004). Poden funcionar tant a l'inici de la transcripció com a etapes posteriors, tals com l'elongació (Blau et al. 1996).

Usualment tenen una estructura modular, on cada mòdul o domini és responsable d'una funció. Típicament contenen un mòdul d'unió a DNA unit a un o varis mòduls d'activació o repressió, així com poden tenir un mòdul de multimerització i un reguladori de la seva activitat. Aquests mòduls s'ha vist que poden mantenir la seva funcionalitat de forma independent de la resta de la proteïna, com demostren estudis de ja fa anys amb el factor de transcripció GAL4 de llevat o VP16 d'herpes virus (Brent et al. 1985; Sadowski et al. 1988).

Els FTs poden ser regulats per modificacions post-traduccionals com és el cas de la fosforilació de, per exemple, CREB (proteïna d'unió a l'element de resposta a l'AMP cíclic, de l'anglès *cyclic AMP response element binding protein*) que activa aquest factor de transcripció (Gonzalez et al. 1989). Altres exemples de modificació post-traduccional de FTs són: l'addició de N-acetilglucosamina O-enllaçada que s'ha relacionat amb la inactivació de diversos factors de transcripció (Yang et al. 2001; Yang et al. 2002); l'acetilació, com el cas de p53 on incrementa l'afinitat d'aquest factor pel DNA (Gu et al. 1997); ubiquitinació i sumoilació (Freiman et al. 2003). També és una característica comuna a molts FTs el poder unir-se a una gran varietat de seqüències, tot i que amb menys afinitat que a les seves seqüències de reconeixement específiques (Walter et al. 1994). A més, els FTs acostumen a reconèixer seqüències de 6-8 pb, que poden ser relativament freqüents. Això suposaria un problema d'especificitat si no fos pel fet que els FTs funcionen millor reconeixent clústers (agrupacions) de les seves seqüències de reconeixement, cosa que afegeix un grau més alt d'especificitat del que s'aconsegueix amb la unió d'un únic FT al DNA. A més, com normalment actuen de forma sinèrgica (Carey et al. 1990), la potència dels FTs es veu augmentada si a la regió reguladora dels gens hi ha aquests clústers, tals com les regions *enhancer*.

En general, es considera que els FTs actuen majoritàriament reclutant cofactors (coactivadors i corepressors) al DNA per mitjà d'interaccions proteïna-proteïna. Aquests cofactors funcionen directa i/o indirectament regulant l'activitat de la maquinària transcripcional de l'RNA Pol II al promotor basal (Kadonaga 2004). Tradicionalment s'ha parlat de que un FT concret és activador o repressor de la transcripció, però aquesta classificació sovint no és vàlida. Un activador pot actuar com a "antirepressor", és a dir, eliminant la repressió prèvia d'un promotor en lloc d'activant-lo realment, com s'ha descrit per exemple pel factor GAGA de *Drosophila* sobre el promotor de *Krüppel* (Kerrigan et al. 1991). També s'ha vist que molts poden actuar com a activadors d'uns gens i repressors d'uns altres, depenent de la

INTRODUCCIÓ

resta d'elements reguladors del promotor. Tal és el cas de la proteïna Dorsal de *Drosophila* que pot actuar com a activador, si recluta el coactivador CBP (*CREB binding protein*), o com a repressor, si recluta el corepressor general Groucho. Que interaccioni amb un o altre cofactor depèn bàsicament de la resta d'elements reguladors del promotor (Valentine et al. 1998).

A.2.7.2. Cofactors o coreguladors

Els cofactors de la transcripció formen un heterogeni grup de proteïnes que són reclutades als promotors per associació amb els factors de transcripció mitjançant unió directa o en el context de complexos multiproteics, i són requerits per a la regulació de l'expressió gènica (O'Malley 2007). La gran varietat de cofactors es podria classificar en dos grans grups, atenent a la seva manera d'actuar: els que funcionen com "adaptadors" entre els FTs i la maquinària basal o factors remodeladors de la cromatina; i els que per si mateixos tenen activitat remodeladora de la cromatina.

Entre els "adaptadors", alguns són components de la maquinària basal de transcripció. Tal és el cas d'alguns TAFs (Albright et al. 2000) i de TFIIA (veure apartat A.2.3.2.A). Altres cofactors, tot i no ser considerats GTFs, també estan estretament lligats a la maquinària basal de transcripció. En aquesta categoria es poden incloure cofactors com PC4 i NC2 (Roeder 1998). En aquesta mateixa categoria destaca especialment el complex multiproteic *Mediator*. Aquest complex no és important únicament per a la regulació de la transcripció, sinó que és absolutament necessari per a gairebé qualsevol transcripció de gairebé qualsevol gen d'RNA Pol II. Interacciona directament amb l'RNA Pol II i amb activadors, servint de pont entre ells i estimulant d'aquesta manera la transcripció. A més, el *Mediator* no és només necessari per a la regulació positiva de la transcripció, també és essencial per a la regulació negativa. Per tant, el *Mediator* és un coactivador, un corepressor i també un GTF, tot en un. Es podria considerar que la seva funció és la de transferir els senyals dels *enchancers* als promotors en tot el rang d'organismes, de llevats a humans (Kornberg 2007).

Un altre grup de cofactors que es podrien catalogar com "adaptadors" són els que estan associats a factors activadors o repressors i són reclutats per aquests a certs promotors, on actuen modulant l'ocupació del DNA o fent d'interfície amb la maquinària basal. Entre aquests hi ha factors com OCA-B (OBF-1/Bob-1), Groucho, CtBP (*C-terminal Binding Protein*) o Notch (Lemon et al. 2000). Cal destacar tant Groucho com CtBP (dCtBP a *Drosophila*) per la seva

importància en dos mecanismes de repressió àmpliament utilitzats: el mecanisme de repressió de llarg abast (*long-range*), mediat freqüentment per Groucho, i el de curt abast (*short-range*), mediat en molts casos per CtBP. El mecanisme de llarg abast està caracteritzat per la repressió de tot un *locus* cromosòmic, de 1 Kpb o més, i és independent de l'orientació dels llocs d'unió dels FTs que promouen aquest tipus de repressió. Per contra, el mecanisme de repressió de curt abast bloqueja l'acció d'activadors propers (fins a uns 100 pb), sense interferir en l'acció d'activadors distals (Courey et al. 2001). Cal tenir en compte que tot i aquesta classificació, existeixen exemples en que un repressor i el seu corepressor de llarg abast poden funcionar com si fossin de curt abast (Nibu et al. 2001).

Entre els cofactors remodeladors de la cromatina podem distingir dos tipus d'activitats: la remodelació ATP-dependent de l'estructura de la cromatina de manera no covalent i la modificació covalent de les cues de les histones. Entre els primers, a *Drosophila* destaquen dSWI/SNF (*switch mating type/sucrose non-fermenting*) i els de la família ISWI (*imitation switch*): NURF (*nucleosome remodeling factor*), ACF (*ATP-utilizing chromatin assembly and remodeling factor*) i CHRAC (*chromatin remodeling and assembly complex*), cadascun dels quals conté de 4 a 8 polipèptids. Complexos similars s'han trobat a llevats i vertebrats, tot i que amb substancials canvis en la composició de les seves subunitats (Kingston et al. 1999). Entre els complexos que modifiquen les histones, a *Drosophila* destaquen el complex amb activitat HAT (*histone acetyltransferase*) SAGA (*Spt-Ada-Gcn5-acetyltransferase*) (Brown et al. 2000), i dos tipus de HDACs (*histone deacetylases*): els complexos tipus Sin3 (*SWI-independent*) i els tipus Mi-2/NuRD (*Nucleosome remodeling deacetylase*) (Knoepfler et al. 1999). Aquests complexos estan conservats de llevats fins a vertebrats. L'activitat HAT està relacionada amb activació, i l'activitat contrària HDAC amb repressió. Tot i la importància de l'acetilació, existeixen altres tipus de modificacions de les histones, produïdes per cofactors, implicades en la regulació gènica, com per exemple la metilació de la lisina 9 de la histona H3 duta a terme pel factor SUV39 de mamífers (homòleg de Su(var)3-9 de *Drosophila* i Clr4 de llevat), que s'ha vist que té efectes repressius (Richards et al. 2002).

A.3. El factor GAGA de *Drosophila melanogaster*

El factor GAGA és una proteïna de *Drosophila* descoberta per la seva capacitat d'unir seqüències d(GA)_n i d(CT)_n als promotors d'*Ultrabithorax (Ubx)* i *Engrailed (en)* i activar la seva transcripció *in vitro* (Biggin et al. 1988; Soeller et al. 1988). Encara que les seqüències reguladores tipus GAGA han estat descrites en una gran varietat d'organismes com eriçó de mar, plantes i mamífers (Frustaci et al. 1995; Bevilacqua et al. 2000; Hagemann et al. 2000; Melfi et al. 2000; Sangwan et al. 2002; Volpi et al. 2002; Santi et al. 2003), on han estat extensivament estudiades és a *Drosophila melanogaster*. Molts gens de *Drosophila*, incloent-hi gens homeòtics i de desenvolupament, contenen seqüències GAGA a les seves regions reguladores (Wilkins et al. 1997; Sun et al. 2003; Negre et al. 2006). Per aquests motius, tot i que GAGA no és la única proteïna a *Drosophila* capaç d'unir aquestes seqüències (Lehmann et al. 1998), ha sigut objecte de molts estudis.

GAGA és una proteïna amb localització estrictament nuclear i d'expressió aparentment ubiqua (Soeller et al. 1993). S'ha vist associada tant a l'eucromatina com a l'heterocromatina. Així, per exemple, GAGA apareix unida a múltiples bandes al llarg dels braços eucromàtics dels cromosomes politènics de glàndules salivals (Tsukiyama et al. 1994; Benyajati et al. 1997). En canvi, durant la mitosi GAGA apareix unida a les regions heterocromàtiques riques en repeticions d(GA)_n dels cromosomes (Platero et al. 1998).

Des de que inicialment GAGA fos descrita com a factor activador de la transcripció (Biggin et al. 1988; Soeller et al. 1988), abundants estudis han anat demostrant que GAGA està implicada en multitud de processos. Així, avui dia es considera que GAGA és una proteïna polifacètica que pot regular l'expressió gènica a múltiples nivells i el seu efecte pot ser tant activador com repressor, depenent del context. Des d'un principi, GAGA ha generat un gran interès per la seva implicació al mecanisme de regulació dels gens *Heat Shock* (Lee et al. 1992b; Tsukiyama et al. 1994; Wilkins et al. 1997), així com dels gens homeòtics (Lehmann 2004), però és clar que aquestes no són les seves úniques funcions.

A.3.1. El gen *Trithorax-like*

La proteïna GAGA està codificada pel gen *Trithorax-like (Trl)* de *Drosophila melanogaster* (Farkas et al. 1994). És un gen de còpia única (Soeller et al. 1993) i amb efecte matern, ja que la

mare diposita a l'embrió una certa quantitat d'mRNA de *Trl* (Bhat et al. 1996). Des de la seva identificació en estudis genètics es va classificar com un gen del grup *Trithorax* (TrxG), ja que els mutants de *Trl* mostraven defectes a l'expressió de gens homeòtics, tals com *Ubx* (Lewis 1978; Farkas et al. 1994). Els mutants nuls per a *Trl*, com el *Trl^{R85}*, no aconsegueixen desenvolupar-se més enllà del tercer estadi larvari i presenten defectes en certs gens homeòtics com *Ubx* i *en*, però no en tots. A més, clons cel·lulars de pèrdua de funció demostren que, durant el desenvolupament larvari, GAGA no es requereix per a la regulació dels gens homeòtics (Bejarano et al. 2004). Els mutants hipomorfs per a *Trl*, com el *Trl^{13C}*, són els únics que arriben a adult. Aquests, a més de tenir una viabilitat i fertilitat reduïda, són *enhancers* de PEV (mostren un augment d'aquest fenomen), fet que concorda amb la capacitat de GAGA per remodelar la cromatina (Farkas et al. 1994). També presenten defectes a l'expressió de gens com *en* i *ftz* (*fushi tarazu*), i greus anomalies a les divisions nuclears embrionàries tals com asincronia, fallada en la condensació cromosòmica, segregació cromosòmica anormal i fragmentació cromosòmica (Bhat et al. 1996).

El gen *Trl* codifica per a dues isoformes de la proteïna GAGA, generades per *splicing* alternatiu: una de 581 aa anomenada GAGA₅₈₁ i una altra de 519 aa anomenada GAGA₅₁₉. Ambdues isoformes només es diferencien a l'extrem C-terminal, poden formar tant complexos homomèrics com heteromèrics entre elles, reconeixen les mateixes seqüències i comparteixen localització a cromosomes politènics (Benyajati et al. 1997). Tot i diferir precisament a la zona descrita com a domini de transactivació (Vaquero et al. 2000)(veure més endavant), totes dues isoformes es mostren igual de competents en activar la transcripció en assajos de transfecció transitòria a cèl·lules de *Drosophila* en cultiu (Benyajati et al. 1997). Només GAGA₅₁₉ es troba a les primeres 6h del desenvolupament embrionari, a partir d'aquest punt GAGA₅₈₁ comença a acumular-se, fins que a les 12h les quantitats s'equiparen i així continuen la resta del desenvolupament (Benyajati et al. 1997). Les funcions de les dues isoformes, tot i que àmpliament solapants, hi ha evidències de que no són exactament idèntiques. Així, mentre les dues isoformes, expressades per separat ectòpicament a mutants defectius per a GAGA, aconsegueixen rescatar el fenotip de letalitat, molts altres fenotips associats a la manca de GAGA no són rescatats en igual mesura per totes dues isoformes (Granok et al. 2001; Greenberg et al. 2001). Estudis comparatius a *D. virilis* mostren com en aquesta espècie GAGA també presenta dues isoformes: una de 556 aa amb un 76% d'identitat amb GAGA₅₁₉ de *D.*

INTRODUCCIÓ

melanogaster; i una de 590 aa amb un 85% d'identitat amb GAGA₅₈₁. A part de l'alt grau d'identitat de seqüència, l'estructura del gen *Trl* a *D. virilis* també està sorprenentment conservada, al igual que el patró de *splicing* que genera les dues isoformes (Lintermann et al. 1998). Aquest nivell de conservació entre dues espècies evolutivament tan allunyades, uns 60 milions d'anys de divergència segons alguns estudis (Beverley et al. 1984), també suggereix algun paper diferencial entre les dues isoformes.

A *Drosophila melanogaster*, *Trl* està situat al cromosoma 3 braç L. Com es pot veure a la figura A.10, GAGA₅₁₉ està formada per quatre exons mentre que GAGA₅₈₁ té un exó més, resultat de la utilització d'un lloc de *splicing* alternatiu dins l'últim exó de GAGA₅₁₉ (Lintermann et al. 1998). A la línia cel·lular S2 de *Drosophila* s'ha vist que la transcripció comença des d'almenys tres llocs diferents (Kosoy et al. 2002).

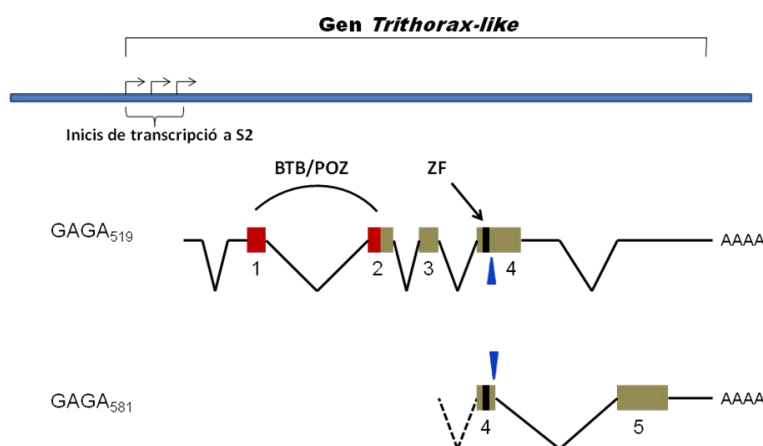


Figura A.10. Estructura genòmica del gen *Trl* de *D. melanogaster* i els seus diferents transcrits. Els números indiquen els diferents exons. En vermell estan els exons que codifiquen pel domini BTB/POZ, un triangle blau indica el lloc de *splicing* alternatiu que dona lloc a les dues isoformes i en negre la seqüència del dit de Zinc (ZF). GAGA₅₈₁ no s'ha representat sencera per simplificar (Lintermann et al. 1998).

A més d'aquestes dues isoformes generades per *splicing* alternatiu, la diversitat de proteïnes GAGA és encara major, com es pot inferir de la resolució electroforètica de la proteïna provinent d'extractes nuclears, tant en una dimensió (Biggin et al. 1988; Soeller et al. 1993) com en dues (Bonet et al. 2005). Aquesta gran varietat de proteïnes GAGA es pensa que vindria donada per la modificació post-traducciona d'ambdues isoformes. Així, fins avui s'ha descrit que GAGA pot estar O-glicosilada (Jackson et al. 1988), fosforilada (Bonet et al. 2005) i recentment al nostre grup també hem vist que GAGA pot estar acetilada (X. Aran, comunicació personal). Aquesta gran diversitat de modificacions post-traduccionals podria estar indicant una acurada regulació de cada una de les múltiples funcions de GAGA.

A.3.1.1. Regulació de l'expressió de *Trl*

L'únic descrit sobre la regulació de l'expressió del gen *Trl* és, no obstant, molt sorprenent, ja que s'ha vist que la mateixa proteïna GAGA reprimeix l'expressió de *Trl* a cèl·lules de *Drosophila melanogaster* en cultiu (Kosoy et al. 2002). No només és sorprenent per estar GAGA implicada en un mecanisme de retroalimentació negativa de la seva pròpia expressió, que s'ha vist que ocorre per a moltes altres proteïnes, sinó pel fet que GAGA actuï com a repressor, sent un factor de la transcripció amb una marcada tendència activadora per als promotors on s'hi uneix.

A cèl·lules de *Drosophila melanogaster* en cultiu la sobreexpressió, per transfecció transitòria, de qualsevol de les dues isoformes de GAGA produeix una dràstica baixada de l'activitat de *Trl*. El promotor de *Trl* conté múltiples seqüències consens d'unió per a GAGA i de fet, per experiments de *footprinting* amb DNAsa I i de ChIP (*Chromatin immunoprecipitation*), s'ha vist que GAGA pot unir-se a aquestes seqüències, tant *in vivo* com *in vitro*. A més, l'expressió transcient a cèl·lules d'RNA d'interferència (RNAi) contra GAGA fa augmentar l'activitat del promotor de *Trl*, indicant que GAGA es troba constitutivament reprimint *Trl* en certa mesura. Per dur a terme la seva funció repressora sobre *Trl*, els dominis POZ i Q de GAGA no són necessaris, ja que formes delecionades per aquests dominis són perfectament competents en repressió. El que sí és imprescindible és que GAGA s'uneixi al DNA, com suggereix el fet que mutacions de GAGA per a la regió del DBD, que aboleixen la capacitat de GAGA d'unir-se al DNA, també aboleixen la capacitat de GAGA per reprimir *Trl*. Aquest és l'únic cas descrit en que GAGA actua com a factor repressor per un promotor d'un gen de *Drosophila*. S'ha suggerit que ha de ser un element específic de la seqüència del promotor de *Trl* el que determini que la unió de GAGA esdevingui en repressió, i no en activació transcripcional (Kosoy et al. 2002).

A.3.2. Estructura de la proteïna GAGA

La proteïna GAGA, com molts factors de transcripció, presenta una estructura composta per mòduls o dominis (veure figura A.11). Totes dues isoformes contenen tres dominis estructurals ben definits: un domini N-terminal anomenat BTB/POZ (*broad complex tramtrack bric-a-brac/poxvirus and Zinc finger*) (Bardwell et al. 1994); un domini d'unió a DNA (DBD) amb

INTRODUCCIÓ

un dit de Zinc i tres regions bàsiques (Pedone et al. 1996); i a C-terminal, un domini ric en residus glutamina, d'aquí que se l'anomeni domini Q. Les dues isoformes difereixen a partir de l'aminoàcid 378. Així, la diferència comprèn l'extrem C-terminal del DBD, que per altra banda no sembla ser necessari per a la unió de GAGA al DNA (Pedone et al. 1996), i tot el domini Q, més llarg en el cas de GAGA₅₈₁, ric en glutamines en ambdues isoformes (Benyajati et al. 1997). Entre el domini POZ i el DBD hi ha una regió, que anomenarem X, on s'ha trobat un possible senyal de localització nuclear (NLS) (Lintermann et al. 1998).

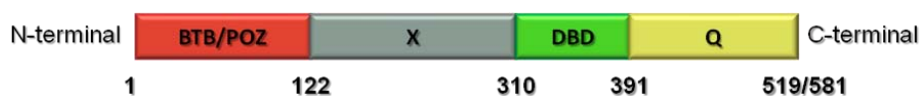


Figura A.11. Esquema dels dominis de les dues isoformes de la proteïna GAGA.

A.3.2.1. Domini d'unió al DNA (DBD)

El domini d'unió al DNA de GAGA comprèn una regió de 82 aa (de la posició 310 a la 391) i està format per un dit de Zinc tipus Cys₂-His₂ flanquejat per dues regions bàsiques (BR) a N-terminal i una a C-terminal (veure figura A.12). S'ha demostrat que els aminoàcids del 310 al 372, que comprenen les dues regions bàsiques a N-terminal (BR1 i BR2) i el dit de Zinc, són suficients per a una unió eficient i específica amb el DNA *in vitro* (Pedone et al. 1996). El domini d'unió al DNA de GAGA difereix del que és típic per a proteïnes d'unió al DNA amb dits de Zinc, ja que acostumen a tenir de 2 a 37 dits de Zinc, on al menys dos es requereixen per a una unió específica i d'alta afinitat amb el DNA (Klug et al. 1995). La manera en la que aquest DBD interacciona amb el DNA va ser resolta per RMN (ressonància magnètica nuclear) per Omichinski i col·laboradors utilitzant el fragment mínim necessari del DBD de GAGA (dels aminoàcids 310 al 372) i un oligonucleòtid contenint la seqüència GAGAG. La característica principal és l'existència de contactes específics amb cada una de les bases del consens GAGAG. Així, el dit de Zinc contacta amb els primers tres nucleòtids (GAGAG) del consens pel solc major del DNA, de manera similar a com ho fan els dits de Zinc d'altres proteïnes, mentre que BR1 i BR2 s'encarreguen d'estabilitzar la unió. BR1 envolta el DNA pel solc menor reconeixent l'A de la quarta posició del consens (GAGAG), mentre que BR2 forma una hèlix que interacciona amb el solc major del DNA reconeixent l'última G del consens (GAGAG) (Omichinski et al. 1997).

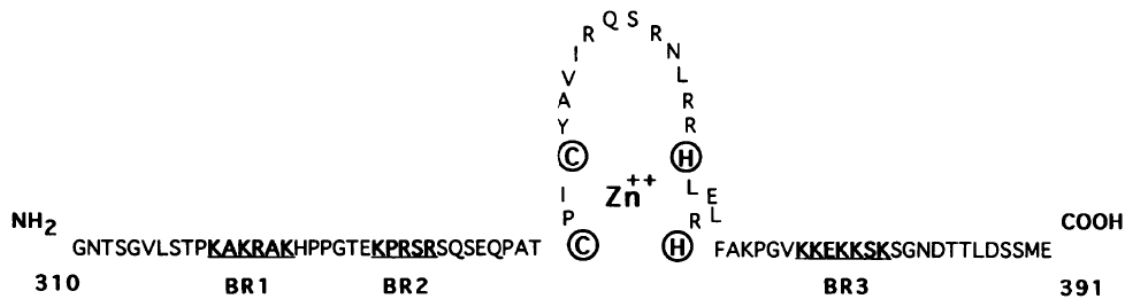


Figura A.12. Representació esquemàtica del DBD de GAGA₅₁₉. Els aminoàcids subratllats són els que formen les BRs que flanquegen el dit de Zinc, coordinat per dues cisteïnes i dues histidines (encerclades) (Pedone et al. 1996).

Primerament es va proposar el pentanucleòtid GAGAG com a seqüència consens mínima per a GAGA, tot i que la majoria de mutacions d'un únic parell de bases, sempre que no sigui el parell G-C central, només afecten a la unió de GAGA de forma moderada (Omichinski et al. 1997). Estudis posteriors varen demostrar que GAGA era capaç d'unir *in vitro* seqüències únicament compostes pel trinucleòtid GAG (Wilkins et al. 1998). Més enllà d'aquestes descripcions *in vitro*, el que s'observa *in vivo* és que les dianes naturals per a GAGA acostumen a estar formades per clústers de llocs GAGA als quals GAGA s'hi uneix de forma cooperativa, mitjançant la formació d'oligòmers a través del seu domini POZ (veure més endavant), que augmenten la seva afinitat pel DNA (Soeller et al. 1993; Espinas et al. 1999; Katsani et al. 1999). Estudis globals a nivell genòmic confirmen la necessitat de clústers de seqüències consens per a que GAGA s'hi uneixi i identifiquen una àmplia varietat de seqüències consens per a GAGA, d'entre les quals el pentanucleòtid GAGAG és la més habitual. A més, aquests estudis també evidencien que *in vivo* el trinucleòtid GAG no és suficient per a la unió de GAGA i que aquesta s'uneix preferencialment a les regions *upstream* de la regió codificant (regions promotores i 5'UTR), així com a regions intròniques, però està virtualment absent a regions exòniques (Sun et al. 2003; van Steensel et al. 2003).

Tot i que el paper predominant del DBD és el de reconèixer i unir el DNA, s'ha vist que també podria contribuir a algunes interaccions proteïna-proteïna descrites per a GAGA. Així, s'ha vist que el factor NURF301, tot i interaccionar majoritàriament amb la regió X, els aminoàcids del 362 al 397 (pertanyents al DBD) també contribueixen a la unió (Xiao et al. 2001). De la mateixa manera, per a l'interacció amb FACT semblen contribuir la regió X, el DBD i part del Q de GAGA₅₁₉ (única isoforma per a la que s'ha caracteritzat aquesta interacció) (Shimojima et al. 2003).

INTRODUCCIÓ

A la regió més C-terminal del DBD, s'ha vist que GAGA₅₁₉ (però no GAGA₅₈₁) pot ser fosforilada tant *in vitro* com molt probablement *in vivo* a les serines de les posicions 388 i en menor mesura 378. Tot i que aquests residus no estan presents a l'estructura d'RMN del complex GAGA DBD-DNA i semblen prescindibles pel reconeixement del DNA (Pedone et al. 1996; Omichinski et al. 1997), la seva fosforilació redueix (però no aboleix) la unió del DBD de GAGA al DNA (Bonet et al. 2005). Més recentment el nostre grup ha trobat que GAGA també pot ser acetilada a certes lisines de les BR1 i BR3, afectant negativament a la unió del DBD de GAGA amb DNA (X. Aran, comunicació personal).

A.3.2.2. Domini BTB/POZ

El domini BTB/POZ (al que ens referirem simplement com a POZ) comprèn uns 122aa de l'extrem N-terminal de GAGA i és una regió evolutivament conservada, trobada a proteïnes d'un ampli rang d'organismes, des de virus fins a humans. És característic d'una família de factors de transcripció amb dits de Zinc, certes proteïnes d'unió a actina i algunes proteïnes de *poxvirus* (Godt et al. 1993; Bardwell et al. 1994; Zollman et al. 1994). La naturalesa hidrofòbica d'aquest domini proporciona una superfície per a les interaccions proteïna-proteïna involucrades en processos tals com activació i repressió de la transcripció (Bardwell et al. 1994; Kaplan et al. 1997; Huynh et al. 1998; Pointud et al. 2001).

S'ha observat *in vitro* que la presència del domini POZ a les proteïnes que el contenen, entre elles GAGA, inhibeix en certa mesura la seva unió al DNA (Bardwell et al. 1994). En contrast amb això, com ja s'ha dit, les seqüències diana per a GAGA al DNA acostumen a presentar-se *in vivo* en forma de clústers de llocs GAGA. Així, el domini POZ, que és el responsable de la interacció de GAGA amb si mateixa formant oligòmers, possibilita el reconeixement cooperatiu d'aquests clústers de seqüències GAGA, augmentant d'aquesta manera l'afinitat i especificitat pel DNA (Soeller et al. 1993; Espinas et al. 1999; Katsani et al. 1999).

A part de servir de superfície per a la homooligomerització de GAGA, el POZ també és el domini per on GAGA interacciona amb multitud de proteïnes. S'han descrit interaccions tant amb altres proteïnes amb dominis POZ com amb proteïnes sense aquest domini. Així, el domini POZ de GAGA s'ha vist que interacciona amb el domini POZ de TTK (Tramtrack) (Bardwell et al. 1994; Pagans et al. 2002), Pipsqueak (Psq) (Schwendemann et al. 2002), Batman (també

conegut com lola-like) (Faucheux et al. 2003) i Mod(*mdg4*) (*modifier of mdg4*) (Melnikova et al. 2004). Entre els factors sense domini POZ que interaccionen amb el POZ de GAGA estan: SAP18 (*Sin3-associated polypeptide 18KDa*), que destaca per ser una proteïna associada al complex corepressor Sin3-HDAC (Espinass et al. 2000); la proteïna Corto, que també interacciona amb algunes proteïnes del grup *Polycomb* (PcG) i que podria jugar un paper en el silenciament mediat per proteïnes PcG (Salvaing et al. 2003); i el dTAF3 (*TBP-associated factor 3* de *Drosophila*), també conegut com a dTAF₁₁₅ o BIP2 (*Bric à brac Interacting Protein 2*), que a més constitueix el primer cas d'interacció directa per part de GAGA amb la maquinària basal de transcripció (Chopra et al. 2008).

Especialment destacables són les interaccions de GAGA amb Psq i Batman. Psq s'ha vist que reconeix les mateixes seqüències al DNA que GAGA, tot i tenir un domini d'unió al DNA diferent (Lehmann et al. 1998). A més d'interaccionar amb GAGA, colocalitza absolutament amb aquesta a cromosomes politènics i els mutants de *Psq* donen fenotips similars als mutants de *Trl*. Tot i això, les funcions de *Trl* i *Psq* no serien solapants, ja que tots dos gens són essencials per a la viabilitat de la mosca, suggerint que actuarien en conjunt de forma general (Schwendemann et al. 2002). D'altra banda, Batman, que pràcticament només està format per un domini POZ, a més d'interaccionar amb GAGA també colocalitza amb aquesta extensivament a cromosomes politènics i mostra les mateixes característiques que GAGA i Psq en quant a la regulació dels gens homeòtics (Faucheux et al. 2003; Mishra et al. 2003). Així, donada l'estreta relació d'aquestes tres proteïnes, s'ha suggerit que podrien estar formant part d'un mateix complex, anomenat GBP (GAGA-Batman-Psq), tot i que encara no hi ha proves bioquímiques de que aquest complex existeixi realment (Lehmann 2004).

A.3.2.3. Domini ric en glutamines Q

Comprèn l'extrem C-terminal de la proteïna i la seva característica més evident és, com el seu nom indica, la seva abundància relativa en residus glutamina. Aquesta característica s'ha vist que és pròpia d'alguns activadors de la transcripció (Granok et al. 1995). Aquest domini és la regió majoritària on difereixen les dues isoformes de GAGA. Tot i que és ric en glutamines per ambdues isoformes, és gairebé el doble de llarg en el cas de GAGA₅₈₁. A més, l'abundància relativa de glutamines del Q de GAGA₅₈₁ també és major que la del de GAGA₅₁₉ (Benyajati et al. 1997).

INTRODUCCIÓ

Inicialment es va descriure que GAGA no funcionava com un activador transcripcional clàssic, en el sentit que la seva capacitat d'activar en realitat era reflex de la seva capacitat de funcionar com a antirepressor, possibilitant un estat obert de la cromatina per a que altres activadors reals actuessin (Croston et al. 1991; Kerrigan et al. 1991). En aquest context, es van descriure primerament tres funcions pel domini Q: distorsió del promotor, unió a DNA monocadena i multimerització, estant implicat sobretot en la formació d'homòmers (Wilkins et al. 1999). A més, en concordança amb això, el domini Q es va veure que *in vitro* era el responsable de la formació de fibres amiloides (una forma agregada de proteïnes) i prescindible per a la remodelació de cromatina (Agianian et al. 1999). Tot i amb això, posteriorment es va veure que l'activació de la transcripció per GAGA podia ocórrer independentment que GAGA induís la remodelació de la cromatina a la zona del promotor, suggerint que la remodelació de la cromatina i l'activació transcripcional, dutes a terme per GAGA, eren fenòmens separats (Pile et al. 2000). Això va donar peu a pensar en GAGA com a un genuí factor activador de la transcripció. Finalment, es va demostrar que el domini Q era el responsable de la capacitat transactivadora de GAGA, tant *in vitro* com *in vivo*, probablement mitjançant l'estabilització del PIC i promovent les reiniciacions de la transcripció (Vaquero et al. 2000; Vaquero et al. 2008).

A.3.2.4. Regió X

Està situada entre el domini POZ i el DBD. És una regió per a la qual no s'ha pogut caracteritzar encara una funció concreta, d'aquí que l'anomenem com a regió X. La primera característica associada a aquesta regió és l'existència d'una possible senyal de localització nuclear (NLS) entre les posicions 209-217. A més d'això, és destacable que la comparació de seqüències entre GAGA de *D. melanogaster* i de *D. virilis* mostra com la regió X està altament conservada entre ambdues espècies, especialment a la part més N-terminal, englobant uns 70 aa adjacents al domini POZ, on les proteïnes d'ambdues espècies són idèntiques (per a una comparació de les seqüències veure Resultats, apartat D.1.1, figura D.1). Això dóna lloc a pensar que aquesta regió és responsable d'alguna funció essencial (Lintermann et al. 1998).

A banda d'això, s'ha vist que la regió X és responsable d'algunes interaccions amb altres proteïnes. Així, juntament amb el domini POZ (veure apartat A.3.2.2.), la part N-terminal de la regió X contribueix a la interacció GAGA-SAP18 (Espinass et al. 2000). A més, s'ha vist que a banda de certs residus del DBD (veure apartat A.3.2.1), la zona de GAGA més important per a

l'interacció amb NURF301 està situada a la regió X (Xiao et al. 2001). Per últim, també s'ha vist que, a més del DBD i part del Q de GAGA₅₁₉ (veure apartat A.3.2.1), la regió X també contribueix a l'interacció amb FACT (Shimojima et al. 2003).

A.3.3. Funcions de GAGA

A.3.3.1. GAGA com a antirepressor i activador transcripcional

Des de la identificació de GAGA com a factor activador *in vitro* de la transcripció dels gens *Ultrabithorax* i *engrailed* (Biggin et al. 1988; Soeller et al. 1988), els estudis fets per descriure el mecanisme pel qual GAGA actua van portar a pensar en GAGA com un activador transcripcional atípic. Els primers experiments al respecte, duts a terme per Kerrigan i col·laboradors, demostraren que *in vitro* GAGA és un activador transcripcional per a *Krüppel* i *Ultrabithorax*, però per a que aquesta activació esdevingui eficient necessita de la presència a la reacció d'un factor repressor inespecífic present a certs extractes. A diferència de factors activadors clàssics (com per exemple la proteïna de fusió GAL4-VP16) que suposadament estimulen la transcripció interaccionant amb la maquinària basal de transcripció, es va proposar que l'activació duta a terme per GAGA es basava en l'eliminació d'aquest factor inespecífic, en lloc d'autèntica activació, pel mecanisme que es va anomenar "antirepressió" (Kerrigan et al. 1991). Posteriorment es va identificar que la histona H1 funcionava com un repressor inespecífic i general de la transcripció per RNA Pol II i que factors de transcripció activadors com Sp1, GAL4-VP16 o GAGA eren capaços de contrarestar aquesta repressió, pel mecanisme de l'antirepressió, però només els dos primers factors eren capaços d'activar en absència d'H1. Així, l'activació real i l'antirepressió es van definir com a fenòmens separats (Croston et al. 1991). El mecanisme d'antirepressió s'ha vist que és general de molts factors activadors de la transcripció i que la unió al DNA, tot i que és suficient per alterar el complex H1-DNA, no ho és per dur a terme l'antirepressió, que requereix a més del domini activador (Croston et al. 1992). A banda d'aquests estudis *in vitro*, posteriorment es va descriure que GAGA podia funcionar com a remodelador de la cromatina a certs promotors tant *in vitro* (Tsukiyama et al. 1994; Wall et al. 1995; Okada et al. 1998) com *in vivo* (Lu et al. 1993; Weber et al. 1997). A més, també es va observar que els mutants hipomorfs per a *Trl*, com *Trl^{13C}*, presentaven un augment del PEV, així com defectes en la divisió cel·lular tals com asincronia, fallada en la condensació cromosòmica,

INTRODUCCIÓ

segregació cromosòmica anormal i fragmentació cromosòmica (Farkas et al. 1994; Bhat et al. 1996), fet que també recolzava la idea de GAGA com a factor implicat en l'estat de la cromatina. Ajuntant totes aquestes evidències, la idea general que va predominar va ser que GAGA funcionava com a antirepressor per la seva capacitat de remodelar la cromatina als promotors on s'hi unia, més que com a activador real (Wilkins et al. 1997).

Actualment sabem que GAGA, a més de funcionar com a antirepressor en certes condicions, també és un activador de la transcripció en el sentit més clàssic, ja que per mitjà del seu domini Q és capaç d'estimular la transcripció. Això s'ha comprovat tant per assajos de transcripció *in vitro* sobre DNA despullat, com *in vivo* per assajos de transfecció transitòria a cèl·lules de *Drosophila*. S'ha proposat que GAGA, per mitjà del seu domini Q, actua estabilitzant la formació del complex de preiniciació i promovent les reiniciacions (Vaquero et al. 2000; Vaquero et al. 2008). A més, s'ha vist que la capacitat activadora de GAGA és major quant major és el número de llocs d'unió existents al promotor. Tenint en compte la tendència que tenen les seqüències d'unió per a GAGA de presentar-se en forma de clústers (van Steensel et al. 2003), i donada la capacitat de GAGA per oligomeritzar amb si mateixa, la unió a aquestes seqüències es realitzaria de manera cooperativa per diverses molècules de GAGA actuant sinèrgicament (Soeller et al. 1993; Espinas et al. 1999; Katsani et al. 1999). En concordança amb això, a part del seu domini de transactivació Q, s'ha vist que GAGA necessita del seu domini de oligomerització POZ per poder dur a terme una activació de la transcripció més eficient (Espinas et al. 1999; Vaquero et al. 2000). A més d'això, l'oligomerització de GAGA s'ha vist que està implicada en l'apropament de regions *enhancer* i promotores molt distants i fins i tot els oligòmers de GAGA són capaços d'unir, per mitjà del reconeixement de les seves seqüències, dues cadenes de DNA independents, donant així fenòmens d'activació en *trans* (Mahmoudi et al. 2002).

A.3.3.2. GAGA com a remodelador de la cromatina

Les primeres indicacions de que GAGA podria estar implicada en la remodelació de l'estructura de la cromatina provenen de l'estudi dels promotors de les *heat shock proteins* (*hsp*), sobretot *hsp26* i *hsp70*. Aquests promotors, en estat de no inducció, contenen l'RNA Pol II aturada, pròxima al promotor i havent transcrit uns 25 nucleòtids. La transcripció només continua en condicions d'inducció (Rougvie et al. 1988). Aquest fenomen es coneix com a PrPP (veure apartat A.2.5.2). Ja inicialment es va veure que GAGA es trobava al promotor d'*hsp70* i

hsp26 in vivo en condicions de no inducció, suggerint un paper per a GAGA en el manteniment de l'estat de parada d'aquests gens (Gilmour et al. 1989). Posteriorment es va confirmar que les seqüències GAGA estaven implicades en el procés de PrPP del promotor d'*hsp70* (Lee et al. 1992b) i analitzant el promotor d'*hsp26* es va veure que la proteïna GAGA estava implicada en la formació de llocs hipersensibles a la digestió amb DNAsa I, indicatius de regions lliures de nucleosomes (Lu et al. 1993). Es va trobar que la capacitat de desplaçar nucleosomes de GAGA venia donada pel factor NURF. Aquesta activitat és la responsable de la formació de les zones hipersensibles dels promotors dels gens de *Heat shock*. NURF és un complex remodelador de la cromatina dependent d'ATP format per quatre subunitats, de les quals ISWI és la que li confereix l'activitat ATPasa, essencial per a la funció remodeladora de la cromatina (Tsukiyama et al. 1995a; Tsukiyama et al. 1995b). A més, com ja s'ha comentat, GAGA interacciona amb la subunitat major del complex NURF, NURF301 (Xiao et al. 2001). En el cas dels promotors dels gens *Heat shock*, s'ha vist que la unió de GAGA al promotor produeix la remodelació de la cromatina necessària per a l'entrada de TFIID i l'RNA Pol II al promotor (Shopland et al. 1995; Weber et al. 1997). Per que es pugui donar el PrPP a aquests promotors, sembla que també estarien implicats els factors NELF i DSIF (veure apartat A.2.5.2). En estat d'inducció per xoc tèrmic, HSF (*Heat shock factor*) s'uneix al promotor per reconeixement de les seves seqüències diana, els HSE (*Heat shock elements*). HSF és el responsable directe de l'activació dels gens de *Heat shock*, tot i que GAGA sembla que estabilitza la unió de HSF amb el DNA (Mason et al. 1997). D'altra banda, GAGA s'ha trobat associada a la regió codificant d'aquests gens en estat induït, coincidint amb la localització de l'RNA Pol II, el que suggereix que el paper de GAGA com a remodelador de la cromatina pot ser important també durant la fase d'elongació (O'Brien et al. 1995). El fenomen del PrPP, lluny de només afectar als gens de *Heat shock*, avui es pensa que podria ser una fase més necessària per la transcripció de molts gens i GAGA, juntament amb altres factors, sembla estar estretament implicada en aquest procés (Wang et al. 2005; Lee et al. 2008).

La capacitat de GAGA per remodelar la cromatina no és només un fet aïllat dels promotors de *Heat Shock*, ja que s'ha vist que GAGA, juntament amb l'ATPasa ISWI, està implicada en la remodelació de la cromatina responsable de l'activació del gen *fushi tarazu* (*ftz*) (Okada et al. 1998). No obstant això, la demostració de que l'activació per part de GAGA d'un promotor no comporta necessàriament la remodelació de la cromatina, indica que el fet que

INTRODUCCIÓ

GAGA actua com a remodelador de la cromatina, ajudi al PrPP, o activi l'expressió de manera directa, depèn del context de cada promotor concret i de la resta de factors que en ell convergeixin (Pile et al. 2000).

A més de la capacitat de GAGA per remodelar la cromatina dependent de NURF, també s'ha vist que GAGA pot interaccionar amb el complex FACT amb funció chaperona d'histones. Aquesta interacció facilita la remodelació de la cromatina dependent de GAGA i té implicacions en la regulació de la transcripció de gens com *Ultrabithorax (Ubx)* i *Abdominal-B (Abd-B)* (Shimojima et al. 2003). Més recentment s'ha vist que tant GAGA com FACT estan implicats en el PEV. Aquests dos factors dirigeixen l'intercanvi de l'histona H3 per la H3.3 a certes zones, fet que contraresta l'expansió de l'heterocromatina (Nakayama et al. 2007).

A.3.3.3. GAGA com a proteïna TrxG i PcG

Durant el desenvolupament dels animals, la identitat cel·lular i la correcta realització del patró corporal depenen de la informació posicional donada pels patrons d'expressió dels gens homeòtics, també coneguts com a gens *Hox*. El manteniment epigenètic d'aquests patrons al llarg de les successives divisions cel·lulars es creu que involucra modificacions estables i heretables de la cromatina. Cerques genètiques a *Drosophila* han identificat dos grups de proteïnes que contribueixen a aquestes modificacions: les proteïnes del grup *Trithorax (TxG)*, requerides pel manteniment de l'expressió dels gens homeòtics; i les proteïnes del grup *Polycomb (PcG)*, requerides pel silenciament d'aquests gens (Simon et al. 2002). Les proteïnes dels PcG i TrxG interaccionen amb el DNA mitjançant unes seqüències anomenades *Polycomb Response Elements (PREs)* i *Trithorax Response Elements (TREs)* respectivament. De tota manera, la diferenciació entre PRE i TRE no sembla ser molt clara i de fet sembla que es solapen en molts casos, el que és consistent amb el fet que proteïnes TrxG i PcG comparteixin molts llocs d'unió (Francis et al. 2001). Per aquests motius, en molts casos els elements PRE/TRE s'anomenen simplement PRE o elements de manteniment.

Un dels *loci* genòmics on més s'ha estudiat l'acció de les proteïnes TrxG i PcG és el complex *Bithorax (BX-C)* de *Drosophila* que comprèn tres gens i les seves regions reguladores: *Ultrabithorax (Ubx)*, *Abdominal-A (Abd-A)* i *Abdominal-B (Abd-B)*. El BX-C és el responsable de la identitat dels segments a la regió posterior de l'embrió i comprèn una regió de més de 300 Kpb (veure figura A.13). L'expressió dels tres gens del BX-C esta regulada acuradament per nou

dominis reguladors, ordenats de la mateixa manera que els parasegments que controlen: *abx/bx*, *bxd/pbx*, *iab-2*, *iab-3*, *iab-4*, *iab-5*, *iab-6*, *iab-7* i *iab-8,9* (Mihaly et al. 1998). Degut a que *Trl* és essencial per a la correcta expressió dels gens homeòtics *Ubx* i *Abd-B*, des de la seva caracterització va ser classificat com a gen TxG (Farkas et al. 1994). Es va trobar que certs PREs de BX-C contenen seqüències GAGA i que aquesta s'hi unia de forma constitutiva, amb preferència pels més actius transcripcionalment, en concordança amb la seva funció com a proteïna TrxG (Strutt et al. 1997). Sorprenentment, també es va trobar que GAGA era requerida al PRE del *iab-7* per a formar una estructura oberta de la cromatina que permetés l'entrada de proteïnes PcG i el posterior silenciament. Així doncs, en aquest context, GAGA funciona com una proteïna PcG, ja que possibilita el silenciament (Hagstrom et al. 1997). A més, es va veure que GAGA formava part d'alguns complexos PcG i participa en la seva unió al PRE de *bxd*, per mitjà del reconeixement de les seqüències GAGA presents en aquest PRE (Horard et al. 2000; Hodgson et al. 2001). Posteriorment s'ha vist que GAGA forma part tant de complexos TxG com PcG i la unió d'aquests a alguns PREs depèn de GAGA (Poux et al. 2002). De totes maneres, tot i que GAGA forma part d'aquests complexos, sembla que per portar-los als PREs requeriria de la participació d'altres factors, ja que no s'ha descrit que GAGA interaccioni directament amb proteïnes TrxG o PcG. En aquest sentit, s'ha descrit que Batman, identificat inicialment com a proteïna PcG, interacciona amb GAGA, està implicat tant en l'activació com en la repressió dels gens homeòtics, i el complex que forma amb GAGA s'uneix cooperativament al PRE de *bxd* (Faucheux et al. 2003). A més, s'ha trobat que Batman interacciona amb certes proteïnes PcG i, juntament amb GAGA, se l'ha trobat formant complexos amb aquestes. Això suggereix que Batman podria funcionar com a mediador de GAGA en el reclutament de les proteïnes PcG als PREs (Mishra et al. 2003). Una altra de les proteïnes que, mitjançant la seva interacció amb GAGA, podria estar implicada en portar proteïnes PcG als PREs és Corto. A més d'interaccionar amb GAGA, Corto s'ha vist que pot interaccionar amb certes proteïnes PcG i podria per tant servir de nexa d'unió entre GAGA i els complexos PcG (Salvaing et al. 2003). A banda d'això, la interacció GAGA-SAP18 també podria jugar un paper important en el mecanisme de silenciament donat per certs PREs, ja que SAP18 sembla que interacciona amb el complex corepressor Sin3-HDAC (Espinassé et al. 2000). De totes maneres, tot i la importància de GAGA en el mecanisme de la regulació gènica dut a terme per proteïnes PcG i TrxG, estudis de *ChIP on Chip* recents indiquen que aquesta no seria la funció predominant de GAGA, ja que la seva

INTRODUCCIÓ

colocalització amb *Polycomb* (PC) i *Polyhomeotic* (PH), dues proteïnes PcG, és molt limitada, suggerint que GAGA no és un component necessari de la cromatina als PREs (Negre et al. 2006; Schuettengruber et al. 2009).

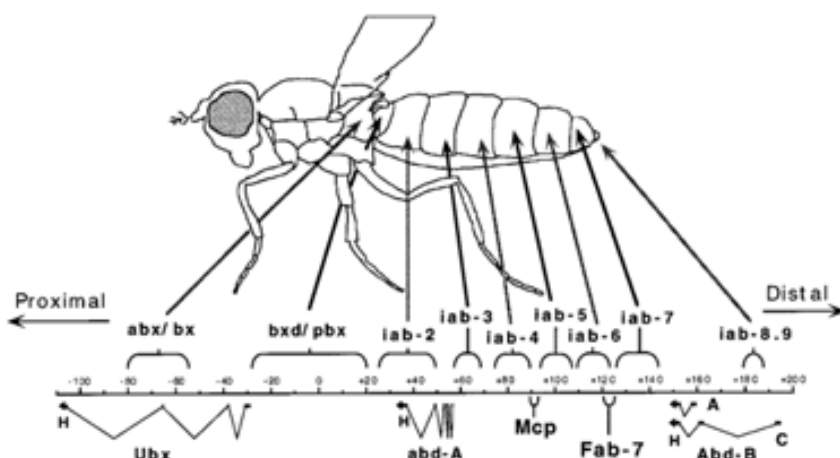


Figura A.13. Esquema de la regió reguladora del complex *Bithorax*. El BX-C està format per tres gens, l'*Ubx*, l'*Abd-A* i l'*Abd-B*, i pels dominis que els regulen. Les fletxes indiquen el segment regulat per cada domini (Mihaly et al. 1998).

A.3.3.4. GAGA als *insulators*

Els *insulators* són elements de seqüència del DNA que estableixen dominis independents d'activitat transcripcional, als genomes eucariotes. Venen definits per dues propietats funcionals. En primer lloc, bloquegen l'activitat d'*enhancers* i silenciadors quan són inserits entre aquests i un promotor (veure figura A.14.A). En aquest sentit, els *insulators* són els responsables de la important funció de restringir l'acció dels *enhancers* i silenciadors al promotor adient. En segon lloc, els *insulators* protegeixen l'expressió gènica dels efectes positius i negatius de la cromatina adjacent a un gen o *locus* genòmic (veure figura A.14.B). Els mecanismes per la protecció dels efectes de la cromatina propera poden implicar el bloqueig directe d'elements reguladors o la restricció de l'extensió de complexos silenciadors, associats freqüentment amb la cromatina repressiva. Els elements amb propietats *insulator* es distribueixen al llarg de tot el genoma dels eucariotes, com cabria esperar de la seva funció establint els límits de dominis de cromatina, dins de l'eucromatina i l'heterocromatina (Kuhn et al. 2003).

Alguns estudis indiquen que les mutacions en gens TrxG augmenten l'efecte dels *insulators* en *enhancers* adjacents, mentre que mutacions en gens PcG tenen l'efecte contrari. Això ha portat a pensar que les proteïnes PcG i TrxG podrien jugar un paper en la regulació dels *insulators* (Gerasimova et al. 1998). La primera evidència de que GAGA podria estar implicada

en la funció dels *insulators* sorgeix de l'observació de que GAGA, mitjançant les seves seqüències d'unió al DNA, s'uneix a elements *insulator* presents a la regió reguladora del gen α -1 *tubulina*. Aquests *insulators* confereixen al gen α -1 *tubulina* la capacitat d'expressar-se ubiquament, de forma independent de l'estat de la cromatina del seu entorn. Essent GAGA un factor associat a la remodelació de la cromatina, es va especular que aquesta propietat podria ser important per la funció d'aquests elements *insulator* (O'Donnell et al. 1994). Posteriorment, mitjançant construccions artificials, es va veure que GAGA era responsable del bloqueig de la comunicació entre un *enhancer* i el promotor d'*even skipped* (*eve*), a l'embrió de *Drosophila*, suggerint un paper de GAGA en la funció de l'element *insulator* d'aquest promotor (Ohtsuki et al. 1998). Estudis més actuals han confirmat el paper de GAGA en la funció de certs *insulators*. Així, s'ha trobat que les seqüències GAGA són essencials per a la funció d'un *insulator* situat a la regió intergènica entre *Sex combs reduced* (*Scr*) i *fushi tarazu* (*ftz*), que confereix independència a aquests dos gens. A més, mutants hipomorfs per a GAGA tenen la funció d'aquest *insulator* disminuïda (Belozarov et al. 2003). També s'ha trobat que GAGA és important per la funció *insulator* del Fab-7 del BX-C, tot i que GAGA no seria l'únic element implicat en aquesta funció (Schweinsberg et al. 2004). A més d'això, recentment s'ha vist que el complex GAGA-FACT està implicat en l'intercanvi de l'histona H3 per l'H3.3, que es dona a partir de certes regions *insulator*. Això podria constituir un mecanisme per evitar la propagació de l'heterocromatina adjacent (Nakayama et al. 2007).

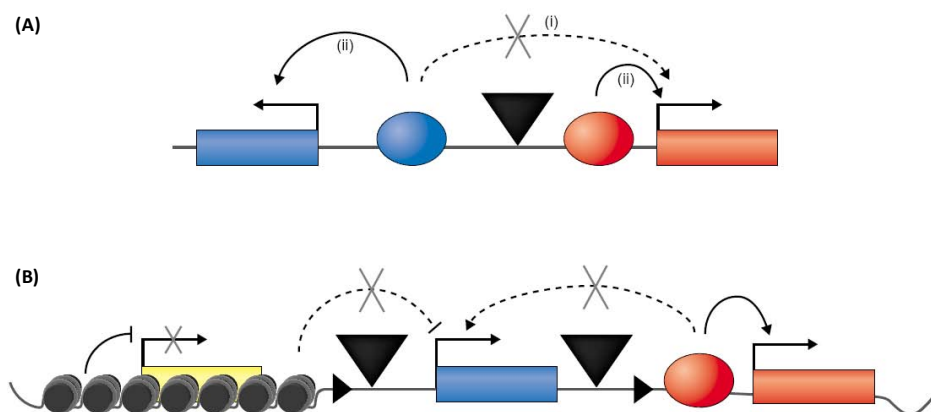


Figura A.14. Propietats dels *insulators*. **(A)** Els *insulators* (triangles negres) bloquegen l'acció dels *enhancers* de manera dependent de posició. (i) Un *insulator* inserit entre un *enhancer* (cercle blau) i el promotor d'un gen (rectangle vermell) bloqueja l'activació de la transcripció donada per l'*enhancer*. (ii) Per contra, l'*insulator* no té efecte si es situa *upstream* de l'*enhancer* (cercle vermell). **(B)** Els *insulators* protegeixen d'efectes de posició cromosòmica. Un transgèn (rectangle blau) flanquejat per *insulators* està protegit de la cromatina repressiva (esquerra) i de l'acció directa d'elements reguladors com els *enhancers* (dreta). Les interaccions *enhancer*-promotor es representen per fletxes, mentre que les interaccions bloquejades es representen com a fletxes discontinües amb una creu (Kuhn et al. 2003).

B. OBJECTIUS

En aquest treball ens vàrem proposar els següents objectius:

- Analitzar els mecanismes de regulació del gen *Trl* de *Drosophila melanogaster*, centrant l'atenció en l'autoregulació duta a terme per GAGA.
- Realitzar un estudi funcional de la proteïna GAGA, per mitjà de *microarrays* a cèl·lules S2, per tal d'identificar els diversos processos biològics en que participa.

C. MATERIALS I MÈTODES

C.1. Materials

C.1.1. Reactius i enzims

La majoria de productes que hem utilitzat en la realització dels experiments són de qualitat analítica i provenen de les cases comercials Merk, Fluka i Sigma. Els enzims que hem fet servir són de Fermentas, Roche, Promega, Gibco, New England Biolabs (NEB) i Biotools.

C.1.2. Soques bacterianes

En aquest treball s'han utilitzat les següents soques d'*Escherichia coli*:

- Soques per a l'amplificació de plasmidis:
 - DH5 α : deoR endA1 gyrA96 hsdR17 ($r_k^- m_k^+$) recA1 relA1 supE44 thi-1 F'[traD36 proAB⁺ lac I^q lac Z Δ M15].
 - HB101: F- supE44 ara14 galK2 lacY1 Δ (gpt-proA)62 leuB6 rpsL20 (Str^r) xyl-5 mtl-1 recA13 Δ (mcrC-mr^r).
- Soca per a l'expressió de proteïnes:
 - BL21 (DE3): *E. coli* B F⁻ dcm ompT hsdS ($r_B^- m_B^-$) gal λ (DE3).

C.1.3. Plasmidis

C.1.3.1. Plasmidis de clonatge

- . **pGL3-Basic (Promega)**: conté el cDNA de la luciferasa de la cuca de llum (*Photinus pyralis*) com a *reporter*.
- . **pActppa**: plasmidi d'expressió a cèl·lules de *Drosophila* que conté la regió promotora del gen Actina 5C de *Drosophila melanogaster*. Cedit amablement pel Dr. Gerardo Jiménez.
- . **pCR2.1-TOPO (Invitrogen)**: utilitzat pel clonatge de fragments de PCR pel sistema TOPO-TA (Invitrogen).
- . **pTopo-Lac I**: conté el cDNA de Lac I inserit al pCR2.1-TOPO. Pel clonatge es va amplificar per PCR el Lac I (d'origen bacterià) utilitzant un parell d'encebadors amb dianes Not I (l'encebador de 5') i Nde I (l'encebador de 3'). Realitzat per Joan Font.

MATERIALS I MÈTODES

. **pTopo-Lac I-GAGA-DBD**: conté el cDNA de Lac I, amb el cDNA del DBD de GAGA inserit en pauta a 3'. Vàrem extreure el cDNA corresponent al DBD de GAGA (aminoàcids del 310 al 391) del pET14b-GAGA-DBD digerint-lo amb Bam HI, omplint els extrems amb Klenow, i digerint amb Nde I. Aquest insert el vàrem lligar al pTopo-Lac I, digerit amb Xho I (amb els extrems omplerts per Klenow) i Nde I.

. **pTopo-Trl vir**: conté la regió promotora de *Trl* de *Drosophila virilis*. A partir d'extractes de DNA genòmic de cèl·lules WR DV1, vàrem amplificar els 2979 pb *upstream* del primer ATG de la regió codificant de *Trl* de *D. virilis*. Vam utilitzar els oligonucleòtids upTrlvir i lowTrlvir per amplificar aquest fragment per PCR, i posteriorment el vam inserir al pCR2.1-TOPO.

. **pTopo-GAGA-A**: conté el cDNA de la isoforma GAGA-A de *Drosophila virilis*. El vàrem amplificar a partir d'extractes d'RNA total de cèl·lules WR DV1, per RT-PCR amb els encebadors upGAGAvir i lowGAGAvir. Posteriorment el vam inserir al pCR2.1-TOPO.

. **pBlueScribe M13+ (Stratagene)**: plasmidi utilitzat pel clonatge de fragments de cDNA que posteriorment s'utilitzaran per a produir dsRNA.

. **pEGFP-N1 (BD Biosciences Clontech)**: conté el cDNA de l'*Enhanced Green Fluorescent Protein* (EGFP). Empleat per a la construcció de fusions de proteïnes amb EGFP a C-terminal.

. **pECFP-Nuc (BD Biosciences Clontech)**: conté el cDNA de l'*Enhanced Cyan Fluorescent Protein* (ECFP) amb tres *Nuclear Localization Signal* (NLS). Empleat per a la construcció de fusions de proteïnes amb ECFP i tres NLS a N-terminal.

. **pCDNA3 (Invitrogen)**: conté el promotor de CMV. Utilitzat per expressar proteïnes a la línia de cèl·lules HeLa, ja que el promotor de CMV és molt més potent a HeLa que no pas el d'Actina 5C present al plasmidi Actppa.

. **pET14b (Novagen)**: permet l'obtenció de proteïnes fusionades a una cua de 6 histidines en posició N-terminal.

. **pET14b-GAGA₅₁₉**: conté el cDNA de GAGA₅₁₉. Realitzat pel Dr. Emilio Jiménez (Tesi Doctoral, 1998).

. **pET14b-GAGA-DBD**: conté el cDNA corresponent al DBD de GAGA, de l'aminoàcid 310 al 391. Realitzat pel Dr. Emilio Jiménez (Tesi Doctoral, 1998).

. **pET28a (Novagen)**: permet l'obtenció de proteïnes fusionades a una cua de 6 histidines en posició N-terminal.

. **pGEX-3X-Gro**: conté el cDNA de la proteïna Groucho, inserit a un pGEX-3X (Amersham-Pharmacia-Biotech). Cedit amablement pel Dr. Gerardo Jiménez.

. **pCaSpeR4**: generat pel Dr. Pirrota, s'ha utilitzat per a clonar les construccions del promotor de *Trl* fusionat a EGFP, emprades en la generació de línies de mosques transgèniques.

. **pUAST**: generat per Brand i Perrimon (Brand et al. 1993), conté cinc llocs òptims d'unió per a la proteïna GAL4, seguits de la caixa TATA del gen *hsp70* i d'un lloc d'inici de la transcripció. Empleat per al clonatge de plasmidis d'expressió sota el control de GAL4 a mosques transgèniques.

C.1.3.2. Plasmidis reporter

. **pGL3-Trl full / pGL3-p2b / pGL3 "-204/+737" / pGL3 "-3470/+347"**. Les quatre construccions contenen respectivament els següents fragments del promotor de *Trl* de *Drosophila melanogaster* (prenent com a +1 l'inici de transcripció més distal): de -3470 a +737; de -345 a +737; de -204 a +737; de -3470 a +347. Generades per la Dra. Ana Kosoy (Tesi Doctoral, 2002).

. **pGL3 "-298/+737"**: generat a partir del pGL3-Trl full per PCR amb la DNA polimerasa Expand (Roche) utilitzant els oligonucleòtids p(-) i FII(4-3) com a encebadors. El primer va ser dissenyat per la Dra. Ana Kosoy i el segon ha sigut dissenyat per a aquest clonatge. Tots dos contenen una diana per a l'enzim de restricció Asp 718. L'oligonucleòtid p(-) s'aparella just a l'extrem 5' de l'insert de pGL3-Trl full i dirigeix l'amplificació en sentit *upstream* de l'insert, mentre que FII(4-3) s'aparella a la zona entorn a -298 dirigint l'amplificació *downtream*. Així, mitjançant aquesta PCR divergent, s'amplifica la totalitat del plasmidi pGL3-Trl full a excepció de la zona a deleccionar. Seguidament es digereix el fragment amplificat per PCR amb Asp718 i es lliga amb lligasa T4 (Fermentas), circularitzant així el fragment i obtenint un plasmidi pGL3 complet amb la regió -298/+737 com a insert, que controlarà l'expressió del cDNA del *reporter* luciferasa.

. **pGL3 "-345/+71" / pGL3 "-345/+14" / pGL3 "-345/-33"**: aquestes tres construccions han estat generades a partir del pGL3-p2b, per PCR amb la DNA Polimerasa Expand. Com a encebadors s'han utilitzat, per una banda, l'oligonucleòtid p2b (dissenyat per la Dra. Ana Kosoy), en els tres casos, i per altra banda els oligonucleòtids Trl pr+71, Trl pr+14, Trl pr-33, a cada una de les construccions, respectivament. L'encebador p2b conté una diana de restricció

MATERIALS I MÈTODES

Asp718 i els altres tres una diana Bgl II, totes introduïdes per a dur a terme el clonatge ja que, després de l'amplificació de cada fragment per PCR, es van digerir amb Asp718 i Bgl II i es van lligar a pGL3-Basic digerit de la mateixa manera.

. **pGL3-p2b-303:** generat a partir del pGL3-Trl full per PCR amb la DNA polimerasa Expand. Vàrem amplificar dos fragments solapants. El primer, per mitjà dels oligonucleòtids p2b i Down-303, dóna un fragment que va de l'inici de p2b (-345) fins a la zona de -303. El segon, per mitjà dels oligonucleòtids Up-303 i luciferasa revers (luc rev), dóna un fragment que va de la regió de -303 fins al principi del cDNA de la luciferasa. Els oligonucleòtids Up-303 i Down-303 són en gran part solapants, però amplifiquen en sentits oposats (hibriden en cadenes oposades) i a més porten una diana per a l'enzim de restricció Eco RV, que introdueix les mutacions a la zona de -303 que persegueix aquest clonatge. El següent pas és ajuntar els productes d'ambdues PCRs i realitzar una nova PCR emprant només els encebadors p2b i luc rev. Això dóna com a resultat el fragment p2b, que incorpora les mutacions a -303 (comprovables per la presència de la diana EcoRV) i la part inicial del cDNA de la luciferasa. Seguidament, es digereix aquest fragment amb Asp718 i Nco I (situat al principi del cDNA de la luciferasa, abans d'on s'uneix luc rev) i el resultat es lliga amb un pGL3-Basic digerit de la mateixa manera.

. **pGL3-p2b-265:** generat a partir del pGL3-Trl full per PCR amb la DNA polimerasa Expand. Vàrem amplificar dos fragments solapants. El primer, per mitjà dels oligonucleòtids p2b i Down-265, dóna un fragment que va de l'inici de p2b (-345) fins a la zona de -265. El segon, per mitjà dels oligonucleòtids Up-265 i luciferasa revers (luc rev), dóna un fragment que va de la regió de -265 fins al principi del cDNA de la luciferasa. Els oligonucleòtids Up-265 i Down-265 són en gran part solapants, però amplifiquen en sentits oposats (hibriden en cadenes oposades) i a més porten una diana per a l'enzim de restricció Eco RI, que introdueix les mutacions a la zona de -265 que persegueix aquest clonatge. El següent pas és ajuntar els productes d'ambdues PCRs i realitzar una nova PCR emprant només els encebadors p2b i luc rev. Això dóna com a resultat el fragment p2b, que incorpora les mutacions a -265 (comprovables per la presència de la diana EcoRI) i la part inicial del cDNA de la luciferasa. Seguidament, es digereix aquest fragment amb Asp718 i Nco I (situat al principi del cDNA de la luciferasa, abans d'on s'uneix luc rev) i el resultat es lliga amb un pGL3-Basic digerit de la mateixa manera.

. **pGL3-p2b-303-265:** per generar aquest plasmidi es repeteix el mateix procediment que per generar p2b-265, però utilitzant com a plasmidi de partida pGL3-p2b-303. Així, el plasmidi resultant portarà les mutacions a -303 i -265.

. **pGL3-p2b-227:** generat a partir del pGL3-Trl full per PCR amb la DNA polimerasa Expand. Vàrem amplificar dos fragments solapants. El primer, per mitjà dels oligonucleòtids p2b i Down-227, dóna un fragment que va de l'inici de p2b (-345) fins a la zona de -227. El segon, per mitjà dels oligonucleòtids Up-227 i luciferasa revers (luc rev), dóna un fragment que va de la regió de -227 fins al principi del cDNA de la luciferasa. Els oligonucleòtids Up-227 i Down-227 són en gran part solapants, però amplifiquen en sentits oposats (hibriden en cadenes oposades) i a més porten una diana per a l'enzim de restricció Stu I, que introdueix les mutacions a la zona de -227 que persegueix aquest clonatge. El següent pas és ajuntar els productes d'ambdues PCRs i realitzar una nova PCR emprant només els encebadors p2b i luc rev. Això dóna com a resultat el fragment p2b, que incorpora les mutacions a -227 (comprovables per la presència de la diana Stu I) i la part inicial del cDNA de la luciferasa. Seguidament, es digereix aquest fragment amb Asp718 i Nco I (situat al principi del cDNA de la luciferasa, abans d'on s'uneix luc rev) i el resultat es lliga amb un pGL3-Basic digerit de la mateixa manera.

. **pGL3-p2b-214:** generat a partir del pGL3-Trl full per PCR amb la DNA polimerasa Expand. Vàrem amplificar dos fragments solapants. El primer, per mitjà dels oligonucleòtids p2b i Lowmotif FII, dóna un fragment que va de l'inici de p2b (-345) fins a la zona de -214. El segon, per mitjà dels oligonucleòtids Upmotif FII i luciferasa revers (luc rev), dóna un fragment que va de la regió de -214 fins al principi del cDNA de la luciferasa. Els oligonucleòtids Upmotif FII i Lowmotif FII són en gran part solapants, però amplifiquen en sentits oposats (hibriden en cadenes oposades) i a més porten una diana per a l'enzim de restricció EcoRV, que introdueix les mutacions a la zona de -214 que persegueix aquest clonatge. El següent pas és ajuntar els productes d'ambdues PCRs i realitzar una nova PCR emprant només els encebadors p2b i luc rev. Això dóna com a resultat el fragment p2b, que incorpora les mutacions a -214 (comprovables per la presència de la diana Eco RV) i la part inicial del cDNA de la luciferasa. Seguidament, es digereix aquest fragment amb Asp718 i Hind III (situat al principi del cDNA de la luciferasa, abans d'on s'uneix luc rev) i el resultat es lliga amb un pGL3-Basic digerit de la mateixa manera.

MATERIALS I MÈTODES

. **pGL3-p2b-120(M7/8)**: generat a partir del pGL3-p2b per PCR amb la DNA polimerasa Expand (Roche). Vàrem utilitzar dos oligonucleòtids, M7/8Up i M7/8Low, que s'uneixen consecutivament a la seqüència entorn a -120 del promotor de *Trl* i que dirigeixen l'amplificació en sentit divergent. Aquests oligonucleòtids porten als seus extrems 5' cinc nucleòtids canviats respecte a la seqüència original, el que donarà lloc a 10 nucleòtids mutats al fragment amplificat, que consta de tot el plasmidi pGL3-p2b. Seguidament vam eliminar l'adenina final que deixa la polimerasa gràcies a la T4 DNA polimerasa i el resultat el vàrem lligar.

. **pGL3-p2b+41**: generat emprant la mateixa metodologia que en el cas anterior. Vam partir de pGL3-p2b i vàrem utilitzar com a encebadors els oligonucleòtids "+41 Mlu I Fwd" i "+41 Nhe I Rev" que porten a la seva part 5' diversos canvis nucleotídics, respecte a la seqüència del promotor de *Trl*, que generen unes dianes pels enzims de restricció Mlu I i Nhe I, respectivament.

. **pGL3 "-345/+71"(+41)**: generat utilitzant exactament la mateixa metodologia i els mateixos materials que en el cas anterior, però partint del plasmidi pGL3 "-345/+71".

. **pGL3-*Trl* vir**: conté els 2962 pb *upstream* del primer codó ATG del gen *Trl* de *Drosophila virilis*. Partint del plasmidi pTopo-*Trl* vir vàrem alliberar l'insert per digestió amb els enzims de restricció Acc65I i Xho I i el vam lligar a un pGL3 idènticament digerit.

. **pGL3-*eve***: conté 1824 pb de la regió promotora de la banda 2 del gen *even skipped* (*eve stripe 2*) de *Drosophila melanogaster*. Generat per la Dra. Sara Pagans (Tesi Doctoral, 2002).

C.1.3.3. Plasmidis per a l'expressió de proteïnes a bacteris

. **pGEX-KG (Amersham-Pharmacia-Biotech)**: permet l'expressió a bacteris de proteïnes fusionades pel seu extrem N-terminal a la Glutatió S-Transferasa (GST). També serveix per a l'expressió a bacteris de GST.

. **pGex-KG-TTK69**: conté el cDNA de TTK69 amb el cDNA de GST a N-terminal. Realitzat per la Dra. Sara Pagans (Tesi Doctoral, 2002).

. **pGex-KG-POZTTK69**: conté el cDNA corresponent al domini POZ de TTK69 amb el cDNA de GST a N-terminal. Vam digerir el pGex-KG-TTK69 amb Eco RI i Bpu 1102I, alliberant així tot el cDNA de TTK69, a excepció de la seqüència que correspon al seu domini POZ. Finalment, vam relligar el plasmidi.

. **pGex-KG-GAGA₅₁₉**: conté el cDNA de GAGA₅₁₉ fusionat en fase amb el cDNA de GST a N-terminal. Realitzat per la Dra. M^a Lluïsa Espinàs.

. **pET28a-His-Gro**: conté el cDNA de la proteïna Groucho, inserit a un pET28a. A partir del pGEX-3X-Gro, vàrem extreure l'insert digerint amb Bam HI, omplint els extrems amb el fragment Klenow de la polimerasa, i digerint amb Sma I. Aquest insert el vàrem lligar a un pET28a digerit amb Eco RI, amb els extrems omplerts amb Klenow.

C.1.3.4. Plasmidis per a l'expressió de proteïnes a cèl·lules eucariotes

. **pCMV-β-gal (Stratagene)**: conté el cDNA de la β-galactosidasa sota el control del promotor de Citomegalovirus (CMV). Utilitzat com a control per a normalitzar els valors luciferasa obtinguts a les transfeccions transitòries.

. **pActppa-GAGA₅₁₉**: conté el cDNA corresponent a la isoforma GAGA₅₁₉ de *Drosophila melanogaster*. Generat pel Dr. Àlex Vaquero (Tesi Doctoral, 2000).

. **pActppa-GAGA-A**: conté el cDNA corresponent a la isoforma GAGA-A de *Drosophila virilis*. A partir del plasmidi pTopo-GAGA-A, vàrem alliberar l'insert digerint amb els enzims de restricció Bam HI i Eco 32I i el vam lligar a un pActppa igualment digerit.

. **pActppa-GAGA₅₁₉ΔPOZ245**: plasmidi dissenyat per a expressar GAGA₅₁₉ però amb els primers 245 aminoàcids a N-terminal deletats. Generat pel Dr. Àlex Vaquero.

. **pActppa-GAGA₅₁₉ΔX1**: plasmidi dissenyat per a expressar GAGA₅₁₉ però amb una deleció de l'aminoàcid 122 al 208, que comprèn la part N-terminal de la regió X. Generat pel Dr. Jordi Bernués.

. **pActppa-GAGA₅₁₉ΔX2**: plasmidi dissenyat per a expressar GAGA₅₁₉ però amb una deleció de l'aminoàcid 222 al 303, que comprèn la part C-terminal de la regió X. Generat pel Dr. Jordi Bernués.

. **pActppa-GAGA-DBD**: plasmidi dissenyat per a expressar els aminoàcids del 310 al 391 de GAGA₅₁₉, que corresponen al domini d'unió a DNA. Vàrem extreure el DBD de GAGA del plasmidi pET14b-BDGAGA per mitjà d'una digestió amb Nde I, omplint els extrems amb Klenow, i finalment digerint amb Bam HI. Seguidament, vàrem lligar aquest fragment al pActppa digerit amb Eco RV i Bgl II.

MATERIALS I MÈTODES

. **pActppa-GAGA-DBD-EGFP:** dissenyat per a expressar el DBD de GAGA amb l'*Enhanced Green Fluorescent Protein* (EGFP) fusionada a C-terminal. Generat pel Dr. Jordi Bernués.

. **pECFP-Nuc-GAGA-DBD:** dissenyat per a expressar el DBD de GAGA₅₁₉ amb l'*Enhanced Cyan Fluorescent Protein* (ECFP) i tres NLS fusionats a N-terminal. Vàrem extreure el cDNA corresponent al DBD de GAGA digerint amb els enzims de restricció Nde I i Bam HI el plasmidi pET14b-BDGAGA i el vam inserir al pECFP-Nuc, el qual prèviament havíem digerit amb Bam HI i omplert els extrems protuberants resultants amb el fragment Klenow de la polimerasa (Roche).

. **pActppa-Lac I:** dissenyat per a expressar Lac I. Partint del pTopo-Lac I, vam extreure l'insert amb Nsi I, omplint els extrems protuberants amb T4 DNA Polimerasa, i finalment digerint amb Bam HI. L'insert el vam lligar al pActppa digerit amb Bam HI i Eco 32I.

. **pActppa-Lac I-GAGA-DBD:** dissenyat per a expressar el DBD de GAGA₅₁₉ amb Lac I fusionat a N-terminal. A partir del pTopo-Lac I-GAGA-DBD, vàrem extreure l'insert amb Nsi I, omplint els extrems protuberants amb T4 DNA Polimerasa, i finalment digerint amb Bam HI. L'insert el vam lligar al pActppa digerit amb Bam HI i Eco 32I.

. **pActppa-TTK69:** conté el cDNA de la proteïna TTK69. Realitzat per la Dr. Sara Pagans (Tesi Doctoral, 2002).

. **pActppa-His-Gro:** dissenyat per a expressar la proteïna Groucho amb una cua de 6 histidines a N-terminal. Partint del pET28a-Gro, vam extreure l'insert digerint amb Sal I, omplint els extrems amb el fragment Klenow de la polimerasa, i digerint amb Bgl II. L'insert el vàrem lligar al pActppa digerit amb Bam HI i Eco 32I.

. **pActppa-dCtBP-S:** conté el cDNA de la forma curta de dCtBP. Realitzat pel Dr. Jordi Bernués.

. **pActppa-dCtBP-L:** conté el cDNA de la forma llarga de dCtBP. Realitzat pel Dr. Jordi Bernués.

. **pcDNA3-dCtBP-S:** dissenyat per a expressar la forma curta de dCtBP sota el control del promotor CMV. Partint del pActppa-dCtBP-S, vam extreure l'insert digerint amb Xba I i el vàrem inserir al pcDNA3 igualment digerit.

. **pcDNA3-GAGA₅₁₉:** dissenyat per a expressar GAGA₅₁₉ sota el control del promotor CMV. Partint del pET14b-GAGA₅₁₉, vam extreure l'insert digerint amb Nde I i Esp I. Posteriorment vàrem omplir els extrems protuberants amb el fragment Klenow de la

polimerasa i vàrem lligar l'insert al pcDNA3 digerit amb Eco RI al que també vàrem omplir els extrems protuberants amb Klenow.

. **pcDNA3-GAGA Δ Q**: dissenyat per a expressar GAGA₅₁₉ sense el seu domini Q, sota el control del promotor CMV. A partir del pcDNA3-GAGA₅₁₉ digerim amb Eco RV, el que elimina el domini Q de GAGA, i relliguem el plasmidi, obtenint així una forma de GAGA₅₁₉ amb una deleció de 94 aminoàcids a C-terminal i, per tant, de 425 aminoàcids.

. **pcDNA3-GAGA₅₁₉ Δ POZ**: dissenyat per a expressar, sota el control del promotor CMV, GAGA₅₁₉ sense els primers 122 aminoàcids a N-terminal, corresponents al domini POZ. Realitzat pel Dr. Jordi Bernués.

C.1.3.5. Plasmidis per a la producció de dsRNA

. **pBS-RNAi-GFP**: plasmidi utilitzat per a produir dsRNA de GFP. Conté tot el cDNA de GFP inserit al pBlueScribe M13+. Realitzat per Glòria Palou.

. **pBS-RNAi-Psq**: plasmidi utilitzat per a produir dsRNA de *Pipsqueak* (Psq). Conté els primers 480 pb del cDNA de Psq. Vàrem extreure l'insert del pET21a-Psq (cedit amablement pel Dr. Alexander Schwendemann) amb Sac I i el vam lligar al pBluescribe M13+ digerit de la mateixa manera.

. **pBS-RNAi-dCtBP**: plasmidi utilitzat per a produir dsRNA de dCtBP. Conté 476 pb de dCtBP. Partint del pActppa-dCtBP-S, vàrem digerir aquest plasmidi amb Hind III (digereix a 269 pb de l'inici del cDNA) i Eco RI (digereix a 745 pb de l'inici del cDNA). Aquest insert el vàrem lligar al pBlueScribe M13+ digerit també amb Hind III i Eco RI.

C.1.3.6. Plasmidis per a la producció de línies de mosques transgèniques

. **pCaSpeR4-Trl full-GFP**: conté la regió promotora de *Trl*, de -3470 a +737, controlant l'expressió del cDNA de GFP. Realitzat pel Dr. Jordi Bernués.

. **pUAST-GAGA₅₁₉**: conté el cDNA de GAGA₅₁₉ sota el control d'un promotor amb llocs d'unió per a la proteïna GAL4. Realitzat per la Dra. Dori Huertas.

MATERIALS I MÈTODES

. **pUAST-GAGA Δ Q**: conté el cDNA de GAGA₅₁₉ amb una deleció que implica la pèrdua dels 94 aminoàcids més a C-terminal, que constitueixen el domini Q. Aquest cDNA està sota el control d'un promotor amb llocs d'unió per a la proteïna GAL4. Realitzat pel Dr. Jordi Bernués.

. **pUAST-GAGA₅₈₁**: conté el cDNA de GAGA₅₈₁ sota el control d'un promotor amb llocs d'unió per a la proteïna GAL4. Realitzat pel Dr. Jordi Bernués.

. **pUAST-GAGA₅₁₉DBDKO**: conté el cDNA de GAGA₅₁₉ amb una mutació puntual que provoca el canvi de la histidina 361 a prolina, cosa que aboleix la capacitat d'unió al DNA de GAGA. Aquest cDNA està sota el control d'un promotor amb llocs d'unió per a la proteïna GAL4. Realitzat pel Dr. Jordi Bernués.

. **pWIZ-RNAiGAGA**: conté un *hairpin* d'RNA que cobreix part dels dominis POZ i X de GAGA (que actua com a RNA d'interferència) sota el control d'un promotor amb llocs d'unió per a la proteïna GAL4. Realitzat per la Dra. M^a Lluïsa Espinàs segons les instruccions dels dissenyadors del pWIZ (Lee et al. 2003).

C.1.3.7. Plasmidis per a la producció de línies de cèl·lules S2 establiment transfectades

. **pMK33-CTAP**: conté la resistència a Hygromicina, que permet la selecció de les cèl·lules establiment transfectades, i el promotor de la metal·lotioneïna, que permet l'expressió, induïble per metalls pesants, de proteïnes amb un TAP (*tandem affinity purification*) a C-terminal. Dissenyat per Veraksa i col·laboradors (Veraksa et al. 2005).

. **pMK33-GAGA₅₁₉**: dissenyat per a l'obtenció d'una línia de cèl·lules S2 establiment transfectades que sobreexpressin GAGA₅₁₉ sota el control del promotor de la metal·lotioneïna. Partint del pET14b-GAGA₅₁₉, vàrem extreure l'insert corresponent al cDNA de GAGA₅₁₉ digerint amb Nde I i Esp I. Seguidament vam omplir els extrems protuberants amb el fragment Klenow de la polimerasa i vàrem lligar l'insert al pMK33-CTAP digerit amb Eco 32I. La presència del codó STOP al cDNA de GAGA₅₁₉ evita l'expressió del TAP a C-terminal, present a pMK33-CTAP.

. **pMK33-EGFP**: dissenyat per a l'obtenció d'una línia de cèl·lules S2 establiment transfectades que expressin EGFP sota el control del promotor de la metal·lotioneïna. Partint del pEGFP-N1, vàrem extreure l'insert corresponent al cDNA d'EGFP digerint amb Not I, seguidament omplint els extrems protuberants amb Klenow, i finalment digerint amb Bam HI.

MATERIALS I MÈTODES

L'insert el vam lligar al pMK33-CTAP digerit amb Bam HI i Eco 32I. La presència del codó STOP al cDNA de EGFP evita l'expressió del TAP a C-terminal, present a pMK33-CTAP.

. **pMK33-RNAiGAGA:** dissenyat per a l'obtenció d'una línia de cèl·lules S2 establiment transfectades que, sota el control del promotor de la metal·lotioneïna, expressin un *hairpin* d'RNA que cobreix part dels dominis POZ i X de GAGA i que s'incorporarà a la maquinària de l'RNA d'interferència. Partint del pWIZ-RNAiGAGA, vam extreure l'insert corresponent al *hairpin* de GAGA digerint amb Spe I i Xba I, i posteriorment vam omplir els extrems protuberants amb el fragment Klenow de la polimerasa. L'insert el vam lligar al pMK33-CTAP digerit amb Eco 32 I.

MATERIALS I MÈTODES

C.1.4. Oligonucleòtids

Per a la realització d'aquest treball s'han emprat oligonucleòtids sintetitzats per SIGMA. A la taula C.1 es fa una relació de tots els oligonucleòtids emprats i dels experiments en que s'han fet servir.

Nom de l'oligonucleòtid	Seqüència de 5' a 3'	Experiments en que s'ha utilitzat
p(-)	5' AGCTCGGTACCTATCGATAG 3'	Clonatge de pGL3 "-298/+737"
p2b	5' CGACTTTTTTGTGGTACCCTCTG 3'	Clonatge de: pGL3 "-345/+71", pGL3 "-345/+14", pGL3 "-345/-33", pGL3-p2b-303, pGL3-p2b-265, pGL3-p2b-303-265, pGL3-p2b-227 i pGL3-p2b-214
luciferasa reverse (luc rev)	5' GCGTAAGTGATGTCACC 3'	Clonatge de pGL3-p2b-303, pGL3-p2b-265, pGL3-p2b-303-265, pGL3-p2b-227 i pGL3-p2b-214
UpTrlvir	5' GAGCATTGCGCCACTGGACT 3'	Clonatge del promotor de Trl de D. virilis
LowTrlvir	5' GTCCAAATCCAACAGTTGTTCTC 3'	
UpGAGAvir	5' TGGATTGGACACGACAGCC 3'	
LowGAGAvir	5' GTGTTAAGGTGGAATGATTAG 3'	
FII(4-3)	5' CTTCTGCGGTACCTCTGGCAATGCC 3'	Clonatge de pGL3 "-298/+737"
Trl pr+71	5' TTTTTGTTTGTAGATCTGTTTTGG 3'	Clonatge de pGL3 "-345/+71"
Trl pr+14	5' TTTCAAGTATCTGGGTGATTTTGTG 3'	Clonatge de pGL3 "-345/+14"
Trl pr-33	5' GCTTAGATCTCTGTTGTTTGTCC 3'	Clonatge de pGL3 "-345/-33"
Down-303	5' CAGATTGATATCAGGAAGAGTATTTTAGG 3'	Clonatge de pGL3-p2b-303
Up-303	5' TCTTCTGATATCAATCTGGCAATGCC 3'	
Down-265	5' TGGCAGGAATCTTGAAGTACAGTTGTCGAG 3'	Clonatge de pGL3-p2b-265 i pGL3-p2b-303-265
Up-265	5' TTCAAGAATCTCTGCCAGGCAGGTCCTC 3'	
Down-227	5' CGCCTGTGAGGCTGCGCTATGAGAGAAGTGACC 3'	Clonatge de pGL3-p2b-227
Up-227	5' TTCTCTATAGCGCAGGCTGACAGGCGCGCAC 3'	
Lowmotif FII	5' GTGCTGGTGTGAGATATCATTGTTATCTGTGATCTATG 3'	Clonatge de pGL3-p2b-214
Upmotif FII	5' CAGATACAAATGATATCTCACACAGCAGCAAATAACC 3'	Clonatge de pGL3-p2b-214
M7/8Up	5'GTAGCCCGGAGCAGCGAAATTCGATC 3'	Clonatge de pGL3-p2b-120(M7/8)
M7/8Low	5' AGTTAGCGGCCAGTTTTCAGTTCTC 3'	
"+41 Mlu I Fwd"	5' CGACGCGTAGATAATTTCAAAAACC 3'	Clonatge de pGL3-p2b+41 i pGL3 "-345/+71"(+41)
"+41 Nhe I Rev"	5' GCTAGCAAAGAAAATTTCAAGTGGC 3'	
FI-1	5' TTTGAGCTTATTACAGAGTGG 3'	Footprinting genòmic amb DMS
FI-2	5' TCCGATAGTTATAGAACCGGAACC 3'	
FI-3	5' TAGAACCGGAACCTACGGATCCCG 3'	
FIIA-1	5' TACGGATCCCGTTGACC 3'	
FIIA-2	5' CACATGTGTTTTTTCATCCGTGAAG 3'	
FIIA-3	5' TTTTCATCCGTGAAGTCGGCCACG 3'	
FIIB-1	5' TCCATCATTCAATATTACCACC 3'	
FIIB-2	5' TTGCCGACTTTTTTGTCTTGCCG 3'	
FIIB-3	5' TTTGTCTTCCGCTTTGCTCTCTC 3'	
FIIC-1	5' ATACTTCTCTGCGCTCC 3'	
FIIC-2	5' ACAACTGTACTTCAAAGAGCGTTG 3'	
FIIC-3	5' TCAAAGAGCGTTGCCAGGCAGGTC 3'	
FIID-1	5' ATACACAGATACAGGCGC 3'	
FIID-2	5' CACCAGCAGCAAATAACCACC 3'	
FIID-3	5' CAAAATAACCACCACGAGTACAG 3'	
Linker small	5' GAATTCAGATC 3'	RT-PCRs
Linker large	5' GCGGTGACCCGGGAGATCTGAATTC 3'	
U6 upper	5' GCTTCGGCAGAACATATAC 3'	
U6 lower wt	5' AATGTGGAACGCTTACAG 3'	
TrlSUTR-70	5' TAAATCGTCGTGTTGGAGAC 3'	
Trlexo2+319	5' TGTTCAAGCACTGGGCAG 3'	Real Time RT-PCRs
Trl q fwd	5' AGGAAGAGCAAAGGGAGTTGG 3'	
Trl q rev	5' TTCTACTCTCTCCACGACG 3'	
U6 q fwd	5' CGATACAGAGAAGATTAGCATGGC 3'	
U6 q rev	5' GATTTGCGTGTATCCTTGC 3'	

Taula C.1. Relació dels oligonucleòtids emprats en aquest treball, indicant els experiments en que s'han fet servir.

C.1.5. Anticossos específics

En aquest treball s'han utilitzat diferents anticossos per als assajos de *Western blot* i immunolocalització. A la taula C.2 s'indiquen tots els anticossos específics emprats així com la seva procedència.

	Anticòs	Característiques	Fabricant	Experiments en que s'han utilitzat
Anticossos primaris	anti-GAGA(rata)	Anticòs policlonal de rata. Reconeix totes les isoformes de GAGA, així com les seves formes truncades	Generat al laboratori	<i>western blot</i> / immunolocalització
	anti-GAGA(conill)	Anticòs policlonal de conill. Reconeix totes les isoformes de GAGA, així com les seves formes truncades	Generat al laboratori	<i>western blot</i>
	anti-tubulina	Anticòs monoclonal de ratolí. Reconeix la específicament la β -tubulina i presenta una forta reacció creuada amb la β -tubulina de multitud d'espècies	Chemicon International (Ref. MAB3408)	<i>western blot</i>
	anti-TTK69	Anticòs policlonal de conill. Reconeix la proteïna Tramtrack69 (TTK69)	Generat al laboratori	<i>western blot</i>
	anti-Psq	Anticòs policlonal de conill. Reconeix les diverses isoformes de la proteïna Pipsqueak (Psq)	Cedit per Celeste Berg, Univ. of Washington, USA	<i>western blot</i>
	anti-sickle	Anticòs policlonal de conill. Reconeix la proteïna sickle	Cedit per E. Alnemri, Kimmel Cancer Institute. Philadelphia, USA	immunolocalització
	anti-Lac I	Anticòs monoclonal de ratolí. Reconeix la proteïna bacteriana Lac I, així com proteïnes de fusió que continguin el tag Lac I	Upsate (Millipore) (Ref. 05-503)	<i>western blot</i> / immunolocalització
	anti-GFP	Anticòs policlonal de conill. Reconeix específicament la Green Fluorescent protein (GFP)	Invitrogen (Ref. SKU# A-6455)	immunolocalització
anti-Flag M2	Anticòs monoclonal de ratolí. Reconeix l'epítip Flag	Sigma (Ref. F 3165)	<i>western blot</i>	
Anticossos secundaris	anti-conill-HRP	Reconeix la cadena pesada dels anticossos de conill. Està conjugat a la peroxidasa de rabe (<i>horseradish peroxidase</i> , HRP)	Amersham Biosciences (Ref. NA 934)	<i>western blot</i>
	anti-ratolí-HRP	Reconeix la cadena pesada dels anticossos de ratolí. Està conjugat a la peroxidasa de rabe (<i>horseradish peroxidase</i> , HRP)	Jackson ImmunoResearch (Ref. 715-035-150)	
	anti-rata-HRP	Reconeix la cadena pesada dels anticossos de rata. Està conjugat a la peroxidasa de rabe (<i>horseradish peroxidase</i> , HRP)	Jackson ImmunoResearch (Ref. 712-035-150)	
	anti-rata Cy2	Reconeix la cadena pesada dels anticossos de rata. Està conjugat a Cy2	Jackson ImmunoResearch (Ref. 712-225-150)	immunolocalització
	anti-rata Cy3	Reconeix la cadena pesada dels anticossos de rata. Està conjugat a Cy3	Jackson ImmunoResearch (Ref. 712-165-150)	
	anti-rata Cy5	Reconeix la cadena pesada dels anticossos de rata. Està conjugat a Cy5	Jackson ImmunoResearch (Ref. 712-175-150)	
	anti-conill Cy2	Reconeix la cadena pesada dels anticossos de conill. Està conjugat a Cy 2	Jackson ImmunoResearch (Ref. 711-225-152)	
	anti-conill Cy5	Reconeix la cadena pesada dels anticossos de conill. Està conjugat a Cy 5	Jackson ImmunoResearch (Ref. 711-175-152)	

Taula C.2. Relació dels anticossos emprats en aquest treball, indicant les seves característiques, el seu origen i el tipus d'experiment on s'han utilitzat.

MATERIALS I MÈTODES

C.1.6. Cèl·lules eucariotes

Per a la realització d'aquest treball s'han fet servir les següents línies de cèl·lules eucariotes en cultiu:

S2 (també anomenades SL2).- Línia cel·lular originaria de *Drosophila melanogaster* provinent d'embrions tardans de 20 a 24h, establerta per Imogen Schneider (Schneider 1972). Són cèl·lules d'aspecte esfèric que creixen tant enganxades a la superfície del recipient en que es troben com en suspensió.

- Línies de cèl·lules S2 establiment transfectades generades durant la realització d'aquest treball són: **S2-GAGA**, **S2-EGFP** i **S2-RNAiGAGA**. Expressen, respectivament, GAGA₅₁₉, EGFP i un *hairpin* d'RNA que cobreix part dels dominis POZ i X de GAGA, que actua com a RNA de doble cadena, incorporant-se a la maquinària de l'RNA d'interferència. L'expressió és induïble, sota el control del promotor de la metal·lotioneïna.

WR DV1.- Línia cel·lular originaria de *Drosophila virilis* provinent d'embrions, establerta per Imogen Schneider (no publicat). Són cèl·lules d'aspecte esfèric que creixen tant enganxades a la superfície del recipient en que es troben com en suspensió.

HeLa.- Línia cel·lular d'origen humà, procedent de cèl·lules epidermoides de carcinoma de cèrvix humà, cultivades inicialment per George O. Gey (Scherer et al. 1953). Tenen aspecte de cèl·lules epidèrmiques i creixen enganxades a la superfície del recipient en que es troben, formant una monocapa.

C.1.7. Stocks de mosques

C.1.7.1. Cromosomes balancejadors i marcadors

Els balancejadors són cromosomes que presenten un elevat nombre de reordenacions i inversions que impedeixen que es produeixi recombinació. Normalment, els balancejadors contenen algun marcador que presenta un fenotip particular i que permet que siguin identificats respecte el cromosoma salvatge. D'altra banda, també disposem de marcadors que no s'associen a un cromosoma balancejat. Tots aquests marcadors s'utilitzen per al manteniment d'*stocks* de mosques, així com per al mapatge de transgens. Tots ells,

cromosomes balancejadors i marcadors, són letals en homozigosi. A la taula C.3 s'indiquen els diferents balancejadors i marcadors utilitzats al llarg d'aquest treball.

Balancejador	Cromosoma	Reorganitzacions	Marcador associat	Fenotip
CyO	II	In(2LR)O, Cy[1] dp[lvl] pr[1] cn[2]	Cy	Ales corbades
TM3	III	In(3LR)TM3, kni[ri-1] p[p] sep[1] I(3)89Aa[1] Ubx[bx-34e] e[1]	Sb	Quetes toràciques curtes
TM6B	III	In(3LR)TM6B, Antp[Hu] e[1]	Hu, Tb	Quetes a les espatlles i larves arrodonides

Marcador	Cromosoma	Fenotip
Irregular face (If)	II	Ulls petits i rugosos
Lyra (Ly)	III	Ales tallades pels marges
MKRS	III	Quetes curtes

Taula C.3. Balancejadors i marcadors genètics de *Drosophila* utilitzats en aquest treball.

C.1.7.2. Soques de *Drosophila melanogaster* utilitzades

En aquest treball s'han utilitzat també diferents soques de *Drosophila melanogaster*. A la taula C.4 s'indiquen els diferents *stocks* utilitzats, així com les característiques que presenten i la seva procedència (moltes provenen de l'*Stock Center* de Bloomington).

Stock	Cromosoma	Característiques	Referència
<i>white</i> ⁻	X	<i>Wild type</i> . Té una mutació al gen <i>white</i> que elimina tota la proteïna. Ulls blancs.	<i>w</i> ¹¹¹⁸
<i>w</i> ⁻ ; If/CyO;Ly/TM3	Doble balancejat	Presenta balancejadors i marcadors als cromosomes II i III sobre un fons <i>white</i> ⁻ .	B-4911
<i>w</i> ⁻ ; If/CyO;MKRS/TM6B	Doble balancejat	Presenta balancejadors i marcadors als cromosomes II i III sobre un fons <i>white</i> ⁻ .	B-3703
ptcGal4	II	Expressa Gal4 sota el control del promotor <i>patched</i> .	Cedides pels Drs. J. Casanova, M. Llimargas i M. Milán
prdGal4/TM3	III	Expressa Gal4 sota el control del promotor <i>paired</i> .	
dppGal4/TM6B	III	Expressa Gal4 sota el control del promotor <i>decapentaplegic</i> .	
MS1096Gal4	X	Expressa Gal4 sota el control del promotor MS1096.	
ActinaGal4/CyO	II	Expressa Gal4 sota el control del promotor <i>actina</i> .	
UAS-GAGA ₅₁₉	II/III	Expressa GAGA ₅₁₉ sota el control d'uns llocs d'unió de Gal4.	Generades per a la realització d'aquesta Tesi Doctoral
UAS-GAGA ₅₈₁	II/III	Expressa GAGA ₅₈₁ sota el control d'uns llocs d'unió de Gal4.	
UAS-GAGA ₅₁₉ ΔQ	II/III	Expressa GAGA ₅₁₉ ΔQ sota el control d'uns llocs d'unió de Gal4.	
UAS-GAGA ₅₁₉ DBDKO	II/III	Expressa GAGA ₅₁₉ DBDKO sota el control d'uns llocs d'unió de Gal4.	
UAS-RNAiGAGA	X/II/III	Expressa l'RNAi de GAGA sota el control d'uns llocs d'unió de Gal4.	
Trl full-GFP	X/II/III	Expressa GFP sota el control de la versió més llarga clonada del promotor <i>Trl</i>	

Taula C.4. *Stocks* de mosques emprats per a la realització d'aquest treball. S'indiquen el cromosoma, les característiques principals i la procedència.

C.2. Mètodes

C.2.1. Tècniques de manipulació de DNA plasmídic

C.2.1.1. Tècniques de DNA recombinant

Durant la realització d'aquest treball s'han seguit les tècniques de DNA recombinant descrites per Sambrook (Sambrook et al. 1989).

C.2.1.2. Transformació de cèl·lules competents

Incubar les cèl·lules competents amb 5 – 20 µL de la lligació durant 20 minuts en gel. Fer un xoc tèrmic de dos minuts a 42°C en un bany. Recuperar les cèl·lules afegint 300 µL de medi LB sense antibiòtic i incubant-les durant 30 minuts a 37°C (temps que triguen els bacteris a produir la resistència a l'antibiòtic). A la flama, plaquejar la transformació en plaques d'LB amb l'antibiòtic adient. Deixar tota la nit a 37°C.

LB (Luria-Bertani)

10 g/L triptona
5 g/L extracte de llevat
10 g/L NaCl
1 mM NaOH

C.2.1.3. Extracció de DNA plasmídic pel mètode de la lisi alcalina ràpida

Inocular una colònia aïllada de bacteris transformats amb el plasmidi d'interès a 4 mL de medi LB amb antibiòtic. Incubar tota la nit a 37°C en agitació. Passar aproximadament 1 mL del cultiu a un tub eppendorf i centrifugar a 9000 rpm (centrífuga eppendorf) durant 1 minut. Treure el sobrenedant amb una pipeta Pasteur i resuspendre en 100 µL de GTE vortexant. Afegir 200 µL de solució 0,2 N NaOH, 1% SDS, invertir dos o tres cops i deixar en gel durant 5 minuts. Afegir 150 µL d'acetat sòdic 3 M pH 5,0, invertir dos o tres cops i deixar en gel 5 minuts més. Centrifugar 5 minuts a 14000 rpm en una centrífuga eppendorf i passar el sobrenedant a un tub eppendorf nou. Afegir 450 µL de fenol/cloroform, vortexar i centrifugar 2 minuts a 14000 rpm. Recuperar la fase aquosa (superior) i afegir-hi 1 mL d'etanol 100% fred. Invertir dos o tres cops i centrifugar 5 minuts a 14000 rpm. Treure el sobrenedant i rentar el *pellet* amb 300

µL d'etanol 70% fred. Centrifugar 3 minuts a 14000 rpm, descartar el sobrenedant i deixar eixugar el *pellet* a 37°C. Resuspendre el *pellet* en 50 µL de TE amb 1 µg d'RNAsaA. Digerir l'RNA amb l'RNAsaA durant 15 minuts a 37°C.

TE	GTE
10 mM Tris HCl pH 8,0	50 mM glucosa
1 mM EDTA	10 mM EDTA
	25 mM Tris HCl pH 8,0

C.2.1.4. Maxipreparació de DNA plasmídic

Inocular 500 mL de medi LB (amb antibiòtic) amb un precultiu de 6-10 h de la soca bacteriana transformada amb el plasmidi que es vol amplificar. Incubar tota la nit a 37°C en agitació. Recollir en pots de 250 mL i centrifugar en una centrífuga Beckman a 6000 g, 10 minuts. Descartar el medi i procedir a l'extracció i purificació del DNA plasmídic amb el kit Qiagen® Plasmid Maxi Kit, segons les instruccions.

Deixar assecar el *pellet* de DNA i resuspendre'l en 500 mL d'aigua milliQ. Per a conèixer la concentració del DNA ([DNA]) obtingut, es dilueix 1/100 la mostra i llegim l'absorbància a 260 nm en un espectrofotòmetre UV-1603 (Shimadzu). Obtenim la [DNA] segons la següent relació:

$$[\text{DNA}] \text{ mg/mL} = \text{Abs } 260\text{nm} \times 5$$

Es comproven la integritat i puresa del DNA carregant 200 ng de la mostra amb tampó de càrrega Orange G a un gel 1 % agarosa en 1X TBE.

TBE (10X)	Orange G
90 mM Tris base	50% glicerol
24 mM EDTA	20 mM EDTA
0,8 M àcid bòric	Orange G

C.2.2. Mètodes de treball amb cèl·lules en cultiu

C.2.2.1. Manteniment de cultius cel·lulars

Les línies cel·lulars de *Drosophila* es mantenen en medi Schneider (*Schneider's insect Medium*, Sigma) suplementat amb un 10% de sèrum boví fetal (FBS, GIBCO) i un 1% de Penicil·lina/Streptomicina (GIBCO). Es mantenen a una densitat entre 1 i $5 \cdot 10^6$ cèl·lules/mL, en una estufa a 25°C. Dos cops per setmana, es passa 1 mL de cèl·lules a un flascó de cultiu

MATERIALS I MÈTODES

(Corning) de 25 cm², amb 4,5 mL de medi nou. No és necessari tripsinitzar-les, ja que aquestes cèl·lules són semiadherents.

La línia cel·lular HeLa es manté en medi DMEM (Dulbecco's Modified Eagle Medium, GIBCO) suplementat amb un 10% de sèrum boví fetal (FBS, GIBCO) i un 1% de Penicil·lina/Streptomycina (GIBCO), en una estufa a 37°C amb un 5% de CO₂. Al créixer enganxades a la superfície del flascó de cultiu en forma de monocapa no és necessari contar-les i per tant es mantenen fins que arriben a un 70-80 % de confluència, moment en el qual es tripsinitzen i es passen a un flascó nou diluint-les 1/10 en medi nou. Per tripsinitzar-les s'aspira el medi de cultiu eliminant així les cèl·lules en suspensió (mortes), es renta amb PBS 1X estèril, s'afegeix tripsina-EDTA, es deixen 2 minuts a 37°C per a que actuï desenganxant les cèl·lules i s'afegeix medi per a neutralitzar la tripsina.

<u>PBS 1X</u>	<u>Tripsina-EDTA</u>	<u>TD (500 mL)</u>
125 mM NaCl	18,6 mL TD	17,125 mL NaCl 4 M
17 mM Na ₂ HPO ₄	1 mL Tripsina 0,5%	625 µL KCl 4 M
8 mM NaH ₂ PO ₄	0,4 mL EDTA 25 mM	0,087 g Na ₂ HPO ₄ ·2H ₂ O
pH 7,3	Esterilitzar per filtració	10 mL Tris-HCl 1 M pH 7,4

C.2.2.2. Congelació de cèl·lules de *Drosophila*

D' un flascó de 175 cm² amb cèl·lules creixent en fase exponencial (20-30·10⁶ cèl·lules totals), es recullen les cèl·lules i es centrifuguen a un tub falcon de 15 µL. Es descarta el sobrenedant i es resuspenen les cèl·lules en 1350 µL d'FBS. Es passen les cèl·lules a un criotub i s'afegeixen, gota a gota, 150 µL de DMSO (Dimetilsulfòxid). Ràpidament es congelen les cèl·lules, deixant-les de 2 a 4 h a -20°C. Seguidament es passa el criotub a -80°C i es deixa tota la nit. Finalment, es guarda en N₂ líquid.

C.2.2.3. Descongelació de cèl·lules de *Drosophila*

Es treu el criotub amb cèl·lules del N₂ líquid i es descongela ràpidament. Es passen les cèl·lules a un tub falcon de 15 mL que contingui 5 mL de medi de cultiu. Centrifugar 5 minuts a 400 g, treure el sobrenedant i realitzar un segon rentat amb medi nou. Finalment, es resuspenen les cèl·lules en 5 mL de medi de cultiu nou, es passen a un flascó de cultiu de 25 cm² i es deixen a una estufa a 25°C.

C.2.2.4. Comptatge de cèl·lules

Les cèl·lules es recullen en un falcon i es posen 10 μL a cada camp de la cambra de Neubauer. Cada camp està constituït per una quadrícula de 4 x 4 (16 quadrats). Es compten les cèl·lules de cada camp i es fa un promig. Per conèixer la concentració de cèl·lules es multiplica el nombre de cèl·lules d'un camp (els 16 quadrats) per 10^4 . S'obté, d'aquesta manera, el nombre de cèl·lules per mL de medi.

C.2.2.5. Transfecció pel mètode del fosfat càlcic

Aquest mètode facilita l'entrada de DNA a les cèl·lules mitjançant la formació de precipitats de DNA-fosfat. A partir de flascons de cultiu de 175 cm^2 es desenganxen les cèl·lules amb ajuda d'una pipeta i es compten utilitzant una cambra de Neubauer. Per a cada transfecció s'inoculen $3 \cdot 10^6$ cèl·lules per placa de cultiu de 60 mm^2 (Corning) en un volum final de 5 mL i s'incuben 24 h a 25°C.

Es transfecten 20 μg de DNA per placa (obtingut a partir de maxipreparacions amb el kit de Qiagen) en un volum final de 50 μL . Com que cada transfecció es fa per duplicat, les barreges s'escalen a 2,5 transfeccions. Es col·loquen les quantitats adequades de cada plasmidi, afegint si fa falta plasmidis buits per arribar als 20 μg de DNA, i s'ajusta el volum a 125 μL (50 μL x 2,5 transfeccions) amb H_2O estèril. A això se li afegeix 1 mL de CaCl_2 0,25 M (esterilitzat per filtració). Aquesta barreja s'afegeix gota a gota a un tub falcon de 15 mL al qual prèviament se li ha afegit 1 mL de HeBs 2X (esterilitzat per filtració). Aquest tub falcon s'ha de mantenir votexant suament mentre afegim gota a gota la barreja del CaCl_2 i els plasmidis. D'aquesta manera es produeixen uns petits precipitats d'aspecte blanquinós. Aquests s'incuben 35 minuts a 25°C i, un cop passat el temps, es resuspenen els precipitats pipetejant varies vegades i s'afegeixen 850 μL d'aquesta barreja per placa, gota a gota i procurant escampar-ho homogèniament. Posteriorment, les plaques es deixen a 25°C. Per a les transfeccions transitòries, les plaques s'incuben usualment 48h a 25°C.

HeBs 2X

NaCl 250 mM

KCl 9 mM

Na_2HPO_4 1,5 mM

D-Glucosa 10 mM

HEPES 50 mM pH 7,1

MATERIALS I MÈTODES

C.2.2.6. Preparació d'extractes cel·lulars per a assajos d'activitat luciferasa i β -gal

Transcorregut el temps necessari després de la transfecció (usualment 48h), es recullen les cèl·lules i es centrifuguen a 400 g. Els pellets cel·lulars es renten dos cops resuspenent-los en PBS 1X i centrifugant a 400 g. Seguidament, els pellets cel·lulars es resuspenen en 250 μ L del tampó de lisi del kit " β -Gal reporter Gene Assay" (Boehringer-Manheim) i s'incuben 30 minuts a temperatura ambient. Transcorregut aquest temps, es centrifuguen els lisats 5 minuts a 14000 rpm a una centrífuga eppendorf. Els sobrenedants es transfereixen a tubs eppendorf nous de 1,5 mL i es mantenen en gel per continuar amb els assajos d'activitat luciferasa i β -gal.

C.2.2.7. Assaig d'activitat luciferasa

Es col·loquen 25 μ L de reactiu luciferasa diluït 1:2 del kit *Luciferase assay* (Promega) en tubs de poliestiré de 3 mL. Es deixa temperar el reactiu, protegint-lo de la llum, uns 15-30 minuts abans de realitzar l'assaig. S'afegeixen 15 μ L de cada extracte (provinents de la lisi amb el tampó de lisi del kit " β -Gal reporter Gene Assay") i es mesura la seva activitat luciferasa en un luminòmetre Sirius (Berthold). Per reduir els errors deguts al pipeteig, la medicació de cada mostra es realitza per duplicat, tenint en compte després la mitjana d'aquests valors.

C.2.2.8. Assaig d'activitat β -galactosidasa

Es col·loquen 100 μ L de reactiu substrat del kit " *β -Gal reporter Gene Assay*" en tubs de poliestiré de 3 mL. A continuació s'afegeixen 50 μ L de cada extracte, es cobreix amb parafilm per evitar l'evaporació i s'incuben 1h a temperatura ambient. Finalment, s'afegeixen 50 μ L del reactiu d'iniciació del kit i es mesura l'activitat β -galactosidasa en un luminòmetre Sirius (Berthold). Per reduir els errors deguts al pipeteig, la medicació de cada mostra es realitza per duplicat, tenint en compte després la mitjana d'aquests valors.

C.2.2.9. Tractament de cèl·lules S2 amb TSA i Nicotinamida

Vàrem dur a terme tractaments amb TSA (Tricostatina-A) i Nicotinamida a fi de comprovar el seu efecte sobre l'activitat transcripcional del promotor de *Trl*, així com sobre

l'efecte que provoca la sobreexpressió de GAGA i dCtBP, en aquest promotor. Vàrem transfectar transitòriament cèl·lules S2 tal qual s'explica a l'apartat C.2.2.5, amb els plasmidis adients. Les transfeccions es van fer per triplicat: sense tractar, tractades amb TSA i tractades amb Nicotinamida. 24 h després de la transfecció vàrem tractar les cèl·lules corresponents amb TSA, amb nicotinamida o no les vam tractar, i 48 h després de transfectar es van fer els assajos luciferasa i β -gal, per determinar l'activitat del promotor de *Trl*. Pel tractament amb TSA es va utilitzar un *stock* a 33 μ M diluït en etanol, afegint el necessari per arribar a una concentració final de 300 nM al cultiu cel·lular. Pel tractament amb Nicotinamida es va utilitzar un *stock* a 1 M diluït en H₂O, afegint el necessari per arribar a una concentració final de 5 mM al cultiu cel·lular.

C.2.2.10. Producció de línies cel·lulars establement transfectades

A partir de cèl·lules S2 en cultiu, vàrem establir tres línies cel·lulars establement transfectades. Per a la producció d'aquestes línies es van utilitzar el plasmidis pMK33-GAGA₅₁₉, pMK33-EGFP i pMK33-RNAiGAGA. Aquests plasmidis estan realitzats a partir del pMK33-CTAP (veure figura C.1). Conté la resistència a Hygromicina, que permet la selecció de les cèl·lules establement transfectades, i el promotor de la metal·lotioneïna, que permet l'expressió, induïble per metalls pesants, de la proteïna clonada, fusionada a un TAP (*tandem affinity purification*) a C-terminal. L'expressió d'aquest TAP no es dóna en pMK33-GAGA₅₁₉ i pMK33-EGFP perquè al clonar aquests cDNAs es va conservar el seu codó STOP original.

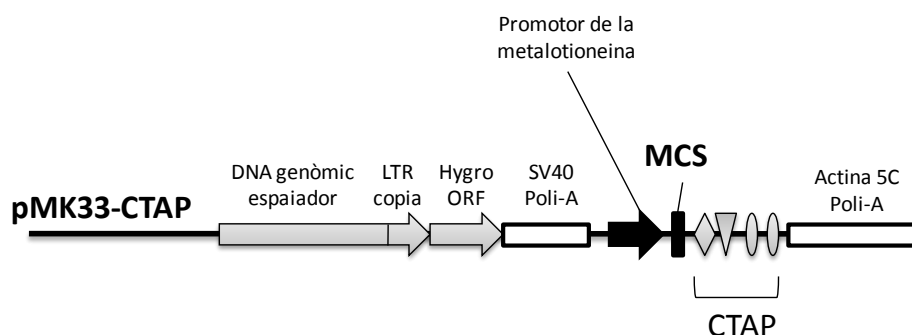


Figura C.1. Esquema del plasmidi pMK33-CTAP (Extret de Veraksa et al. 2005).

El primer pas consisteix en sembrar $3 \cdot 10^6$ de cèl·lules en un flascó de cultiu de 25 cm² amb 5 mL de medi Schneider suplementat, i transfectar les cèl·lules pel mètode del fosfat càlcic

MATERIALS I MÈTODES

(veure apartat C.2.2.5), amb 20 µg del plasmidi d'interès. A les 24 h recuperem les cèl·lules i les centrifuguem a 100 g durant 5 minuts. El *pellet* el resuspenem en 5 mL de medi Schneider suplementat fresc i tornem les cèl·lules al mateix flascó de cultiu anterior. Passades 48 h més, repetim el pas anterior, però en lloc de resuspendre en medi Schneider suplementat normal, resuspenem en 5 mL de medi Schneider + 10% FBS + 300 µg/mL d'Hygromicina-B (Sigma) esterilitzada per filtració, i tornem les cèl·lules al mateix flascó de cultiu anterior. Aquest últim pas es repeteix cada 5-7 dies fins que les cèl·lules comencin a proliferar de manera visible (al voltant de 3 setmanes), moment en el que podem passar-les a un flascó nou i fer créixer el cultiu fins a tenir suficients cèl·lules per a congelar en N₂ líquid. A partir d'aquí la línia cel·lular ja està establerta i per tant, les cèl·lules han integrat múltiples còpies dels plasmidis transfectats. En aquest punt ja es pot substituir l'Hygromicina-B per l'antibiòtic normal (Penicil·lina/Estreptomicina). Convé tornar a utilitzar l'Hygromicina-B com a antibiòtic, almenys durant un canvi de medi cada dos mesos, per evitar que les cèl·lules perdin capacitat d'expressar la nostra construcció.

C.2.2.11. Inducció de les cèl·lules establement transfectades amb

CuSO₄

Per induir l'expressió de les còpies transfectades i integrades de GAGA₅₁₉, EGFP i l'RNAi de GAGA, que hem integrat de manera estable al genoma de cèl·lules S2, afegim CuSO₄ (Sigma) esterilitzat per filtració, a una concentració final de 500 µM al medi de cultiu, durant el temps corresponent, depenent de l'experiment.

C.2.3. Tècniques d'anàlisi genètic de *Drosophila*

Les mosques utilitzades en aquest treball han estat mantingudes a 18°C i 25°C en cambres climàtiques. En la manipulació i el creixement de les mosques s'han seguit les tècniques descrites per Sullivan, Ashburner i Hawley (Sullivan et al. 2000). Les línies transgèniques generades al llarg d'aquest treball s'han realitzat amb el suport tècnic d'Esther Fuentes Díaz.

Per a l'anàlisi fenotípic de les mosques, a més de l'observació directa, hem procedit a muntar les ales per a l'observació i presa de fotografies amb un microscopi òptic Eclipse E-600 (Nikon) amb una càmera DXM 1200F (Nikon) acoblada.

Pel muntatge de les ales, primer es deixen els individus tota la nit en medi SH (glicerol i etanol en proporció 1:3 respectivament). Al dia següent col·loquem les mosques en PBS fred i s'extreuen les ales col·locant-les a un portaobjectes amb medi Fauré (veure figura C.2). Seguidament es col·loca un cobreobjectes i es cobreix amb paper Watmann 3mm. Finalment es deixa tota la nit amb un pes lleuger a sobre, per assecar l'excés de medi.

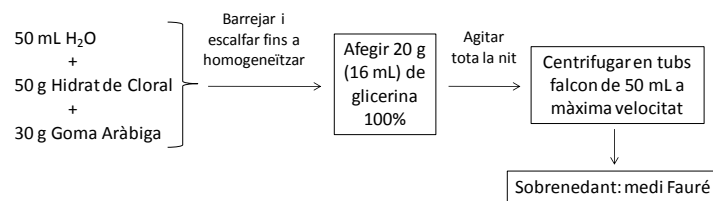


Figura C.2. Mètode de preparació del medi Fauré.

C.2.4. Tècniques d'anàlisi de DNA

C.2.4.1. Reacció en Cadena de la Polimerasa (PCR)

La tècnica de PCR (*Polymerase Chain Reaction*) permet amplificar regions de DNA a partir de mostres de DNA. Aquesta tècnica consisteix en un procés cíclic de desnaturalització del DNA i d'amplificació de les dues cadenes resultants, de manera que a cada cicle d'amplificació el nombre de cadenes de DNA es duplica.

Aquesta reacció en cadena és possible gràcies al descobriment de les polimerases termostables, encarregades de la replicació del DNA d'organismes adaptats a altes temperatures, com el bacteri *Thermus aquaticus*. Aquestes polimerases mostren una activitat òptima a elevades temperatures, característica imprescindible per assegurar l'especificitat de la reacció.

Per a que l'amplificació tingui lloc, cal que uns encebadors (oligonucleòtids) s'uneixin als extrems de la seqüència que es vol amplificar, de manera que cada encebador és complementari a una de les cadenes de DNA. Per a que aquests encebadors s'uneixin de forma específica a la seva seqüència complementària de la mostra, es realitza un aparellament (*annealing*) a una temperatura elevada, de manera que s'eviten les unions inespecífiques. Un cop els encebadors s'han unit a la seva seqüència complementària, la polimerasa pot començar

MATERIALS I MÈTODES

el procés d'amplificació, que és dependent d'ATP i Mg^{2+} , afegint desoxiribonucleòtids a l'extrem 3' de l'encebador. Aquest procés es produeix de forma cíclica gràcies a un aparell termociclador, que condueix canvis molt ràpids de temperatura de la mostra.

A continuació (veure taula C.5) es detallen els reactius i el procés d'una reacció de PCR típica. S'ha de tenir en compte que per a cada parell d'encebadors s'ha d'ajustar la temperatura d'aparellament o *annealing* i que variar aquesta temperatura així com la concentració de $MgCl_2$ a la reacció pot fer variar la seva especificitat i eficiència.

Reactius	Cicles
Mostra de DNA, de l'orde de nanograms	2' a 95°C ← <i>Hot Start</i>
Tampó de la reacció (1X)	30X { 1' a 95°C ← desnaturalització 1' a 55°C ← <i>annealing</i> 1' s 72°C ← amplificació
$MgCl_2$ 2 mM	
dNTPs 0,4 mM	
Encebadors 0,5 pmol/ μ L, cada un	10' a 72°C ← amplificació final
DNA polimerasa 0,05 U/ μ L	

Taula C.5. Reacció de PCR en unes condicions estàndard.

C.2.4.2. Extracció de DNA genòmic de cèl·lules

En aquest treball s'ha extret DNA genòmic de cèl·lules S2 per a la realització dels controls *in vitro* dels assajos de *footprinting*. A més, també s'ha extret DNA genòmic de cèl·lules WR DV1 per a clonar el promotor de *Trl* de *D. virilis*.

Partint de $5 \cdot 10^7$ cèl·lules, es centrifuguen a 400 g, 10 minuts. Es renta el *pellet* dos cops amb 10 mL de PBS 1X, centrifugant a 400 g, 10 minuts i descartant el sobrenedant. Resuspendre el *pellet* en 4 mL de tampó de lisi (idèntic al utilitzat als *footprintings*, veure més endavant, apartat C.2.4.4.A), es vortexa la mostra i s'incuba a 37°C durant 3 h. Es realitzen 3 extraccions amb 1 volum (4 mL) de fenol/cloroform (1:1) vortexant la mostra i centrifugant-la 10 minuts a unes 4000 g. Recuperar l'última fase aquosa en un tub de centrífuga de 50 mL i precipitar el DNA afegint: 1 volum (4 mL) d'acetat amònic 7,5 M + 5 volums (20 mL) d'etanol 100% (fred). Es vortexa i es centrifuga 30 minuts a 20000 g, a 4°C. Es renta el *pellet* amb 5 mL d'etanol 70% (fred), vortexant i centrifugant 10 minuts a 20000 g i 4°C. Es resuspèn el *pellet* en 300 μ L d' H_2O milliQ i es passa a un tub eppendorf. Es torna a precipitar la mostra afegint: 30 μ L NaCl 4M + 3 μ L acrilamida 0,25% + 1mL etanol 100% (fred). Es vortexa i es centrifuga (a una centrífuga eppendorf) a 14000 rpm 15 minuts, a 4°C. Es renta el *pellet* amb 1 mL d'etanol 70%

(fred), vortexant i centrifugant a 14000 rpm 5 minuts. Finalment, s'asseca el *pellet* i es resuspèn en 500-1000 μL d' H_2O milliQ. La concentració obtinguda es mesura a l'espectrofotòmetre UV-1603 (Shimadzu), obtenint la [DNA] tal com s'explica a l'apartat C.2.1.4.

C.2.4.3. Extracció de DNA genòmic d'embrions

En aquest treball s'ha extret DNA genòmic d'embrions control per a la realització dels controls *in vitro* dels assajos de *footprinting*.

Es recullen els embrions de 0-24 h, arribant a un volum d'uns 100 μL d'embrions. Immediatament, es decorionen tractant-los amb hipoclorit de sodi (lleixiu) aproximadament al 5% durant 2 minuts i rentant-los amb PBS 1X. En acabat, es resuspensen els embrions amb 0,5 mL de tampó de lisi (idèntic al utilitzat als *footprintings*, veure més endavant, apartat C.2.4.4.A) i es disgreguen les seves cèl·lules amb ajuda d'un petit èmbol de mida eppendorf (Pellet Pestle®), es vortexen i es deixen 3 h a 50°C. A partir d'aquí es continua de la mateixa manera que per a realitzar extractes de DNA genòmic de cèl·lules (veure apartat anterior).

C.2.4.4. *Footprinting* genòmic *in vivo* amb DMS

Per analitzar el grau d'ocupació *in vivo* de les regions FII i FIII del promotor de *Trl*, vàrem realitzar *footprintings* genòmics amb DMS (Dimetilsulfat) *in vivo*, tant a cèl·lules S2 en cultiu, com a embrions transgènics de *Drosophila melanogaster*. El DMS és un compost que metila les purines, amb una marcada preferència per la guanina. Així, el *footprinting* amb DMS es basa en un tractament limitat amb DMS on, idealment, es metili només una guanina per cadena de DNA, a la regió a analitzar. Posteriorment, es depurinen les bases metilades tractant la mostra amb piperidina, fet que parcialment trenca les cadenes de DNA. Partint d'aquesta mostra, s'amplifiquen els fragments de DNA resultants per *linked-mediated* PCR (LM-PCR), amb oligonucleòtids específics de la zona que es vol analitzar, i els fragments resultants es marquen radioactivament als extrems, per *primer extension* amb encebadors amb fòsfor 32 (P^{32}) als seus extrems 5'. Finalment, per electroforesi en un gel de seqüència, separem els fragments de DNA resultants per mida, i al exposar aquest gel a una placa autoradiogràfica s'obté el patró de bandes que correspon a la seqüència de citosines de la regió analitzada. Si algun factor està situat sobre el DNA en el moment del tractament amb DMS, probablement dificultarà l'accés

MATERIALS I MÈTODES

del DMS a les guanines de la zona, el que es reflectirà en una zona lliure de bandes al gel de seqüència final. Aquesta zona lliure de bandes l'anomenem "protecció". Eventualment, la presència de factors units al DNA poden provocar un canvi de conformació de la doble cadena que faciliti l'accés del DMS a les guanines presents a la zona, o a una zona adjacent a aquesta. Això donarà com a resultat que les bandes corresponents a les citosines d'aquesta zona apareguin més intenses al gel de seqüència final, el que anomenem "hipersensibilitat". A la figura C.3 es representa esquemàticament un experiment de *footprinting in vivo* amb DMS. Amb aquests tipus d'experiments es pot obtenir una idea aproximada del nivell d'ocupació de la zona de DNA analitzada, a més de fer-se evidents canvis en aquesta ocupació en diferents condicions, com és en el nostre cas la sobreexpressió del factor de transcripció GAGA.

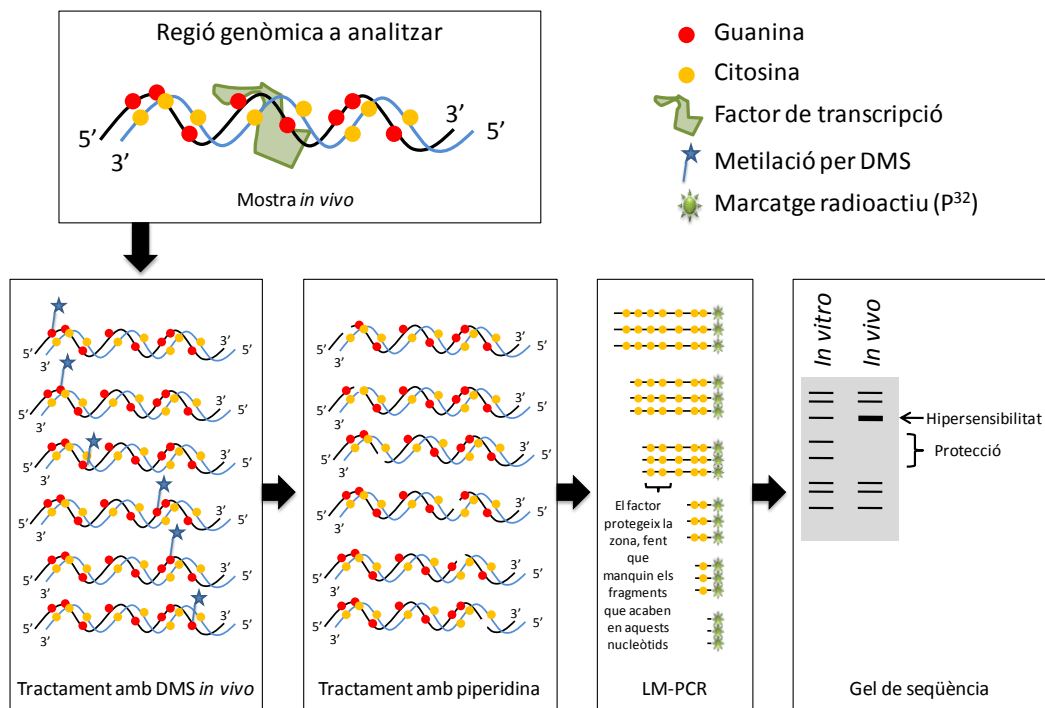


Figura C.3. Esquema d'un experiment de *footprinting in vivo*, on es mostra com un factor de transcripció protegeix una zona de la doble cadena de DNA, impedit la metilació amb DMS i el posterior trencament amb piperidina, fet que fa desaparèixer, al gel de seqüència, les bandes que correspondrien als fragments trencats per les bases protegides (protecció). També és possible que aquest factor de transcripció faciliti l'accés del DMS a altres bases, donant lloc a bandes més intenses al gel de seqüència, corresponents a trencaments en aquestes bases (hipersensibilitat).

C.2.4.4.A. Tractament amb DMS *in vivo* de cèl·lules S2

Partint de dos flascons de cultiu de 175 cm², amb un total de 30-60·10⁶ cèl·lules a cada flascó, es recullen les cèl·lules i s'ajunten en un tub falcon de 50 mL. Es centrifuga a 400 g, 5 minuts, es descarta el sobrenedant i es realitzen 2 rentats amb PBS 1X a temperatura ambient,

centrifugant a 400 g, 5 minuts. L'últim *pellet* cel·lular es resuspèn en 2 mL de PBS 1X a temperatura ambient i es col·loca en un bany d'aigua a 20°C fins que es temperi. D'aquesta manera s'estandaritza la temperatura de totes les reaccions. A una campana de gasos, ja que el DMS és molt tòxic i volàtil, s'afegeix 10 µL de DMS diluït al 10% en etanol als 2 mL de cèl·lules en PBS 1X i es mescla per inversió (0,05% final de DMS). A temps de 2 i 5 minuts s'agafa 1 mL i es col·loca a un eppendorf nou amb 100 µL de solució Stop "*in vivo*". Es mescla per inversió i es deixa el tub immediatament en gel. Un cop les dues reaccions, de 2 i 5 minuts, han estat parades, es centrifuguen a 400 g, 5 minuts, a 4°C. Es descarta el sobrenedant i es renta el *pellet* cel·lular amb 1 mL de PBS 1X (fred) centrifugant a 400 g, 5 minuts. Es descarta l'últim sobrenedant i es resuspenen les cèl·lules en 1 mL de tampó de Lisi. Es vortexen i s'incuben 3h a 50°C. Passat el temps, es refreden els tubs en gel, i es passa el contingut a un tub eppendorf de 2 mL. Es realitzen 3 extraccions amb fenol/cloroform (1:1), afegint un volum a cada mostra, vortexant vigorosament, centrifugant a 14000 rpm 5 minuts a una centrífuga eppendorf i finalment, recuperant la fase aquosa. Precipitar la última fase aquosa afegint 1 volum d'acetat amònic 7,5 M i 5 volums d'etanol al 100% (fred), es vortexa i es centrifuga 30 minuts a 20000 g a 4°C, a una centrífuga Beckman. Es renta el *pellet* amb 5 mL d'etanol al 70% (fred), centrifugant a 20000 g 10 minuts i el *pellet* es deixa assecar. Resuspendre en 300 µL d'H₂O milliQ cada *pellet* i tornar a precipitar-los afegint 30 µL de NaCl 4M, 10 µL d'acrilamida 0,25% i 1 mL d'etanol al 100% (fred). Es vortexen i es centrifuguen 20 minuts a 14000 rpm i 4°C en una centrífuga eppendorf. Els pellets es renten amb 1 mL d'etanol 70% (fred), vortexant-los i centrifugant-los a 14000 rpm i 4°C. Finalment, es deixen assecar els pellets i es resuspenen en 90 µL d'H₂O milliQ, deixant-los així preparats pel tractament amb piperidina.

Stop "*in vivo*"

1 M β-mercaptoetanol
10% BSA (albúmina sèrica bovina) *Faty acid free*
Diluir en PBS 1X

Tampó de Lisi

100 mM NaCl
10 mM Tris-HCl pH 8
0,2 mg/mL Proteïnasa K
25 mM EDTA pH 8
0,5% SDS

C.2.4.4.B. Obtenció i tractament *in vivo* amb DMS dels embrions transgènics de *Drosophila melanogaster*

Amb la realització dels *footprintings* genòmics *in vivo* amb DMS a embrions transgènics de *Drosophila melanogaster* preteníem caracteritzar el grau d'ocupació del promotor de *Trl* endogen d'embrions transgènics que sobreexpressen GAGA₅₁₉, comparant-lo a més amb el grau

MATERIALS I MÈTODES

d'ocupació d'aquest promotor que s'obté a embrions transgènics control. Per aconseguir els individus necessaris per aquest experiment, vàrem realitzar els creuaments esquematitzats a la figura C.4 (tots a 18°C).

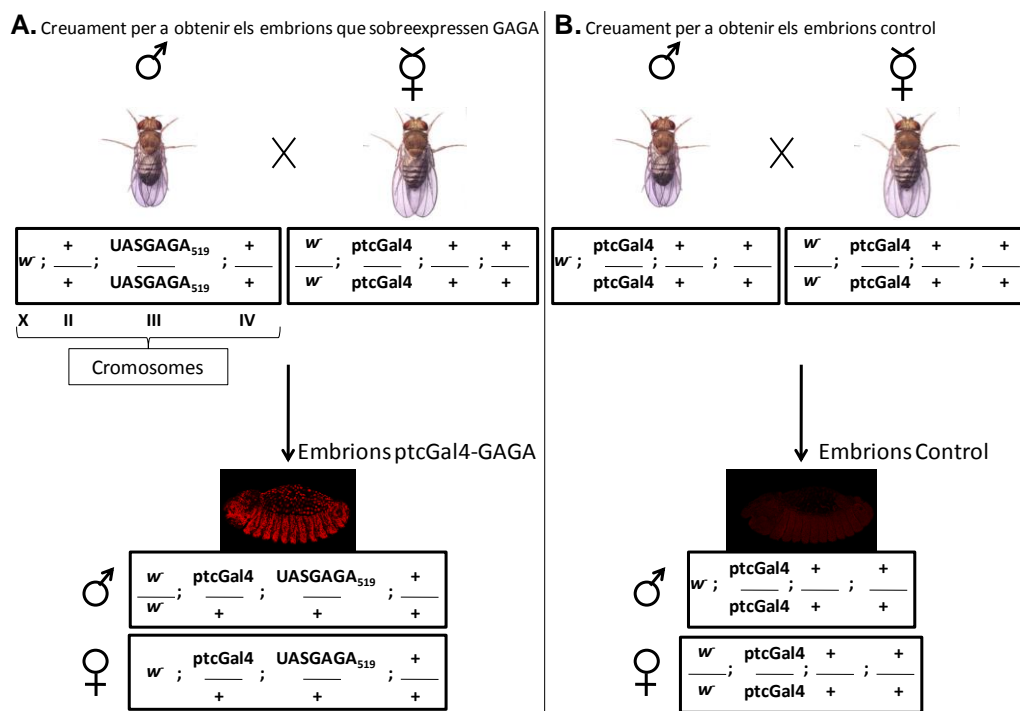


Figura C.4. Esquema dels creuaments fets per a obtenir (A) els embrions ptcGal4-GAGA (GAGA) i (B) control, utilitzats als experiments de *footprinting* genòmic *in vivo* amb DMS. Tots els creuaments es van fer mantenint els individus a 18°C.

♂: mascle; ♀: femella; ♀: femella verge. Amb (+) s'indica un cromosoma salvatge, sense cap transgèn. Les imatges (obtingudes amb microscopi confocal Leica SPE) dels embrions corresponen a immunolocalitzacions amb anticòs anti-GAGA.

Es recullen els embrions de 0-24 h, arribant a un volum d'uns 100 µL d'embrions. Immediatament, es decorionen tractant-los amb hipoclorit de sodi aproximadament al 5% durant 2 minuts i rentant-los amb PBS 1X. Tot seguit, es passen a un tub eppendorf, es resuspenen en 1 mL de PBS1X i es col·loquen les mostres en un bany d'aigua a 20°C per a que es temperin. S'afegeixen 5 µL de DMS 10%, diluït en etanol, arribant així a una concentració de 0,05% de DMS final. Mesclar per inversió i incubar 2 minuts a 20°C. Afegir 100 µL de solució Stop "*in vivo*", mesclar per inversió i deixar en gel. Ràpidament es centrifuguen les mostres a 400 g 5 minuts a 4°C. Descartar el sobrenedant i fer tres rentats amb 1 mL de PBS 1X + Tritó X100 0,1%, centrifugant 5 minuts a 4°C i 500 g. En acabat, es resuspenen els embrions amb 0,5 mL de tampó de Lisi i es disgreguen les seves cèl·lules amb ajuda d'un petit èmbol mida eppendorf (Pellet Pestle®), es vortexen i es deixen 3 h a 50°C. Passat aquest temps, es realitzen 3 extraccions amb fenol/cloroform, afegint un volum a cada mostra, vortexant vigorosament,

centrifugant a 14000 rpm 5 minuts en una centrífuga eppendorf i finalment, recuperant la fase aquosa en un eppendorf de 2 mL. S'afegeix NaCl per a que quedi a una concentració final de 0,4 M i es precipita el DNA de la mostra afegint 3 µL de acrilamida 0,25% i 1,5 mL d'etanol 100% (fred), es vortexa i es centrifuga a 14000 rpm 20 minuts a 4°C. Tot seguit, es renta el *pellet* amb 2 mL d'etanol 70% (fred), vortexant la mostra i centrifugant-la 10 minuts a 14000 rpm. Finalment, es resuspèn la mostra en 90 µL d'H₂O milliQ, deixant-la així preparada pel tractament amb piperidina.

C.2.4.4.C. Tractament *in vitro* amb DMS

Es va realitzar a partir d'extractes genòmics de cèl·lules S2 (obtinguts tal com es descriu a l'apartat C.2.4.2) com a control, ja que s'obté el patró que correspon el DNA sense cap proteïna associada. Donat que el DMS metila guanines, i degut al disseny d'encebadors emprats, el patró que s'obté d'aquestes mostres correspon a les citosines de la seqüència, constituint una escala (*ladder*) de citosines.

Es dilueixen 100 µg de DNA genòmic en un volum final de 198 µL. S'afegeixen 2 µL de DMS 5% (diluint en etanol), quedant així a 0,05% final, i es mescla per inversió. S'incuba durant 5 minuts a 20°C (en bany d'aigua) i tot seguit s'afegeixen 50 µL de solució Stop "*in vitro*". Es mescla per inversió i es deixa en gel. Després s'afegeixen 750 µL d'etanol 100% (fred), es vortexa i es centrifuga a 14000 rpm (a una centrífuga eppendorf), 20 minuts, a 4°C. Es descarta el sobrenedant i es deixa assecar el *pellet* a la campana durant 5 minuts. Seguidament, es resuspèn en 250 µL d'acetat sòdic 0,3M pH 5,0. Es tornen a afegir 750 µL d'etanol 100% (fred), es vortexa i es centrifuga 20 minuts a 4°C i 14000 rpm. Es descarta el sobrenedant i es renta el *pellet* amb 1 mL d'etanol 70% (fred), centrifugant a 14000 rpm, durant 10 minuts. Tot seguit, es descarta el sobrenedant i es deixa assecar el *pellet*. Un cop sec, es resuspèn en 90 µL d'H₂O deixant-lo llest pel tractament amb piperidina.

Stop "*in vitro*"

1,5 M Acetat sòdic pH 5,0

1 M β-mercaptoetanol

C.2.4.4.D. Tractament amb àcid fòrmic *in vitro* (*ladder* G+A)

Es va realitzar a partir d'extractes de DNA genòmic d'embrions control (obtinguts tal com es descriu a l'apartat C.2.4.3). L'àcid fòrmic modifica tant les adenines com les guanines,

MATERIALS I MÈTODES

de tal manera que un tractament amb piperidina posterior produeix trencaments a la cadena de DNA en aquestes adenines i guanines modificades. Així, de manera anàloga al tractament amb DMS *in vitro*, el tractament amb àcid fòrmic *in vitro* mostrarà les citosines i timines complementaries a les guanines i adenines modificades, obtenint així una escala o *ladder* G+A.

Es dilueixen 100 µg de DNA genòmic en un volum final de 32 µL, en H₂O. S'afegeixen 8 µL d'àcid fòrmic al 40% (diluït en aigua), quedant a un 8% final. Mesclar per inversió i incubar 20 minuts a 37°C. Tot seguit, congelar amb l'ajuda de neu carbònica i assecar en un concentrador SpeedVac (Eppendorf) durant 30 minuts a temperatura ambient. Resuspendre en 30 µL d'H₂O milliQ i tornar a congelar en neu carbònica. Assecar una altra vegada en l'SpeedVac 30 minuts a temperatura ambient i finalment, resuspendre en 90 µL d'H₂O milliQ, deixant la mostra preparada pel tractament amb piperidina.

C.2.4.4.E. Tractament amb piperidina

Produeix trencaments de cadena simple a les bases modificades pels tractaments anteriors.

Als 90 µL provinents dels tractaments anteriors, s'afegeixen 10 µL de piperidina (10% final). Es vortexa la mostra i s'incuba 30 minuts a 90°C. Tot seguit es refreda en gel i s'afegeixen 1,2 mL d'n-butanol. Es vortexa i es centrifuga a 14000 rpm (a una centrífuga eppendorf) durant 20 minuts, a 4°C. Es descarta el sobrenedant i es renta el *pellet* amb 1 mL d'etanol 95% (fred) vortexant la mostra i centrifugant-la a 14000 rpm durant 20 minuts i a 4°C. Es descarta el sobrenedant, s'asseca el *pellet* i es resuspèn en 100 µL d' SDS 1%. Es torna a afegir 1,2 mL d'n-butanol, es torna a vortexar i es centrifuga a 14000 rpm durant 20 minuts, a 4°C. Es descarta el sobrenedant i es renta dues vegades amb 1 mL d'etanol 70% (fred), centrifugant a 14000 rpm, 10 minuts. Per últim, s'asseca bé el *pellet* i es resuspèn en 20 µL de TE 1X.

C.2.4.4.F. *Linker-mediated* PCR (LM-PCR)

Amb els tractaments anteriors s'aconsegueixen una sèrie de cadenes de DNA amb trencaments monocadena, allà on les bases han sigut modificades. Per amplificar de forma específica la regió que es vol analitzar, així com per procedir al marcatge radioactiu, es realitza una PCR duta a terme amb l'ajuda d'un DNA *linker*, d'aquí el nom de *linker-mediated* PCR (LM-PCR). La LM-PCR (esquematzada a la figura C.5) consta d'una fase inicial d'extensió amb un

primer encebador. Aquesta reacció és una *primer extension*, ja que no s'amplifica mitjançant PCR sinó que simplement es copia una de les cadenes de DNA. Seguidament, es lliga el DNA *linker*. Aquest *linker* és un petit oligonucleòtid de doble cadena de DNA que s'unirà només als extrems roms, que seran els que s'hagin produït pels tractaments anteriors i la *primer extension* amb el primer encebador. Després es procedirà a amplificar específicament les seqüències d'interès amb un parell d'encebadors: un correspondrà al DNA *linker* i l'altre serà una mica més intern a la seqüència d'interès que el primer encebador (és l'encebador 2). Així s'amplifica una col·lecció de fragments de la regió genòmica d'interès, que finalitzaran als llocs on van ser modificades les bases. Finalment, emprant un últim encebador (encebador 3), una mica més intern a la seqüència d'interès que l'anterior encebador, marcat radioactivament al seu extrem 5' amb P^{32} , i mitjançant una *primer extension* de diversos cicles, s'amplifica i marca radioactivament aquesta col·lecció de fragments. Tot seguit, es resolen mitjançant un gel de seqüència al qual s'exposa una pel·lícula d'autoradiografia Hiperfilm (Amersham-Pharmacia-Biotech). Els encebadors emprats cada cop són més interns a la seqüència per augmentar l'especificitat de l'experiment.

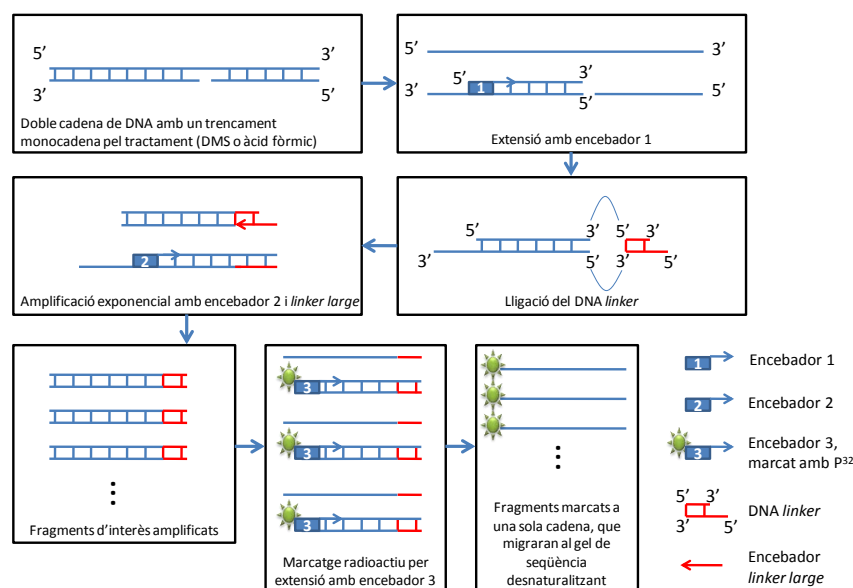


Figura C.5. Esquematzació de la LM-PCR. A la figura només es representa una molècula de DNA amb un únic trencament, realment es parteix de milions de molècules amb diversos trencaments en diferents posicions (idealment un trencament per molècula, a la regió d'interès). Primer, per *primer extension* amb l'encebador 1 es copia la cadena complementària a la que té el trencament, a la regió d'interès, detenint-se la còpia al lloc del trencament, generant així un extrem rom. A aquests extrems roms és on es lliga el DNA *linker* i posteriorment, emprant com a encebador la cadena llarga del DNA *linker* i l'encebador 2, s'amplifica per PCR la regió entre l'encebador 2 i el lloc de trencament. Seguidament, per *primer extension* amb l'encebador 3, marcat radioactivament, es marca la cadena de DNA complementària a la que originàriament portava el trencament. Aquestes seran les que després es veuran al gel de seqüència.

MATERIALS I MÈTODES

➤ Extensió amb l'encebador 1

Partint de 5 µL de cada una de les mostres tractades amb piperidina, s'afegeixen 1 µL d'encebador 1 (0,5 pmol/µL), 0,75 µL de 100 mM MgSO₄, 2,5 µL de tampó 10X de la polimerasa Vent exo- (Roche) i 15,75 µL d'H₂O milliQ. En un termociclador (Applied Biosystems 2400) es deixen 5 minuts a 95°C i 30 minuts a la temperatura d'*annealing* de l'encebador 1. Tot seguit, al mateix termociclador, afegir als tubs 0,5 µL de tampó 10X de la polimerasa Vent exo-, 0,6 µL de dNTPs 10 mM, 0,15 µL de 100 mM MgSO₄, 0,5 µL de polimerasa Vent exo- (Roche) i 3,25 µL d'H₂O milliQ, i deixar-los 10 minuts més a la temperatura d'*annealing* de l'encebador 1 i 15 minuts a 72°C.

➤ Lligació del DNA linker

Als 30 µL resultants de la reacció anterior, afegir: 7 µL de tampó 10X de lligasa (Fermentas), 5 µL de la solució DNA linker, 0,75 µL d'ATP 10 mM, 0,6 µL de T4 DNA lligasa (Fermentas) i 26,65 µL d'H₂O milliQ. S'incuba tota la nit a 18°C.

Solució DNA linker

20 µL de linker large (100 pmol/µL)

20 µL de linker small (100 pmol/µL)

10 µL de tampó 10X M (Boheringer)

10 µL de MgCl₂ 50 mM

40 µL H₂O milliQ

Escalfar 5 minuts a 95°C en un bany d'aigua i deixar refredar tota la nit

➤ Amplificació exponencial

El resultat de la lligació del DNA linker es precipita amb etanol i es resuspèn en 85 µL de la mix A. Es fa un *Hot Start* de 5 minuts a 95°C al termociclador. Després, al termociclador s'afegeixen els 15 µL de la mix B i es realitza el següent programa de PCR:

30 cicles { 40 segons a 95°C
2 minuts a la temperatura d'*annealing* de l'encebador 2
2 minuts a 72°C
Extensió final 7 minuts a 72°C

Les reaccions es poden guardar a -80°C.

Mix A (per mostra)

8,5 µL de tampó 10X de la DNA Polimerasa (Biotools)

3,4 µL de MgCl₂ 50mM

1 µL de BSA 10 mg/mL

2 µL encebador 2 (10 pmol/µL)

2 µL linker long (20 pmol/µL)

68,1 µL H₂O milliQ

Mix B (per mostra)

1,5 µL de tampó 10X de la DNA Polimerasa (Biotools)

0,6 µL de MgCl₂ 50mM

4 µL de dNTPs (10 mM cadascun)

1 µL DNA polimerasa (Biotools)

7,9 µL H₂O milliQ

➤ Extensió amb encebador 3 marcat radioactivament

S'agafen 15 μL de la reacció d'amplificació exponencial i se li afegeixen 5 μL de la mix C. Després, es realitza el següent programa al termociclador:

9 cicles { 2 minuts a 94°C
40 segons a 94°C
3 minuts a la temperatura d'*annealing* de l'encebador 3
5 minuts a 72°C

Mix C (per mostra)

0,5 μL de tampó 10X de la DNA polimerasa (Biotools)
0,75 μL de dNTPs (10 mM cadascun)
0,25 μL de MgCl_2 50mM
0,5 μL de solució d'encebador 3 marcat
0,5 μL de DNA polimerasa (Biotools)
2,75 μL d' H_2O milliQ

Solució encebador 3 marcat

2 μL d'encebador 3 (10 pmol/ μL)
1 μL tampó 10X de la PNK (Roche)
2 μL de γATP^{32} (10 $\mu\text{Ci}/\mu\text{L}$)
0,75 μL PNK (Polinucleòtid quinasa, Roche)
4,25 μL d' H_2O milliQ
S'incuba 45 minuts a 37°C i es guarda a -20°C

En acabat, es fa una extracció amb fenol/cloroform de les mostres, es precipita el DNA amb etanol absolut i es renta amb etanol 70%. Es resuspèn la mostra en 15 μL de tampó de càrrega i s'escalfa 5 minuts a 95°C. Es carreguen 5-10 μL en un gel de seqüència al 6% d'acrilamida.

Tampó de càrrega (2 mL)

2 mL Formamida desionitzada
10 μL SDS 10%
Xilencianol (punta d'espàtula)
Blau de bromofenol (punta d'espàtula)

Gel de seqüència 6% d'acrilamida (100 mL)

48 g d' Urea
10 mL TBE 10X
15 mL Acrilamida 40% (19:1, acrilamida:bisacrilamida)
 H_2O fins a 100 mL

Filtrar i desgasar

El gel de seqüència es va muntar en vidres de 30 cm X 40 cm, siliconitzant un dels vidres i segellant-los amb cinta d'embalar. Per polimeritzar el gel s'afegeixen 500 μL de PSA i 25 μL de TEMED, ràpidament s'aboca a l'interior dels vidres evitant fer bombolles. L'electroforesi es fa a 40 W durant 2-3h. Finalment, s'asseca el gel i s'exposa a una pel·lícula d'autoradiografia Hiperfilm (Amersham-Pharmacia-Biotech) a temps variables, segons la intensitat del marcatge radioactiu.

En els *footprintings* realitzats en aquest treball s'han utilitzat els encebadors de la taula C.6, amb la temperatura d'*annealing* indicada.

MATERIALS I MÈTODES

Joc d'encebadors	Encebador	Temperatura d' <i>annealing</i>
FI	FI-1	50°C
	FI-2	56°C
	FI-3	62°C
FIIA	FIIA-1	56°C
	FIIA-2	62°C
	FIIA-3	68°C
FIIB	FIIB-1	54°C
	FIIB-2	63°C
	FIIB-3	70°C
FIIC	FIIC-1	58°C
	FIIC-2	63°C
	FIIC-3	70°C
FIII	FIII-1	50°C
	FIII-2	55°C
	FIII-3	62°C

Taula C.6. Relació d'encebadors emprats als *footprintings*, i la seva temperatura d'*annealing*.

C.2.5. Tècniques d'anàlisi d'RNA

C.2.5.1. Extracció d'RNA total

Les extraccions d'RNA total de discos d'ala de larves MS1096Gal4; +/+; +/+ i de larves MS1096Gal4; +/+; UASGAGA₅₁₉/+ han estat realitzades per Marta Blanch Lozano, utilitzant el Mini RNA Isolation I Kit™ (Zymo Research) segons les instruccions del fabricant, emprant 20 discos de cada tipus de larva per extracció.

Les extraccions d'RNA total de cèl·lules S2, així com de les diferents línies establiment transfectades provinents d'aquestes, s'han fet amb el kit RNeasy Mini Kit de Qiagen®, segons les instruccions del fabricant, emprant 5-10·10⁷ cèl·lules per extracció.

Per a conèixer la concentració de RNA obtinguda s'han fet tant mesures a un espectrofotòmetre UV-1603 (Shimadzu) (diluïnt 1/100 la mostra, la [RNA] mg/mL = Absorbància a 260 nm x 4), com utilitzant un BioAnalyzer 2100 (Agilent Technologies). Les mesures amb el BioAnalyzer s'han realitzat al servei de transcriptòmica del Parc Científic de Barcelona i a més de servir per a calcular la concentració d'RNA també serveixen per comprovar la qualitat de l'RNA extret.

C.2.5.2. RT-PCR

Tots els assajos d'RT-PCR que es mostren en aquest treball han estat duts a terme mitjançant el kit OneStep RT-PCR de Qiagen®.

La tècnica d'RT-PCR permet realitzar amplifacions d'RNAs específics. Per a que això pugui tenir lloc, cal que l'RNA sigui retrotranscrit a cDNA, que després serà amplificat per PCR convencional. Aquesta retrotranscripció és duta a terme per una DNA polimerasa/RNA-dependent: la retrotranscriptasa o transcriptasa reversa, que utilitza la cadena d'RNA com a motlle per a generar una cadena complementària de DNA.

Per a que la transcriptasa reversa pugui afegir dNTPs, cal que hi hagi uns encebadors units a la cadena d'RNA que es vol amplificar. Per a les RTs, s'utilitzen diferents tipus d'encebadors: l'oligo-dT (format només per residus timina i que s'uneix a tots els RNAs que presenten cua poliA), o encebadors específics, que s'uneixen a l'RNA que es vol amplificar. En aquest treball hem utilitzat encebadors específics, que amplifiquen dos transcrits diferents: U6, un gen transcrit per l'RNA Polimerasa III que codifica per un snRNA, utilitzat com a control, ja que no hauria de veure alterada la seva quantitat d'mRNA; i *Trl*, el gen d'interès.

D'altra banda, l'enzim que es subministra amb el kit que hem utilitzat presenta una funció dual: té tant activitat retrotranscriptasa com activitat DNA polimerasa/DNA-dependent (DNA polimerasa convencional). Aquestes dues activitats es fan servir de manera seqüencial, és a dir, primer s'utilitza la seva activitat retrotranscriptasa, i seguidament s'inactiva aquesta i s'activa la seva activitat DNA polimerasa convencional. La temperatura òptima de retrotranscripció és de 50°C, i és aconsellable mantenir les mostres en gel fins el moment de començar la RT, per a evitar *annealings* i amplifacions inespecífiques. La retrotranscriptasa s'inactiva durant 15 minuts a 95°C, moment en que s'activa la DNA polimerasa convencional.

Finalment, per a cada experiment s'ha realitzat un control de contaminació de DNA genòmic, on els encebadors són afegits després de la reacció d'RT i, per tant, no es produeix el cDNA. A la figura C.6 es mostren les condicions en les que s'han realitzat els experiments.

MATERIALS I MÈTODES

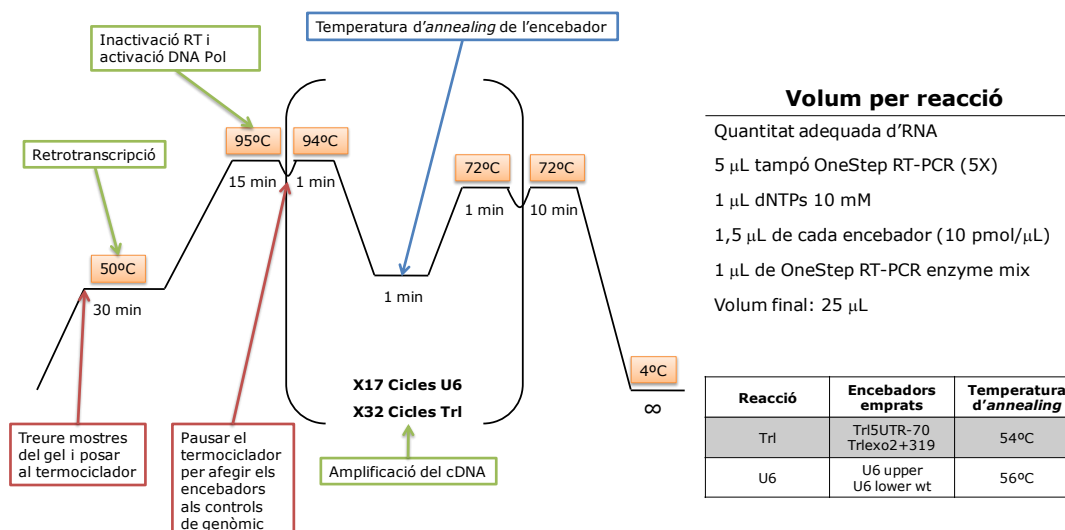


Figura C.6. A l'esquerra, temperatures i cicles de les RT-PCRs realitzades en aquest treball. A la dreta, reactius emprats en les reaccions i temperatures d'annealing específiques de cada reacció.

C.2.5.3. Real Time RT-PCR

La tècnica de la PCR a temps real (*Real Time*) es basa en l'habilitat per monitoritzar el progrés de la PCR mentre la reacció s'està produint, ja que l'amplificació i la detecció es realitzen a la vegada. Per a la detecció del DNA s'utilitza la fluorescència, en concret la produïda per SYBRGreen, una molècula petita i amb capacitat d'unir-se al DNA de doble cadena sense inhibir la reacció de PCR, i que al unir-se el DNA de doble cadena emet fluorescència. D'aquesta manera, la fluorescència que es detecta és directament proporcional al nombre d'amplicons generats a la reacció de PCR i, com més gran sigui el nombre de molècules de partida, més aviat s'observarà un increment de la fluorescència. Per tant, la fluorescència detectada permet calcular la concentració de DNA de la mostra. El valor Ct (*Cycle threshold*) és el cicle de la PCR en el qual la fluorescència detectada és més gran que la fluorescència *background* (soroll de fons).

El primer pas és la retrotranscripció, per produir el cDNA a partir de l'RNA. Per aquest pas, en aquest treball hem utilitzat el kit Omniscript RT de Qiagen®. Aquest kit conté dues transcriptases reverses amb alta afinitat per l'RNA, permetent la reacció fins i tot quan l'mRNA es troba en baixa còpia. Com a encebadors hem emprat oligonucleòtids específics per a dos mRNAs: "U6 lower wt", que amplifica part de l'mRNA d'U6, un gen transcrit per l'RNA Polimerasa III que codifica per un snRNA, utilitzat com a control, ja que no hauria de veure's alterada la seva quantitat d'mRNA; i "Trlexo2+319", que amplifica la regió 5'UTR de *Trl*,

permetent per tant distingir entre l'mRNA de la còpia endògena de *Trl* i l'mRNA de còpies ectòpiques emprades per sobreexpressar GAGA.

El segon pas és l'amplificació i quantificació del cDNA, utilitzant Platinum SYBR Green qPCR Super Mix UDG (Invitrogen) i 300 nM de cada encebador *forward* (fwd) i *reverse* (rev), específics del gen a analitzar. Per a amplificar el cDNA d'U6 s'han emprat els encebadors "U6 q fwd" i "U6 q rev". Per a amplificar el cDNA de *Trl* s'han emprat els encebadors "Trl q fwd" i "Trl q rev". Aquest pas s'ha dut a terme a un aparell ABI Prism 7700 Sequence Detector (Applied Biosystems). Les condicions de PCR han estat: 2 minuts d'incubació amb UDG a 50°C, 2 minuts d'activació de la polimerasa i 40 cicles de 15 segons a 95°C i 30 segons a 60°C. No és necessari un pas d'extensió a 72°C perquè les regions a amplificar són molt curtes (entorn a 50 pb). Cada assaig cal que inclogui corbes patró amb diverses dilucions seriadades per a cada punt (per triplicat) i controls sense DNA (NTC o *non-template control*). També s'analitzen mostres amb les quals no s'ha realitzat el pas previ de retrotranscripció, per a descartar contaminació de DNA. Els resultats han estat analitzats amb el *software* ABI Prism 7000 SDS (Applied Biosystems) i amb Microsoft Excel.

C.2.5.4. Tractament de les cèl·lules amb RNA d'interferència

Per a l'obtenció de cèl·lules S2 que no expressin les proteïnes dCtBP o Psq, hem incubat les cèl·lules amb dsRNA de seqüència homòloga al gen a silenciar. Aquest dsRNA és capaç de difondre a l'interior de les cèl·lules en cultiu, activant llavors la maquinària de l'RNA d'interferència, provocant així l'eliminació activa dels transcrits homòlegs al dsRNA i per tant, silenciant l'expressió del gen diana. Com a control hem utilitzat un dsRNA de GFP.

C.2.5.4.A. Preparació d'RNA de doble cadena (dsRNA)

Per a l'obtenció dels dsRNAs de dCtBP, Psq i GFP hem partir de les construccions pBS-RNAi-dCtBP, pBS-RNAi-Psq i pBS-RNAi-GFP, que contenen part (en el cas de GFP, la totalitat) de la regió codificant d'aquests gens, entre els promotors de les RNA polimerases T3 i T7.

Es digereix 1 µg de cada plasmidi, per tal d'aconseguir fragments lineals que continguin el promotor de T7 o de T3 a una banda, i extrems 5' protuberants a l'altra. A la figura C.7 es mostra un esquema dels plasmidis utilitzats, indicant els promotors de T7 i T3 i les seqüències diana dels enzims de restricció emprats per a linealitzar els plasmidis.

MATERIALS I MÈTODES

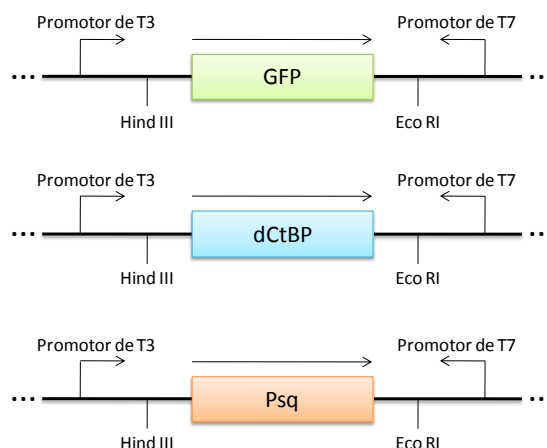


Figura C.7. Esquema dels plasmidis pBS-RNAi-dCtBP, pBS-RNAi-Psq i pBS-RNAi-GFP.

En tots els casos aquestes digestions han estat amb Eco RI per a obtenir el fragment lineal que es transcriurà amb la T3 polimerasa, i amb Hind III per a obtenir el fragment lineal que es transcriurà amb la T7 polimerasa.

Un cop digerits aquests fragments, s'han realitzat transcripcions *in vitro*, incubant la següent preparació durant 2 hores a 37°C:

- 2 µL de 0,5 µg/µL de DNA digerit
- 10 µL de tampó de transcripció 5X (Ambion)
- 10 µL de rNTP (10 mM cadascun) (Amersham)
- 1,5 µL de RNA polimerasa T3 o T7 (Ambion)
- 1 µL anti-RNAsa (Ambion)

Passat aquest temps, s'ajunten les reaccions de transcripció fetes amb les RNA polimerases T3 i T7 i, en un bany d'aigua, s'incuben 5 minuts a 95°C (per desnaturalitzar els transcrits) i es deixen refredar lentament fins a temperatura ambient, dins el mateix bany. En acabat, es fa una extracció amb fenol/cloroform, es precipita amb etanol 100% i es renta amb etanol 70%. Per eliminar el DNA i les restes d'RNA que no siguin dsRNA, es resuspèn la mostra en 50 µl de tampó 1X M (Boehringer) i s'afegeix 1 µl de DNAsal (10 units/µl) i 0.5 µl d'RNasaA (10 µg/µl) (ambdues de Roche). Incubar 15 minuts a 37°C. Tornar a fer fenol/cloroform, precipitar amb etanol 100% i rentar amb etanol 70%. Resuspendre amb uns 30-50 µl d'H₂O milliQ. Es comprova la integritat i concentració del dsRNA carregant una alíquota en un gel d'agarosa a l'1%.

C.2.5.4.B. Tractament de les cèl·lules amb dsRNA

Pel tractament de les cèl·lules amb dsRNA, hem sembrat $3 \cdot 10^6$ cèl·lules S2 en plaques de cultiu de 60 mm^2 , amb 5 mL de medi Schneider suplementat. Les hem transfectat, amb els plasmidis concrets per a cada cas, tal com s'explica a l'apartat C.2.2.5. 24 h després, recollim les cèl·lules en tubs falcon, les centrifuguem a 400 g, 10 minuts. Descartem el sobrenedant i fem un rentat afegint 5 mL de PBS 1X i repetint la centrifugació. El *pellet* cel·lular el resuspenem en 5 mL de medi Schneider suplementat, afegim la quantitat indicada de dsRNA (RNAi) a cada cas, tornem a col·locar les cèl·lules en plaques de cultiu de 60 mm^2 i s'incuben 72 h (pel cas de l'RNAi de dCtBP) o 96 h (pel cas de l'RNAi de Psq), a 25°C. Passat el temps, es recullen les cèl·lules i es realitza un assaig luciferasa i β -gal convencional, tal com s'explica als apartats C.2.2.6-8.

C.2.5.5. Assajos de *microarray*

Per a la realització dels assajos de *microarray* presentats en aquest treball, cada mostra es va preparar extraient l'RNA total de les cèl·lules adients, tal qual s'explica a l'apartat C.2.5.1. Les mostres van ser enviades als "Affymetrix GeneChip® Array-Services" de l'European Molecular Biology Laboratory (EMBL) situats a Heidelberg (Alemanya), on es van dur a terme *arrays* del tipus "Drosophila Genome 2.0 Array", d'Affymetrix.

La informació sobre l'expressió gènica que es va obtenir es va normalitzar via RMA (Irizarry et al. 2003), tal com està implementat a la funció *justRMA* del *software* Bioconductor (Gautier et al. 2004; Gentleman et al. 2004).

Per a la comparació dels resultats de mostres control i tractades (Control vs. GAGA i Control vs. RNAiGAGA) es van realitzar tests estadístics t-test moderats (Smyth 2004), tal com estan implementats a la llibreria *limma* de Bioconductor. Seguidament, es va calcular la probabilitat de que cada gen estigués diferencialment expressat (comparant controls i tractades) per ajust a una densitat t parcial semi-paramètrica (Rossell et al. 2008), i es va obtenir un llistat de gens diferencialment expressats controlant el "Bayesian FDR" (*False Discovery Rate*) per sota del 5% (Müller et al. 2004). Finalment, l'anàlisi dels processos biològics sobrerrepresentats al llistat de gens diferencialment expressats entre control i tractades (*Gene Ontology*) es va dur a terme utilitzant el *software* DAVID (*Database for Annotation,*

MATERIALS I MÈTODES

Visualization, and Integrated Discovery), disponible *on line* (<http://david.abcc.ncifcrf.gov/home.jsp>).

Per a la comparació de les dades dels nostres *microarrays* amb les obtingudes als experiments de *ChIP on chip* de GAGA per Lee i col·laboradors (Lee et al. 2008), primer es van analitzar les dades públicament disponibles, que van ser generades per aquests autors a partir de “GeneChip® Drosophila Tiling 1.0R Arrays”, d’Affymetrix. Aquest anàlisi es va realitzar amb el *software* MAT (Johnson et al. 2006). Vàrem calcular el “MAT score” corrent finestres de 300 pb, seleccionant pics amb un “*p-value*” d’enriquiment $< 0,00001$ i ajuntant els pics a menys de 300 pb. El “*p-value*” llindar correspon a un 1% FDR (*False Discovery Rate*). Finalment, vàrem avaluar en quina mesura els pics enriquits solapaven amb els gens diferencialment expressats als nostres *microarrays*. Això es va dur a terme mapant quin era el gen més proper a cada pic de *ChIP on chip*, utilitzant la funció *findClosestGene* de Bioconductor (Scacheri et al. 2006). Aquests són els gens que vàrem assignar com a “positius” al *ChIP on chip*, ja que GAGA pot estar controlant d’alguna manera la seva activitat.

L’anàlisi estadística de les dades obtingudes dels experiments de *microarray*, l’anàlisi estadística de les dades disponibles dels experiments de *ChIP on chip*, i la comparació d’ambdós tipus de resultats, han estat realitzades pel David Rossell, de la unitat de Bioinformàtica i Bioestadística de l’IRB (*Institute for Research in Biomedicine*), ubicat al PCB (Parc Científic de Barcelona).

C.2.6. Tècniques d’anàlisi de proteïnes

C.2.6.1. Electroforesi en gel de poliacrilamida-SDS (SDS-PAGE)

L’electroforesi en gels de poliacrilamida-SDS s’utilitza per a separar proteïnes segons el seu pes molecular. L’SDS desnatura les proteïnes i cobreix la seva superfície de càrregues negatives de manera que elimina els possibles efectes de la càrrega intrínseca de cadascuna. Així doncs, en un camp elèctric, les proteïnes migren en funció del seu pes molecular.

Els gels de poliacrilamida-SDS estan formats per dues fases: la fase superior o gel d’apilament i la fase inferior o gel separador. Cadascuna d’elles presenta un pH i un percentatge d’acrilamida diferent, havent sigut aquest últim d’un 4% pel gel d’apilament i d’un 10% pel separador. En aquest treball s’han utilitzat gels tipus *Mini-Protean* (8,5 x 8,5 cm).

Composició del gel apilador

acrilamida:bisacrilamida 29:1 4%
Tris-HCl 0,125 M pH 6,8
SDS 0,1%
6 µL de PSA (Persulfat amònic) 10%/mL solució
3 µL TEMED (N,N,N',N'-Tetramethylethylenediamine)/mL solució

Composició del gel separador

acrilamida:bisacrilamida 29:1 10%
Tris-HCl 0,4 M pH 8,7
SDS 0,1%
6 µL de PSA 10%/mL solució
3 µL TEMED/mL solució

A les mostres se'ls afegeix tampó de càrrega desnaturalitzant de proteïnes PLB 5X (*Protein Loading Buffer*) i β -mercaptoetanol (1,5 M final), s'escalfen 5 minuts a 95°C i immediatament es carreguen. L'electroforesi es fa a 25 mA, utilitzant com a tampó Laemmli 1X.

Un cop migrats, es descarta el gel d'apilament i el gel separador es tenyeix amb blau de Coomassie o es transfereix a una membrana de nitrocel·lulosa per a dur a terme un assaig de *western blot*. Per a la tinció amb blau de Coomassie, els gels s'incuben uns 20 minuts en la solució de tinció CBB i després, per eliminar l'excés de colorant, es destenyeixen amb una solució d'àcid acètic al 10%.

Laemmli 1X

0,1% SDS
200 mM Glicina
25 mM Tris base
pH 8,3-8,5

PLB 5X

125 mM Tris-HCl pH 6,8
21,75 % glicerol
5 % SDS
0,25 % blau de bromofenol

CBB

0,05% Coomassie blue brillant
40% etanol
10% àcid acètic

C.2.6.2. Western blot

Aquesta tècnica permet la detecció de proteïnes presents en un gel de poliacrilamida-SDS transferit a una membrana de nitrocel·lulosa mitjançant l'ús d'anticossos específics. Així doncs, el primer pas consisteix en una electroforesi SDS-PAGE i posteriorment es transfereix el gel a una membrana de nitrocel·lulosa Optitran (Schleicher & Schuell). Aquest pas es realitza en una cubeta amb tampó de transferència i aplicant un voltatge de 80 V durant 3 h a 4°C i amb agitació constant.

Un cop finalitzada la transferència, s'incuba la membrana amb agitació d'1 h a tota la nit amb la solució de bloqueig per tal d'evitar la unió inespecífica dels anticossos a la membrana. Posteriorment, es fa 1 rentat de 5 minuts amb PBST (PBS 1X, Tween 20 0,1%) i s'incuba amb la dilució (en PBST) d'anticòs primari adient en cada cas durant 1 h, a temperatura ambient i en agitació constant. Seguidament, es fan 3 rentats de 5 minuts amb PBST i s'incuben amb la dilució (en PBST) adequada d'anticòs secundari durant 1 h, a temperatura ambient i en agitació constant. Finalment, es fan 5 rentats de 5 minuts amb PBST i es procedeix a la detecció mitjançant quimioluminescència.

MATERIALS I MÈTODES

Per a la detecció de l'anticòs secundari conjugat a la peroxidasa de rave (HRP) es barregen 10 mL de solució de Luminol amb 10 µL de solució *Enhancer*, s'apliquen sobre la membrana i s'incuben 1 minut. Finalment, s'asseca la membrana i s'exposa amb una pel·lícula fotogràfica Agfa Curix RP2 el temps adient per obtenir un senyal òptim.

Tampó de transferència

Tris base 25 mM
Glicina 40 mM
SDS 0,05%
Metanol 20%

Solució de Bloqueig

PBS 1X
Llet desnatada en pols 5%
Tween 20 0,1%
Azida sòdica 0,01%

PBST

PBS 1X
Tween 20 0,1%

Luminol

Tris-HCl 0,1 M pH 8,6
Luminol sòdic 1,25 mM
H₂O₂ 2,7 mM

Enhancer

Àcid p-hidroxycumarínic 0,1% en DMSO

C.2.6.3. Expressió de proteïnes a bacteris

Per a aquest treball es va dur a terme l'expressió, a la soca BL21(DE3) d'*E.coli*, de la proteïna Groucho fusionada a una cua d'histidines. Aquestes cèl·lules es van transformar amb el plasmidi pET28a-His-Gro. Amb aquests bacteris s'inoculen 4 mL de medi LB i es fan créixer en agitació a 37°C fins que arribin a una densitat òptica (DO) entre 0,6 i 0,8. En aquest moment s'extreu una alíquota d'1 mL, col·locant-la a un tub eppendorf. S'indueix l'expressió amb IPTG (Isopropyl β-D-1-thiogalactopyranoside) 1 mM, 3 h a 37°C en agitació. Un cop acabada la inducció, s'agafa 1 mL del cultiu induït i es col·loca a un tub eppendorf. Juntament amb l'alíquota del cultiu sense induir, es centrifuguen les cèl·lules 6000 g, 5 minuts. Descartem el sobrenedant i fem dos rentats amb 1 mL de PBS 1X, centrifugant a 6000 g, 5 minuts. Seguidament, el *pellet* cel·lular es resuspèn amb 100 µL de tampó de càrrega de proteïnes (PLB 1X), s'afegeix β-mercaptoetanol (1,5 M final), es vortexen les mostres i s'escalfen a 95°C, 5 minuts. Finalment, es carreguen 20 µL de cada mostra a un gel SDS-PAGE i, acabada l'electroforesi, es tenyeix amb Coomassie.

C.2.6.4. Assaig d'interacció proteïna-proteïna *in vitro* per GST-

Pulldown

L'assaig de GST-*Pulldown* permet analitzar la interacció entre dues proteïnes en condicions *in vitro*. Es tracta d'expressar una de les proteïnes a analitzar fusionada a GST (Glutació S-transferasa) i unir-la a la reïna *Glutathione Sepharose 4B* de manera que s'obté una

columna d'afinitat. Aleshores s'incuba amb un extracte proteic i s'analitzen les proteïnes que queden retingudes. La detecció de les proteïnes que interaccionen específicament amb la columna es caracteritza per mitjà de *Western blot*.

Amb els nostres assajos de *GST-Pulldown* preteníem estudiar la possible interacció de GAGA amb el corepressor dCtBP. Per a tal fi, per una banda vam expressar a bacteris les proteïnes GAGA₅₁₉, TTK69 i el domini POZ de TTK69, fusionades a GST i les vam unir a reïnes *Glutathione Sepharose 4B*. D'altra banda, els extractes proteics de nuclis de cèl·lules S2 que vam incubar amb aquestes reïnes, els vam enriquir en proteïna dCtBP transfectant transitòriament el plasmidi d'expressió pActppa-dCtBP-L, a raó de 10 µg per placa de cultiu de 60 mm². D'aquesta manera s'augmenten les possibilitats de detectar la possible interacció GAGA-dCtBP.

C.2.6.4.A. Preparació d'extractes proteics nuclears de cèl·lules S2 transfectades transitòriament per a la realització de GST-Pulldowns

Primerament, transfectem transitòriament 10 plaques de cultiu de 60 mm² de cèl·lules S2 amb 10 µg de pActppa-dCtBP-L cada una, tal qual s'explica a l'apartat C.2.2.5. Transcorregudes 48 h de la transfecció, passem les cèl·lules a un tub falcon de 50 mL, ajuntant així el volum de totes les plaques. Baixem les cèl·lules centrifugant a 400 g durant 5 minuts i realitzem 3 rentats amb PBS 1X, centrifugant les cèl·lules 5 minuts a 400 g. Descartem l'últim sobrenedant i, de forma aproximada, determinem el volum que ocupa el *pellet* de cèl·lules (V). Es resuspèn aquest *pellet* en 3 vegades el volum V de tampó A. Amb l'ajuda d'un homogeneïtzador Dounce de 2 mL es lisen les cèl·lules utilitzant l'èmbol A, pujant-lo i baixant-lo unes 20 vegades. Després, es centrifuga el lisat a 3200 g 15 minuts. Es descarta el sobrenedant i el *pellet* es resuspèn altre cop en 3 vegades el volum V de tampó A. Es repeteix la centrifugació anterior i es descarta el sobrenedant. El *pellet* es resuspèn en 2 vegades el volum V de tampó B. En principi, la mostra actual conté la fracció nuclear de l'extracte cel·lular. Amb l'ajuda del Dounce i l'èmbol B, es lisen els nuclis pujant i baixant l'èmbol unes 20 vegades. Seguidament, es realitzen 3 cicles de congelació i descongelació de la mostra en neu carbònica, per acabar de trencar les membranes nuclears. La mostra s'homogeneïtza fent-la girar uns 30 minuts a 4°C en una roda de tubs eppendorf. Finalment, es centrifuga la mostra a 4°C 30 min, a 15000 rpm, en una centrífuga eppendorf. Recuperem el sobrenedant, que és l'extracte proteic nuclear.

MATERIALS I MÈTODES

Tampó A

15 mM Tris-HCl pH 7,4
60 mM KCl
15 mM NaCl
0,25 mM MgCl₂
230 mM sacarosa
0,5 mM espermina
0,15 mM espermidina
14 mM β-mercaptoetanol
0,2 mM PMSF
(phenylmethylsulphonyl
fluoride)

Tampó B

25 % glicerol
20 mM HEPES(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)-KOH pH 7,9
450 mM NaCl
1,5 mM MgCl₂
0,2 mM EDTA (àcid etilendiaminotetraacètic)
0,5 mM DTT (Ditiotreitol)
0,1 mM PMSF

C.2.6.4.B. Expressió i purificació de proteïnes fusionades a GST

La Glutatió S-transferasa de *Schistosoma japonicum* és una proteïna d'alta afinitat pel glutatió, fet que permet la purificació de proteïnes fusionades a ella mitjançant la reïna *Glutathione Sepharose 4B*. Per a l'obtenció de les proteïnes de fusió d'interès es van fer servir vectors d'expressió pGEX-KG (Amersham-Pharmacia-Biotech), els quals estan dissenyats per a la producció de grans quantitats de proteïna sota el control del promotor *tac*, induïble per IPTG.

El nostre objectiu era estudiar la interacció de GAGA amb el corepressor dCtBP. Per dur-ho a terme vàrem clonar GAGA₅₁₉ al vector pGEX-KG. Addicionalment, també vam clonar la proteïna TTK69 i el seu domini POZ en aquest vector. Estudis anteriors havien descrit que la interacció de TTK69 amb dCtBP podia donar-se a través del domini POZ de TTK69 (Wen et al. 2000). Així, l'expressió d'aquestes dues proteïnes de fusió ens aportava dos controls positius d'interacció amb dCtBP. En tots els casos les proteïnes de fusió es van expressar en la soca d'*E. coli* BL21(DE3).

Es transforma la soca BL21(DE3) amb el plasmidi corresponent, amb aquests bacteris s'inoculen 500 mL de medi LB i es fan créixer en agitació a 37°C fins que arribin a una densitat òptica (DO) entre 0,6 i 0,8. S'indueix l'expressió amb IPTG 1 mM, 3 h a 37°C, en agitació. La inducció de l'expressió de GST-GAGA₅₁₉ i GST-TTK69 es va fer amb IPTG 1mM durant 4h a 30°C en agitació per millorar el rendiment. Un cop acabada la inducció, es centrifuguen les cèl·lules a 6000 g 10 minuts a 4°C. Es descarta el sobrenedant i el *pellet* es resuspèn amb 10 mL de NETN + 1 mM PMSF + 10 µg/ Leupeptina. Seguidament es sonica la mostra utilitzant un sonicador Digital Sonifier® (Branson) amb el *microtip*, fent 3 cicles de 10 segons amb intervals de 10 segons de descans, i un valor de sortida entre 50 i 60 W. Després es centrifuga 15 minuts a 10000 rpm en una centrífuga Beckman. Es guarda el sobrenedant i el *pellet* es resuspèn en 5 mL de NETN + 2% Sarkosyl + 1 mM PMSF + 10 µg Leupeptina. Es repeteix la sonicació i la

centrifugació i finalment, el sobrenedant s'ajunta amb l'anterior. Aquests dos últims sobrenedants, que constitueixen l'extracte proteic bacterià, es guarden a -80°C.

NETN

20 mM Tris-HCl pH 8

100 mM NaCl

1 mM EDTA

0,5 % NP-40

C.2.6.4.C. Unió de les proteïnes fusionades a GST a la reïna *Glutathione*

Sepharose 4B

Es descongelen els extractes proteics dels bacteris BL21(DE3) induïts i es centrifuguen 10 minuts a 10000 rpm i 4°C, en una centrífuga Beckman. Al acabar recuperem el sobrenedant. Paral·lelament, preparem la reïna *Glutathione Sepharose 4B*. S'empaqueta 1 mL de la reïna en una *Econocolumn* (Biorad) i s'equilibra amb 10 mL de tampó NETN. Tota el procediment es du a terme a 4°C amb l'ajuda d'una bomba peristàtica *LKB Pump P-1* (Pharmacia) a velocitat 6-8 de la posició 1.

Col·locar tot el volum possible de l'extracte proteic a l'*Econocolumn* amb la reïna, resuspenent la reïna. Passar el contingut a un tub falcon de 50 mL i utilitzar la resta d'extracte proteic per acabar de resuspendre la reïna de l'*Econocolumn*, ajuntant-lo al tub falcon de 50 mL. El tub falcon es deixa donant voltes a una roda durant 30 minuts, així les proteïnes de fusió poden unir-se a la reïna. Passat el temps, es torna a posar el contingut del falcon a l'*Econocolumn*, deixant que es filtri el tampó. Fer dos rentats de la reïna amb NETN-0,7 M NaCl (fet afegint NaCl al tampó NETN, fins a una concentració final de 0,7 M) deixant que es filtri. Finalment, s'afegeix el volum necessari de NETN a l'*Econocolum* per evitar que la reïna s'assequi, i es pot guardar a 4°C.

C.2.6.4.D. Elució de les proteïnes fusionades a GST de la reïna *Glutathione*

Sepharose 4B

Donat que la purificació de les proteïnes fusionades a GST es fa mitjançant la mateixa reïna *Glutathione Sepharose 4B* que s'utilitza per a l'assaig de *Pulldown*, seria possible realitzar-lo sense eluir les proteïnes de la reïna. No obstant això, per augmentar el temps de conservació de les proteïnes fusionades a GST purificades, es van eluir de la reïna, podent així conservar-les a -80°C.

MATERIALS I MÈTODES

Per a dur a terme l'elució (sempre a 4°C), es realitza un rentat a l'*Econocolumn* omplint-la de tampó TST + 5 mM DTT i deixant que es filtri. Resuspendre la reïna amb un volum de tampó TST + 5 mM DTT + 20 mM GSH (Glutatió) reduït i vortexar breument. Deixar la mostra 2 minuts en gel, vortexar 1 minut, tornar a deixar en gel 2 minuts i finalment, vortexar enèrgicament. Deixar filtrar recuperant en tubs eppendorf l'eluit, que contindrà les proteïnes fusionades a GST. Resuspendre altre cop la reïna amb tampó TST + 1 mM DTT + 10 mM GSH, vortexar i deixar filtrar recuperant l'eluit en tubs eppendorf. Repetir aquest últim pas 2 vegades més. Així, finalment s'obtenen 4 fraccions eluïdes. Es comprova la concentració de proteïna recuperada a cada fracció carregant una alíquota de cada una en un gel SDS-PAGE i tenyint-lo amb coomassie, descartant les fraccions menys concentrades.

Amb les fraccions eluïdes més concentrades es procedeix a la seva dialització. S'ajunten totes les fraccions en una membrana de diàlisi Spectra/Por (Spectrum Laboratories, Inc) i es col·loca en un vas de precipitats amb 1L de tampó de diàlisi. Es deixa en agitació a 4°C 1h. Passat el temps es canvia la membrana a un vas de precipitats amb tampó de diàlisi nou i es deixa una altra hora a 4°C en agitació. Aquest procés es repeteix una última vegada i finalment, recuperem el contingut de la membrana en tubs eppendorf. Mitjançant SDS-PAGE i tinció amb coomassie es carreguen diferents dilucions de les mostres dialitzades fins aconseguir equiparar les concentracions de les diferents proteïnes fusionades a GST, a fi d'utilitzar la mateixa quantitat de cada una a l'assaig de GST-Pulldown.

TST

50 mM Tris-HCl pH8
150 mM NaCl
0,1% Tritó X-100

Tampó de diàlisi

20 mM HEPES
20% glicerol
0,2 mM EDTA
0,1 mM NaCl
0,1 mM PMSF
1 mM DTT

C.2.6.4.E. Assaig de GST-Pulldown

Col·locar 20 µL de la reïna *Glutathione Sepharose 4B* en tubs eppendorf de 0,5 mL. Afegir 480 µL de tampó D a cada eppendorf. Mesclar 1 minut a una roda i centrifugar 1 minut a 3000 rpm en una centrífuga eppendorf. Descartar el sobrenedant i repetir aquest rentat amb tampó D un cop més. A partir d'aquest punt, tot l'experiment s'ha de realitzar a 4°C. Treure el sobrenedant de la reïna i afegir 20 µL de la proteïna fusionada a GST adient. Afegir 460 µL de tampó D a cada tub i incubar 1 h a una roda. Passat aquest temps, es centrifuguen les mostres a

3000 rpm 1 minut i es descarta el sobrenedant. Després es realitzen tres rentats amb 480 µL de tampó D, deixant 5 minuts agitant els tubs a la roda i centrifugant a 3000 rpm 1 minut. Es treu l'últim sobrenedant i s'afegeixen 20 µL de l'extracte de proteïnes nuclears de cèl·lules S2 (veure apartat C.2.6.4.A) i 460 µL de tampó D. Es deixen els tubs agitant 2 h a la roda. Passades les dues hores, es centrifuguen els tubs a 3000 rpm 1 minut, es descarta el sobrenedant i es realitzen 6 rentats de 5 minuts amb 480 µL de tampó D, centrifugant a 3000 rpm 1 minut i descartant els sobrenedants. Finalment, resuspendre cada reïna en 20 µL de tampó de càrrega de proteïnes PLB 5X (1-2X final) i β-mercaptoetanol (1,5 M final), vortexar-les, escalfar-les a 95°C 5 minuts, centrifugar-les a 3000 rpm 1 minut per a baixar la reïna i carregar 20 µL de cada sobrenedant en un gel SDS-PAGE. Finalment, per *western blot* amb anticossos específics, detectem quines proteïnes fusionades a GST han retingut dCtBP-L a la reïna *Glutathione Sepharose 4B*. En el nostre cas utilitzem l'anticòs anti-Flag, que detecta dCtBP-L fusionada a l'epítip Flag, que havíem transfectat transitòriament a cèl·lules S2. Com a referència de la quantitat de proteïna dCtBP-L fusionada a Flag present a l'extracte, es van agafar 2 µL (un 10% del que es va utilitzar a cada *Pulldown*) de l'extracte de proteïnes nuclears d'S2, es va preparar de la mateixa manera que les altres mostres i es va carregar al gel SDS-PAGE. És el que anomenem "10% input".

Tampó D

0,1 M KCl
20% glicerol
20 mM HEPES-KOH pH 7,9
0,2 mM EDTA pH8
0,05% NP-40
0,1 mM PMSF
0,5 mM DTT

C.2.6.5. Immunolocalitzacions a cèl·lules

Per a la realització d'immunolocalització de proteïnes a cèl·lules S2, primer es tracten els cobreobjectes per a que les cèl·lules puguin unir-se, afegint 50 µL de concanavalina-A (0,5 mg/mL) (Sigma). Es deixa assecat la concanavalina-A (aproximadament 1 h a la campana) i es col·loquen els cobreobjectes a plaques de cultius de 6 pous de 35 mm de diàmetre cada un (Corning). S'afegeix 1 mL de medi Schneider suplementat per pou i seguidament s'afegeixen $1 \cdot 10^5$ cèl·lules a analitzar. Les cèl·lules s'incuben 1-2 h a temperatura ambient, per a que s'enganxin al cobreobjectes amb concanavalina-A. Un cop enganxades les cèl·lules, es descarta el medi i es renta amb PBS 1X, afegint-hi 1 mL i deixant-lo en agitació 10 minuts. A continuació

MATERIALS I MÈTODES

es fixen les cèl·lules traient el PBS, afegint 200 µL de paraformaldehid 4% (diluït en PBS) i incubant 15 minuts a temperatura ambient. Seguidament, es treu el paraformaldehid i es renta amb 1 mL de PBS 1X en agitació 15 minuts. En acabat, fer dos rentats més amb 1 mL de PBS-Tritó-BSA (PBS 1X; 0,1 % Tritó-X100; 1 g/L BSA), 10 minuts en agitació. Tot seguit, es treu el PBS-Tritó-BSA i es posen 200-300 µL de la dilució corresponent de l'anticòs primari (en PBS-Tritó-BSA). S'incuba 1 h a temperatura ambient, i després tota la nit a 4°C, sempre en agitació. Al dia següent es fan tres rentats amb 1 mL de PBS-Tritó-BSA, 10 minuts en agitació. Després, es treu el PBS-Tritó-BSA i s'afegeixen 300 µL d'una dilució 1:400 (en PBS-Tritó-BSA) de l'anticòs secundari corresponent. S'incuba 1 h a temperatura ambient en agitació. En acabat, es renta dues vegades amb 1 mL de PBS-Tritó (PBS 1X; 0,1 % Tritó-X100), 10 minuts en agitació. Tot seguit es fan dos rentats més amb 1 mL de PBS 1X, 10 minuts en agitació. Finalment, es munten els cobreobjectes en portaobjectes, afegint 3-5 µL de DAPI-Mowiol: 100 µL de DAPI (4',6-diamidino-2-phenylindole) a 2 ng/µL + 900 µL de Mowiol (antiFADE Mowiol, Calbiochem-Novabiochem). Les mostres s'han analitzat per microscopia òptica de fluorescència (microscopi Nikon Eclipse E-1000) i/o per microscopia confocal (microscopis Leica TCS SPE i Leica TCS SP5).

C.2.6.6. Immunolocalitzacions a embrions

Es recullen embrions de l'estadi desitjat, en el nostre cas normalment recollíem els embrions de 0-24 h de creuaments a 18°C. Es decorionen 2-3 minuts amb hipoclorit sòdic al 5% aproximadament. Seguidament, es renten amb PBS 1X + 0,1% Tritó-X100. Es recullen els embrions i es passen a un vial de vidre que contingui 2 mL de PEM (0,1 M PIPES pH 6,9; 1 mM EGTA; 1mM MgSO₄), 2 mL d'heptà i 250 µL de formaldehid 37%. S'incuben durant 20 minuts en agitació constant. S'elimina la fase aquosa inferior, s'afegeixen 2 mL de metanol i es vortexa durant 20 segons. Es recullen els embrions que queden al fons del vial, que són els que han estat fecundats i han perdut la membrana vitel·lina, amb una pipeta *pasteur* i es transfereixen a un tub eppendorf amb metanol. Es renten 3 vegades amb metanol i es guarden un mínim de 12 h a 4°C abans de continuar.

Per a la immunolocalització cal transferir els embrions a un tub eppendorf de 0,5 mL. Es renten ràpidament 3 cops amb PBS 1X + 0,3% Tritó-X100, per eliminar el metanol. Seguidament, es bloquegen (3 vegades) amb solució de bloqueig (PBS 1X, 0,3% Tritó-X100, 2% BSA), 10 minuts agitant-se a la roda. S'incuben tota la nit a 4°C en agitació constant a la roda,

amb 600 µL d'anticòs primari, diluït en solució de bloqueig. Al dia següent, es renten els embrions ràpidament 3 cops amb la solució de bloqueig, i 4 cops agitant-se durant 10 minuts a la roda, amb la solució de bloqueig. Tot seguit, s'incuben els embrions 2 h a temperatura ambient amb l'anticòs o anticossos secundaris, diluïts 1:400 en solució de bloqueig, protegits de la llum. Acabada la incubació, els embrions es renten ràpidament 3 cops amb PBS 1X + 0,3% Tritó-X100. Seguidament, es fan 4 rentats de 10 minuts amb la solució de bloqueig. En acabat, es fa un rentat de 10 minuts amb PBS 1X. Després, s'incuben els embrions en 500 µL de PBS 1X + 5 µL de DAPI (2 ng/µL), 15 minuts a temperatura ambient, agitant-se a la roda. Finalment, es renten 3 cops durant 5 minuts amb PBS 1X + 0,3% Tritó-X100, i es munten en antiFADE Mowiol. Les mostres s'han analitzat per microscopia confocal (microscopis Leica TCS SPE i Leica TCS SP5).

C.2.6.7. Immunolocalitzacions a discos imaginals de larves

Les larves, en PBS, es seccionen per la meitat. La meitat posterior es descarta i la meitat anterior es gira deixant en contacte amb el medi els òrgans interns. Es fixen aquests teixits en paraformaldehid 4% (diluït en PBS) 15 minuts a temperatura ambient en agitació en una roda. En acabat, es renten ràpidament 3 cops amb PBS 1X + 0,3% Tritó X-100, i després durant 20 minuts en agitació en una roda. Seguidament, es bloquegen (3 vegades) durant 10 minuts amb solució de bloqueig (PBS 1X, 0,3% Tritó-X100, 2% BSA), agitant-se a la roda. Tot seguit, s'incuben els teixits tota la nit a 4°C en agitació a la roda, amb 600 µL d'anticòs primari, diluït en solució de bloqueig. Al dia següent, es renten ràpidament 3 cops amb la solució de bloqueig, i 4 cops agitant-se durant 10 minuts a la roda, amb la solució de bloqueig. Tot seguit, s'incuben 2 h a temperatura ambient amb l'anticòs o anticossos secundaris, diluïts 1:600 en solució de bloqueig, protegits de la llum. Acabada la incubació, es renten ràpidament 3 cops amb PBS 1X + 0,3% Tritó-X100. Seguidament, es fan 4 rentats de 10 minuts amb la solució de bloqueig. En acabat, es fa un rentat de 10 minuts amb PBS 1X. Després, s'incuben en 500 µL de PBS 1X + 5 µL de DAPI (2 ng/µL), 15 minuts a temperatura ambient, agitant-se a la roda. Finalment, es renten 3 cops durant 5 minuts amb PBS 1X + 0,3% Tritó-X100, es disseccionen els teixits recollint els discos imaginals d'ala, halteri i potes, i es munten en antiFADE Mowiol. Les mostres s'han analitzat per microscopia confocal (microscopis Leica TCS SPE i Leica TCS SP5).

D. RESULTATS

D.1. El factor GAGA de *Drosophila melanogaster* a sistemes heteròlegs

Com ja s'ha vist (veure Introducció, apartat A.3.1.1), la proteïna GAGA, tot i estar caracteritzada com a factor activador de la transcripció per a multitud de gens (Wilkins et al. 1997), actua com a repressor del gen que la codifica (*Trl*). Aquesta sorprenent faceta de GAGA sobre l'activitat del seu propi gen ha estat observada a la línia cel·lular S2 de *Drosophila melanogaster*. En aquest context, s'ha vist per transfecció transitòria que la sobreexpressió de GAGA comporta una ràpida i eficient repressió de l'activitat del promotor de *Trl* fusionat a un gen *reporter* (Kosoy et al. 2002).

Per mirar d'indagar sobre el mecanisme que porta a GAGA a la repressió de *Trl*, el primer que vam fer va ser comprovar si aquest mecanisme és un fet diferencial de *Drosophila melanogaster*, o si per contra està conservat al llarg de l'evolució. Per veure això vàrem dissenyar una sèrie d'experiments de sobreexpressió de GAGA per transfecció transitòria a dues línies cel·lulars d'altres organismes. Vam escollir la línia cel·lular HeLa, provinent d'un tumor de cèrvix humà, i la línia cel·lular WR DV1 de *Drosophila virilis* que, al igual que les S2, són d'origen embrionari. Òbviament, la línia de cèl·lules humanes està molt més allunyada evolutivament de *D. melanogaster* que no pas la línia de *D. virilis*. Tot i així, la divergència evolutiva entre *D. melanogaster* i *D. virilis* és considerable (uns 60 milions d'anys segons alguns estudis (Beverley et al. 1984)).

D.1.1. GAGA a cèl·lules de *Drosophila virilis* reprimeix *Trl*

Per analitzar si el mecanisme de repressió de *Trl* per GAGA es troba conservat a *D. virilis* vàrem comprovar si tant els promotors de *Trl* com les proteïnes GAGA d'ambdues espècies són funcionalment equivalents, és a dir, si els promotors de *Trl* de *D. virilis* i *D. melanogaster* eren igualment reprimits per les proteïnes GAGA d'ambdues espècies. A més, tenint en compte que el fons genètic d'ambdues espècies és diferent, els possibles factors accessoris necessaris per aquest mecanisme de repressió també poden diferir. Per aquesta raó, tots els experiments els vàrem dur a terme tant a cèl·lules de *D. melanogaster* (S2) com de *D. virilis* (WR DV1).

RESULTATS

El primer pas per realitzar aquests estudis va ser clonar tant la regió promotora com el cDNA (DNA complementari, còpia de l'mRNA madur) del gen *Trl* de *D. virilis*. El gen *Trl* a *D. virilis*, al igual que a *D. melanogaster*, s'ha vist que codifica per a dues isoformes, anomenades GAGA-A i GAGA-B, de 556 aa i 590 aa respectivament (Lintermann et al. 1998). Per alineament de seqüència es pot veure com la GAGA-A i GAGA-B de *D. virilis* corresponen a GAGA₅₁₉ i GAGA₅₈₁ de *D. melanogaster*, respectivament (veure figura D.1).

Com es pot veure a la figura D.1 totes dues isoformes estan molt conservades entre espècies, sobretot per la part N-terminal. A més d'aquesta conservació a nivell de proteïna, també s'ha vist que l'estructura d'introns i exons del gen *Trl* és igual en ambdues espècies (Lintermann et al. 1998). Donat que les dues isoformes de GAGA a *D. melanogaster* són igualment competents en reprimir *Trl*, i que per a la repressió el domini Q (que és pràcticament l'únic en que es diferencien les dues isoformes, tant de *D. melanogaster* com de *D. virilis*) és prescindible (Kosoy et al. 2002), pel nostre estudi comparatiu ens hem centrat en la GAGA₅₁₉ de *D. melanogaster* i la seva contrapartida a *D. virilis*, la GAGA-A.

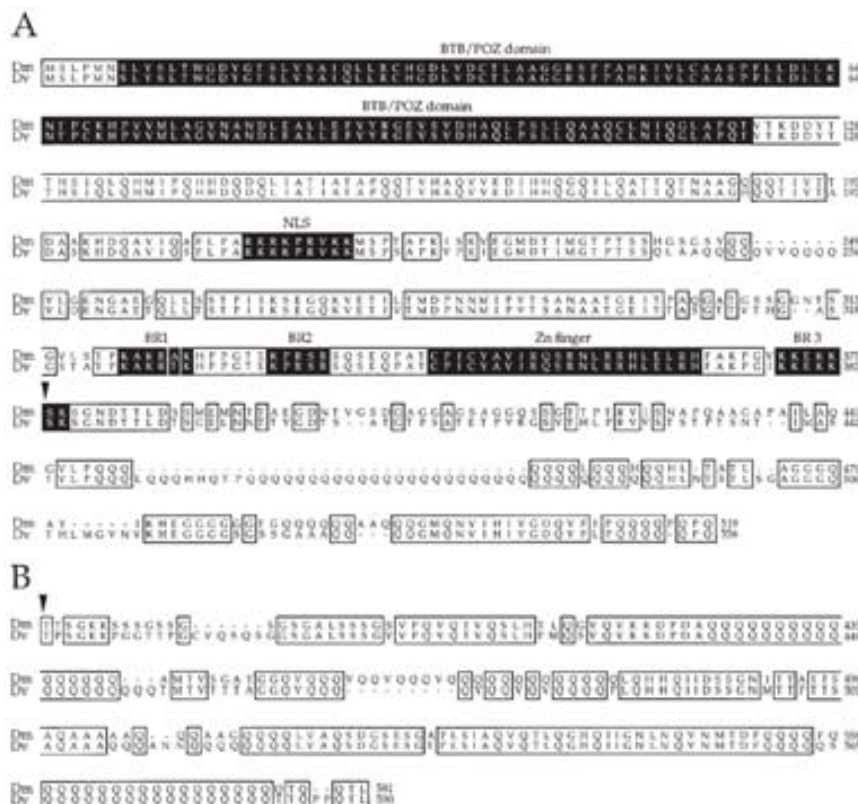


Figura D.1. Comparació de seqüència entre les isoformes de GAGA de *D. melanogaster* i les de *D. virilis*. Les dues isoformes de cada espècie són idèntiques fins l'aa 378 i 383, per *D. melanogaster* i *D. virilis* respectivament (marcat amb un triangle negre). **(A)** Comparació entre GAGA₅₁₉ de *D. melanogaster* i GAGA isoforma A de *D. virilis*. **(B)** Comparació entre GAGA₅₈₁ de *D. melanogaster*, a partir de l'aa 378, i GAGA isoforma B de *D. virilis*, a partir de l'aa 383. Els aa conservats es troben enquadrats. En negre es marquen el domini BTB/POZ, la possible NLS, el dit de Zinc i les regions bàsiques (BRs). Els guions són gaps introduïts per a maximitzar l'alineament (Lintermann et al. 1998).

Per a clonar la regió promotora de *Trl* de *D. virilis*, basant-nos en les dades de FlyBase (<http://flybase.org/>), vam dissenyar un parell d'encebadors que amplifiquessin 2967pb *upstream* del primer codó ATG de la regió codificant. Per PCR sobre DNA genòmic obtingut de la línia cel·lular WR DV1, vam amplificar aquesta zona i la vam inserir al vector pGL3, que conté el gen *reporter* luciferasa, per testar la seva funcionalitat com a regió promotora de *Trl* de *D. virilis*. Un alineament de seqüències entre les regions promotores de *Trl* a *D. melanogaster* i *D. virilis* mostra que la conservació nucleotídica d'ambdues és molt limitada (veure Annexos, apartat H.1.3), tot i que les dues tenen molts llocs potencials d'unió per a GAGA.

Per a clonar el cDNA corresponent a la isoforma GAGA-A vàrem procedir a fer RT-PCR d'extractes d'RNA total de la línia cel·lular de *D. virilis*. El cDNA resultant el vam inserir al vector d'expressió pActppa.

Així doncs, amb les dues construccions corresponents al cDNA i a la regió promotora de *Trl* a *D. virilis*, i amb les seves contrapartides de *D. melanogaster*, vàrem realitzar assajos de transfecció transitòria tant a cèl·lules S2 com a cèl·lules WR DV1. Per a dur a terme aquest assajos vàrem transfectar concentracions fixes del plasmidi que conté el promotor de *Trl* fusionat al gen *reporter* luciferasa, amb les quals cotransfectàvem concentracions creixents del plasmidi d'expressió amb una o l'altra proteïna GAGA, i analitzàvem com s'afectava l'activitat dels promotors (veure figura D.2).

Com es pot veure a la figura D.2, tant la GAGA₅₁₉ de *D. melanogaster* com la GAGA-A de *D. virilis* són capaces de reprimir, de manera dosi dependent, tots dos promotors de *Trl*, tant a cèl·lules S2 com a cèl·lules WR DV1. A més, també es pot veure que la GAGA-A de *D. virilis* sembla lleugerament menys eficient per reprimir qualsevol dels dos promotors. Tot i que a la figura es mostren els valors d'activitat luciferasa dels dos promotors, normalitzada i relativitzada al valor inicial, les activitats luciferasa inicials absolutes dels dos promotors, als dos tipus cel·lulars, són molt similars.

Com a control, es realitzaren *western blots* amb els extractes proteics de les cèl·lules emprades pels assajos de transfecció (veure figura D.2.C, només es mostra el *western blot* d'S2 ja que el de WR DV1 és similar). Com es pot veure, totes dues proteïnes GAGA s'expressen correctament, tot i que la GAGA-A de *D. virilis* apareix menys intensa al *western blot*, cosa que pot voler dir que es degradi més ràpidament que la GAGA₅₁₉ o simplement que l'anticòs tingui menys afinitat per ella, ja que hem emprat l'anti-GAGA fet a rata contra la proteïna GAGA₅₁₉ de

RESULTATS

D. melanogaster. D'altra banda, que hi hagués menys quantitat de GAGA-A als extractes podria explicar la menor activitat repressora de la proteïna de *D. virilis*.

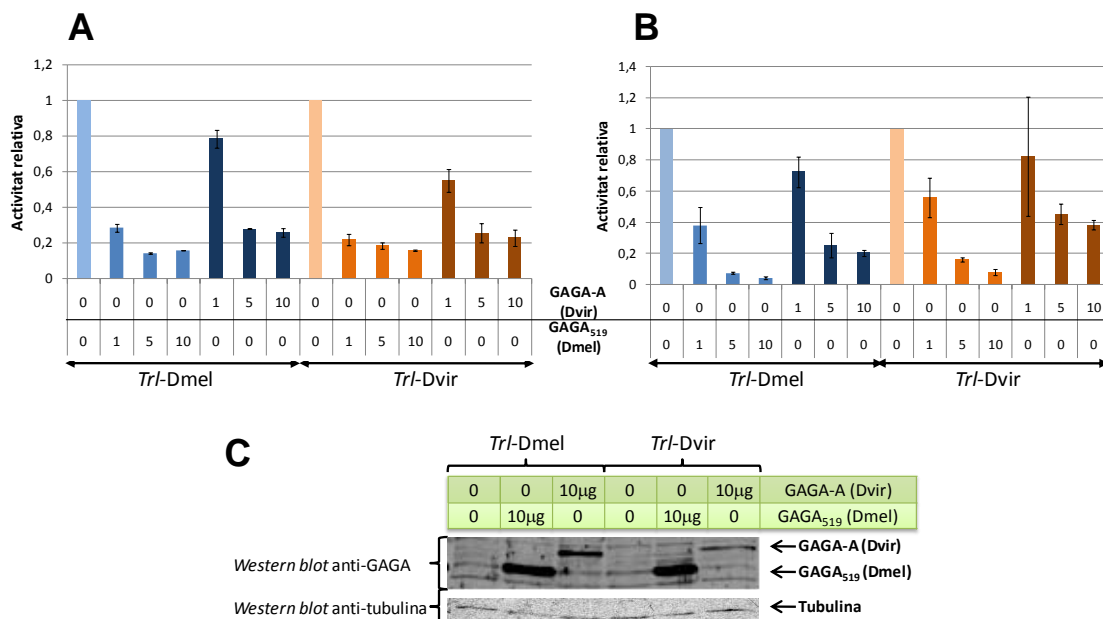


Figura D.2. La repressió de *Trl* per GAGA està conservada a *D. virilis*. Sobre una quantitat fixa de promotor de *Trl* (3 µg) de *D. melanogaster* (*Trl*-Dmel a la figura) o *D. virilis* (*Trl*-Dvir a la figura), es varen afegir quantitats creixents (1, 5 i 10 µg) de plasmidi d'expressió de cada una de les proteïnes GAGA. **(A)** Valors mitjans normalitzats de l'activitat luciferasa, obtinguts a cèl·lules S2 de *D. melanogaster*, relativitzats al primer valor. **(B)** Valors mitjans normalitzats de l'activitat luciferasa, obtinguts a cèl·lules WR DV1 de *D. virilis*, relativitzats al primer valor. Totes dues gràfiques mostren els valors mitjans de quatre experiments independents. **(C)** Control de western blot corresponent a una transfecció (A) revelat tant amb anticossos anti-GAGA de rata com amb anti-tubulina (com a control de càrrega). **Nota:** L'activitat luciferasa inicial absoluta dels dos promotors, als dos tipus cel·lulars, és similar.

D.1.2. GAGA a cèl·lules HeLa humanes activa *Trl*

Per testar si el mecanisme de repressió està conservat al llarg de l'evolució, a més d'emprar cèl·lules d'una altra espècie de *Drosophila*, també vàrem provar el nostre sistema a cèl·lules de mamífer, concretament a la línia de cèl·lules HeLa, d'origen humà. A vertebrats, humans inclosos, no s'ha descrit cap proteïna homòloga a GAGA. No obstant això, certs estudis sí descriuen l'existència de proteïnes d'unió a DNA amb una especificitat similar a GAGA i fins i tot s'ha vist que, a ratolí, una d'aquestes proteïnes podria tenir cert grau d'homologia estructural amb GAGA, ja que els anticossos fets contra la GAGA de *D. melanogaster* són capaços de reconèixer-la (Bevilacqua et al. 2000).

Emprant transfecció transitòria a cèl·lules HeLa, vàrem analitzar l'activitat del promotor de *Trl* en resposta a la sobreexpressió de quantitats creixents de GAGA₅₁₉ (veure figura D.3.A). Tots els experiments els vam fer només amb la isoforma GAGA₅₁₉ per la mateixa raó que el cas

anterior, ja que ambdues isoformes de GAGA són igualment competents en repressió i el domini Q que les diferencia és prescindible per aquesta funció (Kosoy et al. 2002). Com es pot veure a la figura D.3.A, el resultat va ser molt sorprenent ja que no només GAGA no reprimeix *Trl* a HeLa, sinó que l'activa eficientment, de manera dosi dependent. D'aquesta manera, es confirma que el mecanisme que porta a GAGA a comportar-se com a repressor del seu propi promotor no es troba conservat a cèl·lules HeLa. Probablement, el factor o factors que actuen juntament amb GAGA a cèl·lules de *Drosophila* per a donar la repressió de *Trl* no es troben presents a HeLa i per tant GAGA pot actuar com el que és, un activador transcripcional.

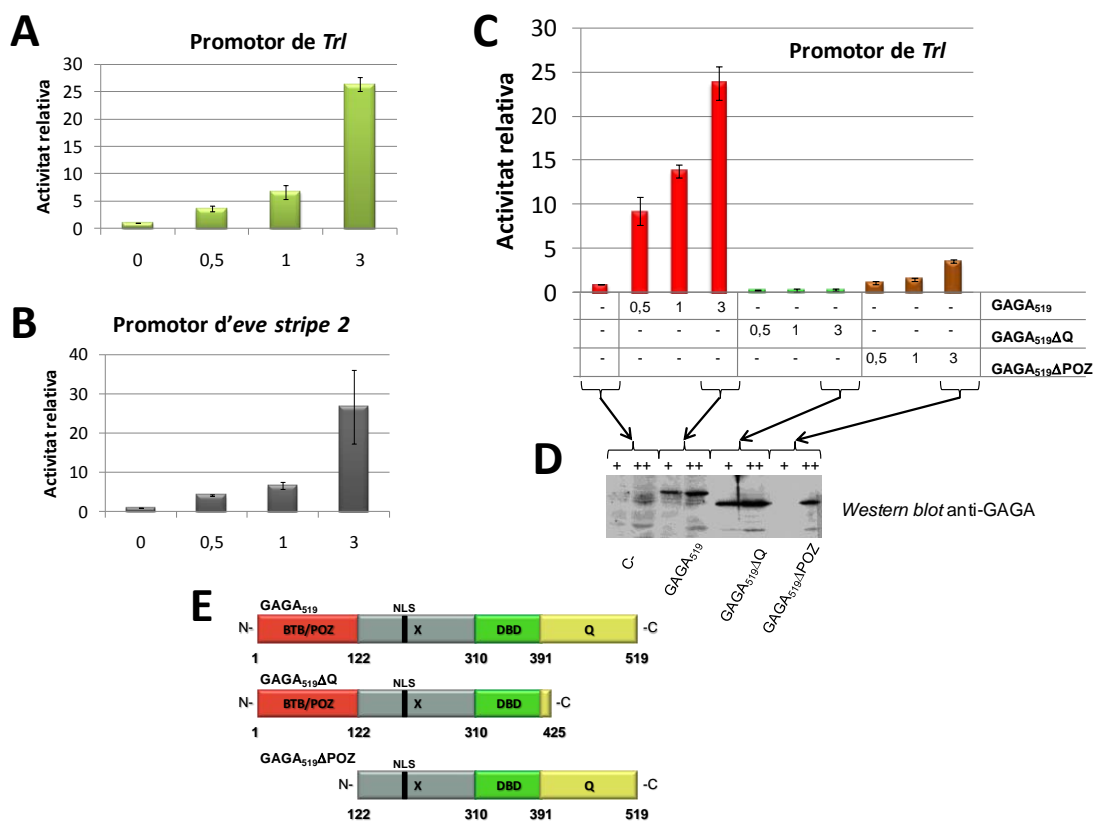


Figura D.3. Transfeccions de GAGA a HeLa. En abscesses s'indiquen els micrograms transfectats de cada plasmidi d'expressió i en ordenades s'indica l'activitat normalitzada del promotor en qüestió, relativitzada al punt d'activitat inicial del promotor, on no es sobreexpressa cap proteïna. **(A)** Efecte de la sobreexpressió per quantitats creixents de vector d'expressió de GAGA₅₁₉ sobre una quantitat fixa d'1 µg del vector amb el promotor de *Trl*. **(B)** Efecte de la sobreexpressió de quantitats creixents de vector d'expressió de GAGA₅₁₉ sobre una quantitat fixa d'1 µg del vector amb el promotor d'*eve stripe 2*. **(C)** Efecte de la sobreexpressió de quantitats creixents dels vectors d'expressió de GAGA₅₁₉, i de les seves formes delecionades, sobre una quantitat fixa d'1 µg del vector amb el promotor de *Trl*. Cada gràfica està feta a partir dels valors mitjans de quatre experiments independents. **(D)** Western blot amb anticòs anti-GAGA d'extractes proteics d'una transfecció com la mostrada a (C). Les fletxes indiquen a que mostres correspon cada carril del gel. Es van carregar dues quantitats de cada mostra (+ i ++). **(E)** Representació esquemàtica de les proteïnes sobreexpressades: GAGA₅₁₉, GAGA₅₁₉ΔQ i GAGA₅₁₉ΔPOZ.

Vàrem voler complementar aquests resultats comprovant si el mecanisme d'activació de GAGA a HeLa funciona de manera anàloga a com funciona a cèl·lules de *Drosophila*, per altres

RESULTATS

promotors activables per GAGA. Per fer-ho vàrem utilitzar el promotor d'*even-skipped* (*eve stripe 2*, el qual anteriorment s'havia descrit que podia ser activat per GAGA de manera dosi dependent, en transfecció transitòria de cèl·lules S2 (Kosoy et al. 2002; Pagans et al. 2002). Així, repetint el mateix assaig de transfecció transitòria a HeLa, però emprant com a plasmidi *reporter* el promotor *eve stripe 2* fusionat al cDNA de la luciferasa, es pot veure com GAGA actua a HeLa de la mateixa manera que a S2, és a dir, activa de manera dosi dependent (veure figura D.3.B). És interessant remarcar que, a HeLa, GAGA activa aquest promotor a un nivell similar a com activa el promotor de *Trl*.

A la llum d'aquests resultats sembla clar que, per a que GAGA actuï com a repressor, és necessari algun factor o factors presents a cèl·lules de *Drosophila*, però no a cèl·lules humanes. Per contra, sembla que el mecanisme pel qual GAGA funciona com a activador sí es troba suficientment conservat a HeLa. A *Drosophila*, s'ha vist que GAGA necessita el seu domini Q per activar i que a més també necessita el seu domini POZ per fer-ho de forma eficient, presumiblement perquè l'oligomerització de GAGA és important per a la seva activitat (Espinass et al. 1999; Vaquero et al. 2000; Pagans et al. 2004). Així doncs, vàrem comprovar si el mecanisme d'activació de *Trl* per GAGA a HeLa té aquests mateixos requeriments. Per fer-ho vàrem repetir els mateixos experiments de transfecció a HeLa, però sobreexpressant també una forma de GAGA sense el seu domini Q (GAGA₅₁₉ΔQ) i una altra forma sense el domini POZ (GAGA₅₁₉ΔPOZ). Com es pot veure a la figura D.3.C, anàlogament a com ja estava descrit a cèl·lules de *Drosophila* pels promotors activables per GAGA, per a l'activació de *Trl* a HeLa GAGA necessita obligatòriament el seu domini Q. A més, d'igual forma, l'activació per una GAGA sense el domini d'oligomerització POZ és molt ineficient.

D.2. Anàlisi detallada del promotor de *Trl* i la seva contribució a la repressió

Ja inicialment vàrem pensar que la diferència que porta a GAGA a reprimir el promotor de *Trl*, en lloc d'activar-lo com fa amb la resta de promotors coneguts on hi actua, ha de residir en algun element de la seqüència d'aquest promotor (Kosoy et al. 2002). Per mirar de recolzar aquesta hipòtesi vàrem dur a terme tota una sèrie d'experiments a fi de trobar l'element o elements de la seqüència del promotor de *Trl* responsables de la repressió. Així, vàrem abordar

la qüestió des de tres fronts diferents. Primer, vam realitzar una cerca bioinformàtica al promotor de *Trl* per tal de definir quins factors de transcripció poden estar unint-se juntament amb GAGA a aquest promotor. Segon, vam dur a terme assajos de *footprinting* genòmic *in vivo* amb DMS (dimetilsulfat), tant a cèl·lules S2 com a embrions de *Drosophila melanogaster*, per tal d'identificar els llocs del promotor ocupats *in vivo*, tant en un estat "normal" o basal del promotor (sense sobreexpressar GAGA) com en un estat "reprimat" del promotor, per mitjà de la sobreexpressió de GAGA. Per últim, amb les dades ja conegudes i les noves provinents dels experiments de *footprinting*, vam dur a terme un estudi funcional exhaustiu del promotor, per mitjà de transfecció transitòria a cèl·lules S2 de formes delecionades i/o mutades del promotor de *Trl*.

D.2.1. Cerca bioinformàtica al promotor de *Trl* en busca de motius d'unió per a factors de transcripció

El primer pas d'aquesta anàlisi exhaustiva del promotor de *Trl* va ser la cerca bioinformàtica de possibles llocs d'unió per a factors de transcripció coneguts. Per a dur-la a terme es varen utilitzar dos programes disponibles *on line*: TFSEARCH (<http://molsun1.cbrc.aist.go.jp/research/db/TFSEARCH.html>) i TESS (<http://www.cbil.upenn.edu/cgi-bin/tess/tess>). A més, també vàrem definir quins dels elements del promotor basal (comentats a la Introducció, apartat A.2.2) estaven presents al promotor de *Trl*. Per dur a terme aquestes anàlisis vàrem emprar un fragment del promotor de *Trl* de -679 pb a +737 pb, respecte al lloc d'inici de la transcripció més distal (descriu a Kosoy et al. 2002). Aquest fragment s'ha demostrat que conté tots els elements necessaris per a una expressió correcta de *Trl*, així com per a la repressió per GAGA. A la figura D.4 es recullen els resultats més destacables d'aquestes cerques.

Com es pot veure a la figura D.4, pel que fa a elements del promotor basal, el primer TSS, indicat per la posició +1, no sembla presentar cap dels coneguts. El segon TSS podria tenir una caixa TATA (situada a uns 29 pb *upstream*) i un DCE (subelement II), que podria no ser funcional, ja que es situa a 24 pb *downstream* del TSS, bastant més lluny del que és habitual (uns 16 pb). El tercer TSS podria contenir un element Inr i un DCE (subelement III), tot i que aquest últim està més a prop del que és habitual, 24 pb *downstream* del TSS, en lloc dels 30 pb

RESULTATS

habituals. Aquests tres TSS són els més probables, però s'ha vist que a cèl·lules S2 podrien existir dos més, utilitzats en menor mesura (Kosoy et al. 2002). Un, situat uns pocs parells de bases *upstream* del TSS més proper a la regió codificant, podria fer servir l'Inr més *upstream* dels dos que hi ha junts, i el DCE (subelement III), que per aquest TSS sí queda a la distància adequada. Un cinquè TSS s'ha vist que podria situar-se a no menys de 50 pb *upstream* del que situem a +1. Així, aquest TSS podria utilitzar l'Inr i el DPE que es troben més *upstream*, entre les posicions -76 i -42.

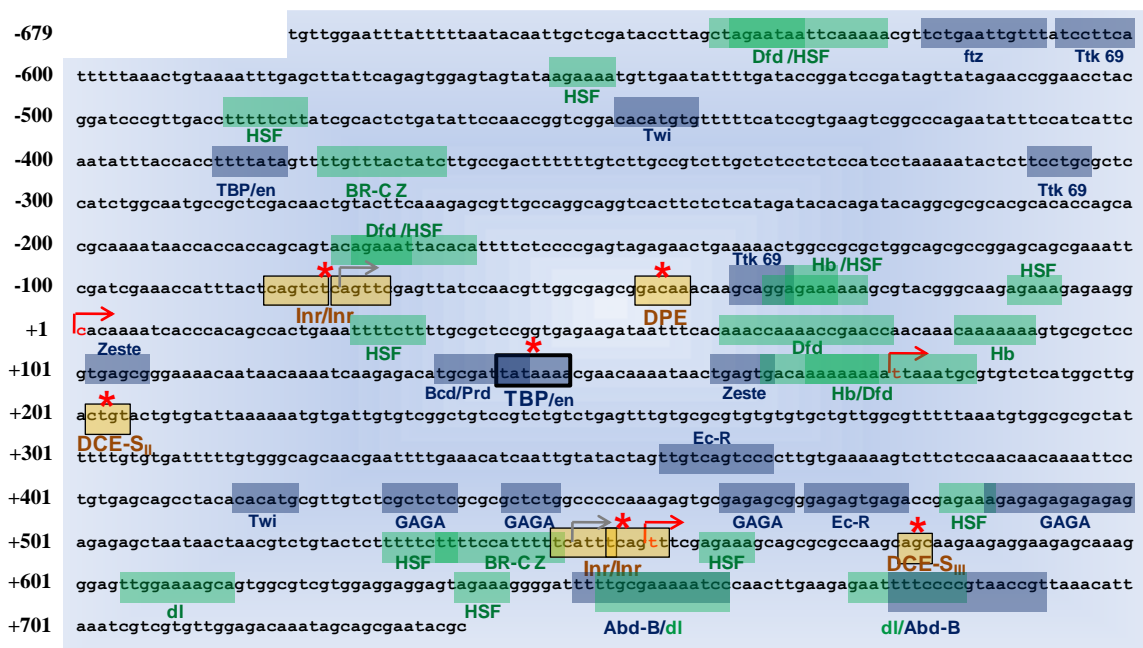


Figura D.4. Resum esquemàtic de la cerca bioinformàtica. Es mostren els possibles llocs d'unió per a factors de transcripció al promotor de *Trl*, de -679 pb a +737 pb. Amb un asterisc vermell es marquen els elements del promotor basal presents: en taronja els Inr, DPE, DCE-S_{II} i DCE-S_{III}; en blau emmarcat en negre la caixa TATA. En verd es poden veure els resultats obtinguts amb el programa TFSEARCH i en blau els resultats del programa TESS. Els llocs d'inici de la transcripció descrits anteriorment (Kosoy et al. 2002) es marquen amb el nucleòtid corresponent en vermell i una fletxa vermella al damunt. A més, també s'indiquen amb fletxes grises els altres dos possibles TSS, descrits també anteriorment. Nota: els llocs d'unió per a TBP i en (engrailed) solapen, ja que *in vitro* s'ha vist que aquests dos factors poden competir per la caixa TATA (Ohkuma et al. 1990a).

En quant a la predicció de llocs d'unió per a factors de transcripció que donen els programes informàtics, s'obtenen resultats força inconsistents. Així, cada algoritme troba uns FTs diferents, sense coincidir pràcticament. Exemple de la seva poca fiabilitat és el fet que, per exemple, TFSEARCH no troba llocs d'unió per a GAGA a la seqüència del promotor i TESS en troba només quatre quan, a simple vista, ja es poden veure multitud més de seqüències consens GAGAG o CTCTC, a les quals GAGA ha demostrat poder unir-se per *footprinting in vitro*

(Kosoy et al. 2002). Tot i amb això, es pot destacar la presència de múltiples llocs per a HSF (*Heat-shock factor*), trobats per ambdós programes (tot i que només es representen els trobats per TFSEARCH, els trobats per TESS pràcticament són els mateixos). A més, també hi ha múltiples llocs per a Ttk69 (Tramtrack 69), proteïna que s'ha vist que interacciona amb GAGA (Pagans et al. 2002; Pagans et al. 2004). Altres proteïnes que hi podrien unir-se són: Twi (Twist), Ftz (Fushi tarazu), En (Engrailed), Hb (Hunchback), Dfd (Deformed), Zeste, Bcd (Bicoid), Prd (Paired), Ec-R (receptor de l'hormona Ecdisterona), BC-Z (Broad-Complex Z), dl (dorsal) i Abd-B (Abdominal-B).

D.2.2. Anàlisi de l'ocupació del promotor de *Trl* *in vivo* per *footprinting* genòmic amb DMS

A banda de les dades obtingudes *in silico* que hem presentat i dels estudis anteriors de *footprinting in vitro* amb la proteïna GAGA (Kosoy et al. 2002), on es veu que GAGA pot unir-se a multitud de llocs al llarg de tota la regió promotora (les caixes subratllades a la figura D.5.A), les dades *in vivo* sobre quins factors podrien unir-se al promotor de *Trl* es limiten a les obtingudes per CHIP, que demostren que GAGA està unida al promotor de *Trl* a cèl·lules S2 (Kosoy et al. 2002). Per aquesta raó, vàrem estudiar quins llocs del promotor de *Trl* estan ocupats *in vivo*, tant en un estadi normal com en un estadi de sobreexpressió de GAGA.

D.2.2.1. *Footprinting* genòmic a cèl·lules S2

Com a primer pas per descriure l'estat d'ocupació del promotor de *Trl in vivo* vàrem dur a terme assajos de *footprinting* genòmic *in vivo* amb DMS (dimetilsulfat) a cèl·lules S2. Vàrem escollir aquesta tècnica perquè permet, per mitjà d'una metodologia relativament senzilla, l'anàlisi *in vivo* de seqüències, amb una resolució molt alta (a nivell de parells de bases).

Així, primer ens vàrem interessar per l'estat basal o inicial del promotor de *Trl* a cèl·lules S2, és a dir, l'estat en que el promotor de *Trl* segueix la seva regulació natural. Cal destacar que sabem que el promotor de *Trl* en estat basal ja presenta un cert nivell de repressió per GAGA, ja que l'RNAi contra la proteïna fa augmentar l'activitat del promotor (Kosoy et al. 2002). Així, l'estudi de l'estat basal del promotor de *Trl* a S2 podria aportar dades sobre el mecanisme de repressió per GAGA.

RESULTATS

S'ha vist que la fusió de certs fragments del promotor de *Trl* confereixen a un promotor heteròleg i insensible a GAGA com és CMV (Citomegalovirus), la capacitat de ser reprimit per GAGA (Kosoy et al. 2002). Aquests fragments en concret són els FII i FIII (veure figura D.5). Essent l'estudi de la repressió per GAGA un dels objectius d'aquest treball i donada la dificultat experimental que comporta abastar grans regions en assajos de *footprinting*, en aquesta primera aproximació vàrem centrar els estudis de *footprinting in vivo* a aquestes regions suposadament implicades en la repressió.

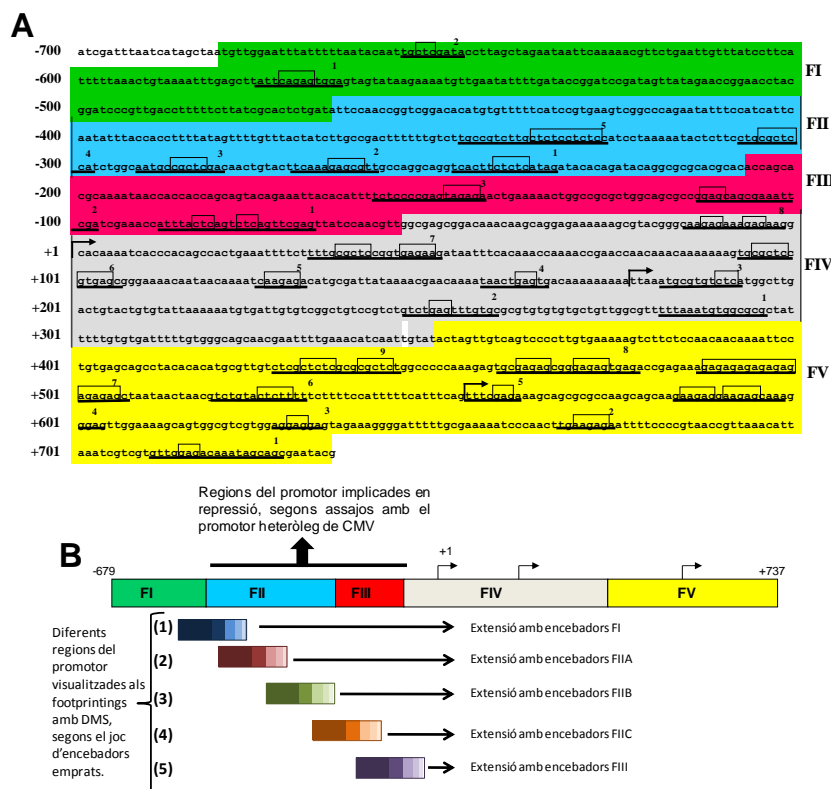


Figura D.5. (A) Seqüència del promotor de *Trl*. Es mostren les subdivisions arbitràries en diferents fragments (FI al FV, marcats amb diferents colors) que s'han utilitzat en diversos experiments, tant anteriors (Kosoy et al. 2002), com del present treball. Les caixes indiquen els llocs d'unió per a GAGA obtinguts per Kosoy i col·laboradors utilitzant *footprintings in vitro* amb DNAsa I. **(B)** Esquema dels diferents fragments. Els fragments FII i FIII es mostren implicats en el mecanisme de repressió per GAGA en assajos amb promotors heteròlegs (Kosoy et al. 2002). També es mostren les regions que es van analitzar per *footprinting* genòmic, indicant el joc d'encebadors emprats. El degradat de les barres representa la gradual pèrdua de resolució de cada *footprinting*, quant més lluny està de l'inici. Les fletxes en angle recte marquen els TSS.

En aquests experiments vam comparar el patró de *footprinting in vitro* amb DMS d'un DNA genòmic nu, extret de cèl·lules S2, amb el patró de *footprinting in vivo* amb DMS de cèl·lules S2 intactes. Amb els encebadors emprats vàrem cobrir les regions FII i FIII, del promotor de *Trl* endogen, a més de part de les regions adjacents (veure figura D.5.B). A la figura D.6.A es poden veure els resultats d'aquests *footprintings*, junt amb un esquema resum a la

figura D.6.B. Als *footprintings* podem trobar zones protegides, on la freqüència de metilació de guanines pel DMS és menor que pel DNA nu, el que es pot interpretar com que algun factor de transcripció unit a la zona està interferint amb aquesta metilació. Aquestes zones apareixeran com a “forats”, ja que les bandes que les formen es veuran menys intenses. D'altra banda també podem trobar zones on la metilació ha sigut més freqüent que pel DNA nu i per tant apareixen més intenses en el *footprinting*. Aquestes zones, que anomenem hipersensibilitats, poden ser reflex d'una major accessibilitat de la seqüència en qüestió en presència de certs factors de transcripció. Aquests factors, que no tenen per que estar units just a sobre de l'hipersensibilitat, provoquen canvis estructurals a la fibra de cromatina, el que afavoreix l'accessibilitat del DMS.

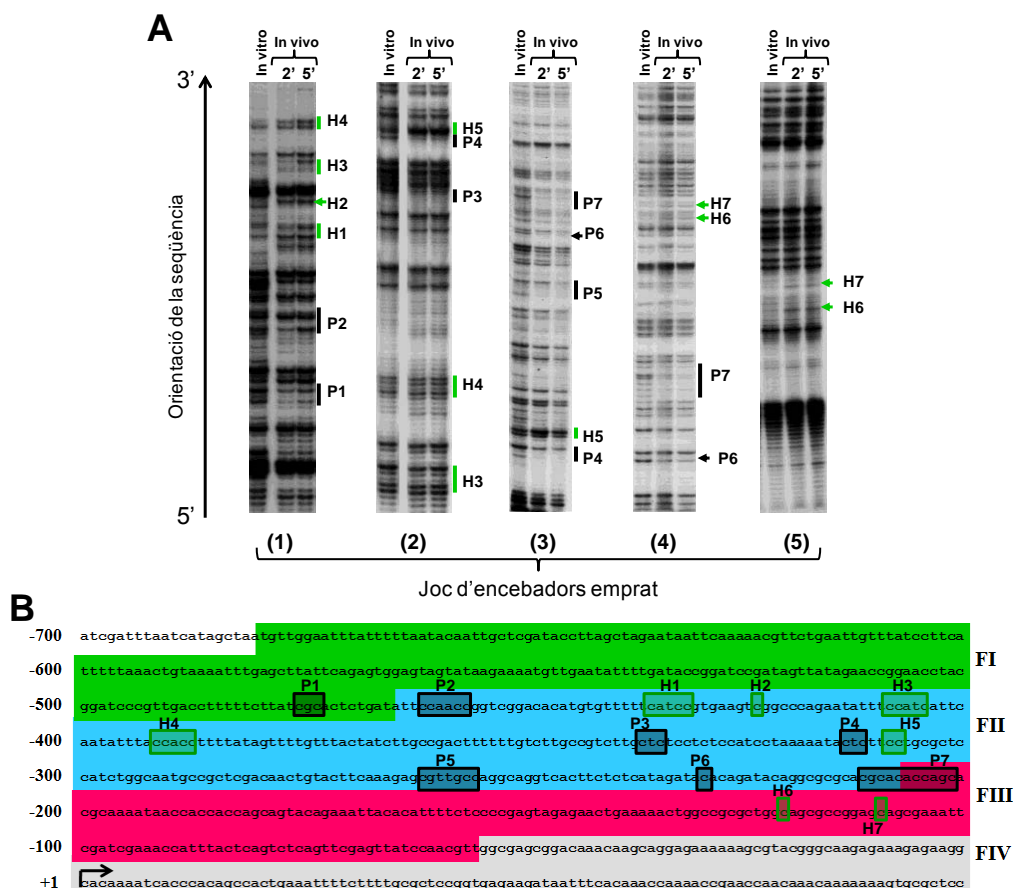


Figura D.6. Anàlisi per *footprinting* genòmic amb DMS de les regions FII i FIII del promotor de *Trl*. **(A)** Diferents regions visualitzades segons el joc d'encebadors emprat (descriu a la **figura D.5.B**). En negre s'indiquen les zones protegides (P) i en verd les zones hipersensibles (H). Es compara el patró obtingut amb DNA nu (*in vitro*) amb l'obtingut a cèl·lules S2 intactes (*in vivo*) utilitzant temps creixents de tractament amb DMS, el que incrementa el senyal de les bandes. Cal destacar com, a la regió solapant entre encebadors, es reproduïxen els patrons de *footprinting*. **(B)** Resum esquemàtic representat sobre la seqüència de nucleòtids, on en negre apareixen les zones protegides i en verd les hipersensibilitats.

RESULTATS

Com es pot veure a la figura D.6, a S2 la seqüència analitzada del promotor de *Trl* presenta tant zones protegides (P) com hipersensibilitats (H). S'ha de fer notar que, degut a que el DMS és específic de G (tot i que també metila A, amb molta menys eficiència) i al disseny dels encebadors, les bandes més intenses que apareixen als gels de *footprinting* corresponen a les C de la figura D.6.B, tot i que també apareixen bandes menys intenses corresponents a T (fruit de la metilació menys específica d'adenines), així com a la resta de nucleòtids, per trencament/metilació espontània de la cadena de DNA. És destacable el fet que sempre es reproduïen els patrons de *footprinting* de les regions solapants entre encebadors, el que confirma la reproduïbilitat dels experiments. Dels resultats obtinguts destaquen especialment la P3, P5 i H6, situades a seqüències on, per *footprinting in vitro* amb DNAsa I, GAGA s'hi uneix (veure figura D.5.A per veure els resultats de *footprinting in vitro* amb DNAsa I, (Kosoy et al. 2002)). També són interessants les P4 i H5, situades a un lloc d'unió per a Ttk 69, trobat per cerca bioinformàtica (veure figura D.4).

D.2.2.2. *Footprinting* genòmic a embrions

Un cop obtinguts els patrons de *footprinting* a cèl·lules S2 vàrem voler comparar aquests amb els que es podien obtenir d'embrions de *Drosophila melanogaster*. A més, l'elecció d'aquest sistema confereix un avantatge afegit, ja que permet la comparació dels patrons de *footprintings in vivo* del promotor de *Trl* en estat basal i en estat reprimat per la sobreexpressió de GAGA. Així, vàrem comparar el patró de *footprinting* d'embrions de 0 a 24 h (incubats a 18°C) de dos tipus de mosques transgèniques: unes control, que porten un transgèn (anomenat *driver*) consistent en el cDNA de GAL4 (una proteïna activadora de llevat) sota el control del promotor de *ptc* (*patched*); i unes altres *ptcGal4-GAGA*, que porten el mateix transgèn de les control i un altre consistent en el cDNA de GAGA₅₁₉ sota el control d'una regió promotora amb múltiples llocs d'unió per a GAL4. Així, als teixits on el promotor de *ptc* és actiu s'expressarà GAL4 i aquest activarà la sobreexpressió de GAGA₅₁₉. L'expressió de GAL4 per si sol no té efectes ni a nivell fenotípic ni a nivell del patró de *footprinting* genòmic *in vivo* amb DMS. Cal remarcar que *ptc* no és d'expressió ubiqua a l'embrió (Hooper et al. 1989) i per tant no totes les cèl·lules de la soca *ptcGal4-GAGA* estaran sobreexpressant GAGA, fet que pot donar uns patrons de *footprinting* només lleugerament diferents als de la soca control. A més, els embrions els vàrem fer créixer a 18°C degut a la forta letalitat que presenten a 25°C els

embrions *ptcGal4-GAGA* (veure més endavant, apartat D.6.2, taula D.1). Aquesta temperatura més baixa fa que la sobreexpressió de *GAGA*₅₁₉ no sigui molt alta, el que augmenta la viabilitat de la soca, però de retruc contribueix a que les diferències observables entre els patrons dels embrions control i els *ptcGal4-GAGA* siguin subtils.

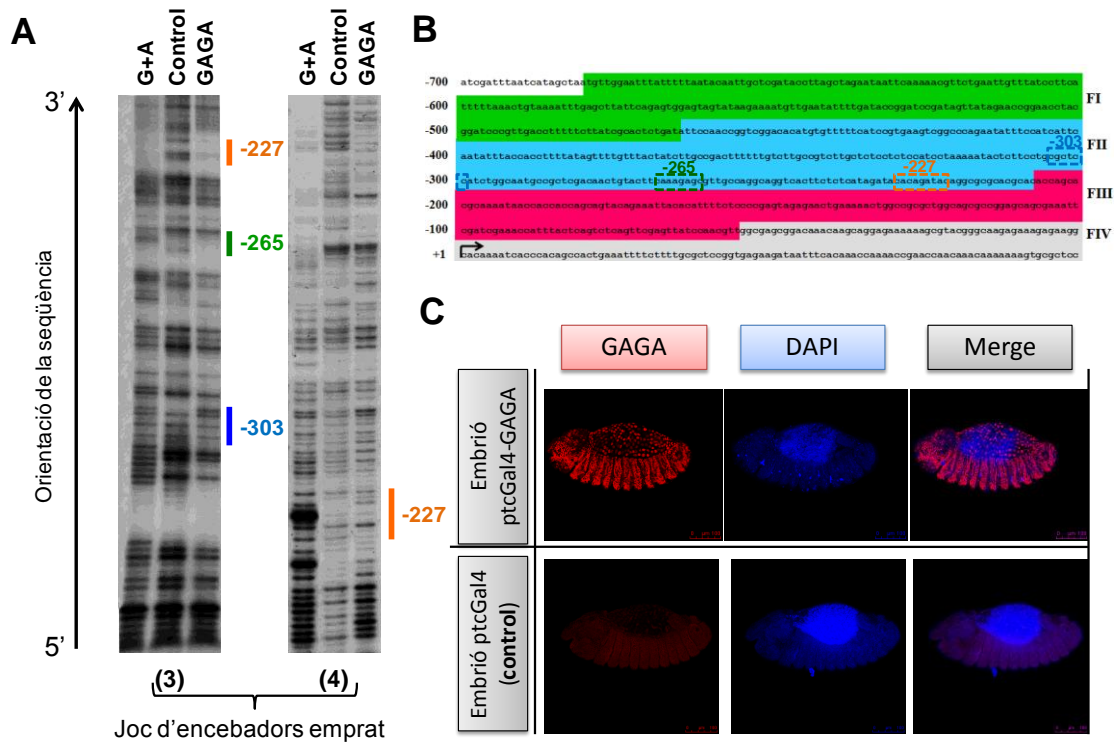


Figura D.7. *Footprinting* amb DMS a embrions. **(A)** Diferents regions visualitzades segons el joc d'encebadors emprat (descriu a la figura D.5.B). Els resultats més interessants s'obtenen amb el joc d'encebadors FII (3), amb el que es poden veure tres llocs que esdevenen protegits en els embrions *ptcGal4-GAGA* (GAGA) respecte als embrions control (control), marcats amb barres de diferents colors. A més, també es mostra l'escala G+A, obtinguda per tractament amb àcid fòrmic *in vitro* (sobre DNA despulcat), que mostra les guanines i adenines de la seqüència. **(B)** Esquema resum sobre la seqüència dels resultats obtinguts on es mostren les tres proteccions obtingudes a (A), amb els colors corresponents. El número utilitzat per anomenar aquestes proteccions és orientatiu de la seva posició relativa al lloc d'inici de la transcripció més distal a la regió codificant. **(C)** Projectió màxima d'imatges obtingudes al microscopi confocal Leica SPE, d'immunotincions amb l'anticòs anti-GAGA, dels embrions emprats als *footprintings in vivo* amb DMS. A dalt, imatge d'un embrió *ptcGal4-GAGA*, que sobreexpressa GAGA sota el patró de bandes que determina el *driver ptcGal4*. A baix, imatge d'un embrió control que únicament porta el *driver ptcGal4*, i que per tant, no sobreexpressa GAGA. A l'esquerra, en vermell el senyal de GAGA, al centre el DAPI i a la dreta la combinació de tots dos senyals (*merge*).

A la figura D.7.A es pot veure un exemple dels resultats obtinguts i a la figura D.7.B es mostra un esquema resum sobre la seqüència del promotor, marcant les seqüències protegides. A la figura D.7.C es mostra el control de la correcta sobreexpressió de *GAGA*₅₁₉, així com el seu patró de sobreexpressió a embrions, conduït pel promotor de *ptc*, que s'expressa a en amples bandes a l'embrió (Hooper et al. 1989). Com es pot observar, vàrem obtenir només proteccions dins la regió FII. Val a dir que les zones protegides que s'indiquen a la figura D.7.B

RESULTATS

no són les úniques que vàrem obtenir, però sí les més reproduïbles, donada l'alta variabilitat que vam trobar. Així, la resta de proteccions que es poden observar a la figura D.7.A, així com moltes altres que hem obtingut als nostres *footprintings*, no les hem tingut en compte, ja que no són reproduïbles. Aquesta alta variabilitat experimental és deguda probablement a l'escàs nivell de sobreexpressió de GAGA₅₁₉ (donat que els individus van ser incubats a 18°C), sumat al fet que, als embrions *ptcGal4-GAGA*, no totes les cèl·lules la sobreexpressen. Les proteccions obtingudes, designades amb la posició aproximada que ocupen al promotor, coincideixen amb alguns dels resultats obtinguts a cèl·lules S2. Així, la protecció que trobem a -303 aproximadament, està molt propera a l'H4 obtinguda a S2, i les proteccions -265 i -227 coincideixen en part amb les P5 i P6 respectivament (veure figura D.6.B). A més, les proteccions -303 i -265 coincideixen amb llocs d'unió per a GAGA, determinats per *footprinting in vitro* amb DNAsa I (veure figura D.5.A per veure els resultats de *footprinting in vitro* amb DNAsa I, (Kosoy et al. 2002)), fet que va en concordança amb l'augment de llocs GAGA ocupats al promotor que cabria esperar per sobreexpressió de GAGA. S'ha de destacar que, tot i que observem protecció a aquestes regions de manera reproduïble, la forma en que apareixen aquestes no és sempre idèntica, com es pot observar en el cas de la protecció -227, que no és evident emprant el joc d'encebadors (4) (veure figura D.7.A).

Tot i amb això, les tres proteccions obtingudes eren bones candidates a ser seqüències implicades en el mecanisme de repressió i per això es van testar per mutació dirigida, com es veurà a l'apartat D.2.3.

D.2.2.3. *Footprinting* genòmic a cèl·lules S2 establement transfectades

Per les característiques dels embrions *ptcGal4-GAGA*, on no totes les cèl·lules sobreexpressen GAGA i les que ho fan, al estar creixent a 18°C, ho fan en poca quantitat, esperàvem que els patrons de *footprinting* variessin poc entre els embrions control i els *ptcGal4-GAGA*. Probablement per aquesta raó només hem pogut descriure tres zones protegides. Per intentar millorar aquests resultats ens vàrem proposar obtenir una línia de cèl·lules S2 establement transfectades que sobreexpressessin GAGA₅₁₉ de forma induïble. Donat que la sobreexpressió de GAGA₅₁₉ a mosques transgèniques és altament letal (veure més

endavant, apartat D.6.2, taula D.1), i per evitar possibles problemes de viabilitat cel·lular, vam obtenir aquesta línia cel·lular inserint el cDNA de GAGA₅₁₉ al plasmidi pMK33 que conté el promotor de la metal·lotioneïna, que s'activa en presència de metalls pesats al medi, tals com el Cu²⁺, el Zn²⁺ o el Cd²⁺, induint així la sobreexpressió de GAGA₅₁₉. Així doncs, vàrem establir aquesta línia de cèl·lules S2 (línia de cèl·lules GAGA), així com una altra línia control que sobreexpressa EGFP (*Enhanced Green Fluorescent Protein*) de la mateixa manera (línia de cèl·lules GFP).

Els resultats obtinguts, amb qualsevol dels jocs d'encebadors emprats van ser del tot inconsistents. Tot i que la línia cel·lular sobreexpressava GAGA₅₁₉ de forma correcta (veure figura D.8) no vàrem ser capaços d'aconseguir patrons de *footprinting* reproduïbles. Potser, la causa d'aquesta irreproduïbilitat venia de que les línies cel·lulars no eren isogèniques, és a dir, que no provenien d'un clon cel·lular, així doncs, no eren genèticament idèntiques. De fet, a la immunolocalització ja es pot veure com no totes les cèl·lules sobreexpressen GAGA₅₁₉ de la mateixa manera (veure figura D.8.B). A més, el propi tractament amb metalls pesats podria estar afectant al promotor i això afectar als patrons de *footprinting* de manera incontrolable.

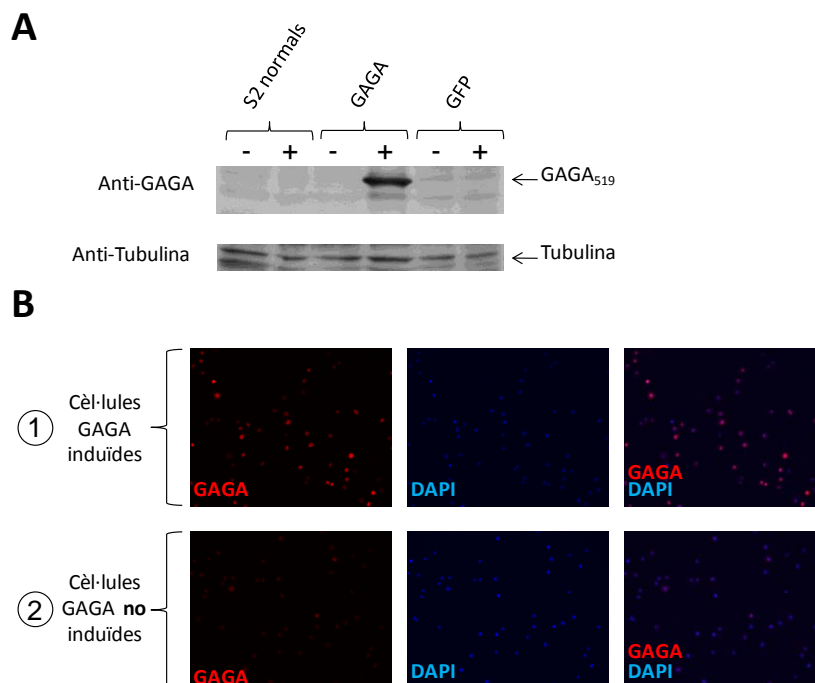


Figura D.8. Caracterització de la línia de cèl·lules GAGA. **(A)** *Western blot* amb anti-GAGA de cèl·lules sense tractar (-) i tractades amb 500µM de CuSO₄ al medi durant 24h (+). D'esquerra a dreta: cèl·lules S2, cèl·lules de la línia GAGA i cèl·lules de la línia GFP. A baix, *western blot* amb anti-tubulina de la mateixa membrana. **(B)** Immunolocalització de cèl·lules de la línia GAGA emprant anti-GAGA. (B1) Cèl·lules GAGA tractades amb 500µM de CuSO₄ durant 24h. (B2) Cèl·lules GAGA sense tractar. Imatges preses amb un microscopi òptic de fluorescència Nikon Eclipse E-1000.

RESULTATS

D.2.3. Anàlisi del promotor de *Trl* per mitjà de delecions i mutacions

Experiments anteriors ja havien delimitat, per mitjà de successives delecions, la regió promotora mínima per a que *Trl* tingués tota la seva activitat transcripcional, tant a cèl·lules S2 com a embrions de *Drosophila*. A la figura D.9 es pot veure un resum d'aquestes delecions i la seva activitat relativa a cèl·lules S2, mesurada per mitjà d'assajos de transfecció transitòria (Ana Kosoy, tesi doctoral 2002).

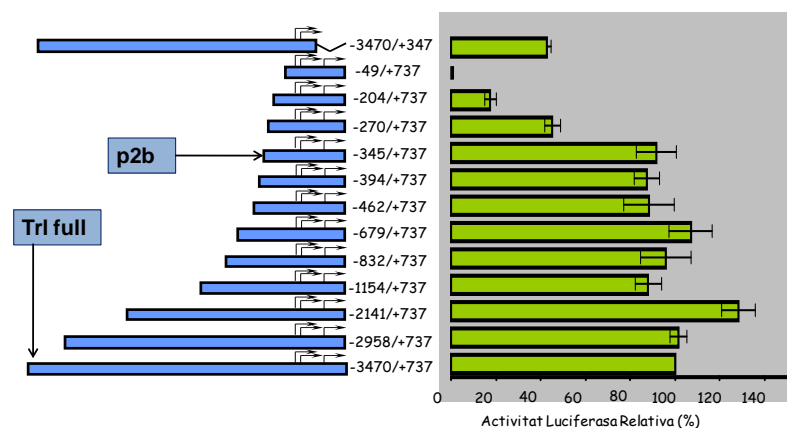


Figura D.9. Resum dels resultats obtinguts per la Dra. Ana Kosoy realitzant successives delecions del promotor de *Trl* i assajant la seva activitat relativa a cèl·lules S2. Els números fan referència a la regió del promotor que comprèn cada construcció emprada, esquematitzades a l'esquerra de la figura. A la dreta es representa l'activitat relativa de la proteïna *reporter* luciferasa, per a cada construcció. Com es pot observar, la regió mínima del promotor que conserva tota la seva activitat és la que va de -345 a +737 del promotor, que anomenem p2b. Per *Trl* full anomenem al fragment més gran clonat de la regió promotora de *Trl*. Extret de la tesi doctoral d'Ana Kosoy (2002).

Nosaltres, per a l'estudi del mecanisme de repressió de *Trl*, vàrem analitzar la funcionalitat del promotor de *Trl* seguint dues estratègies. Primer, vàrem realitzar diferents delecions de la regió promotora, per tractar d'acotar encara més la regió imprescindible del promotor, tant per a la seva activitat normal, com per a la repressió per GAGA. Segon, basant-nos en part en els resultats de *footprinting* obtinguts a embrions (veure apartat D.2.2.2), vàrem realitzar tot un seguit de mutacions al promotor de *Trl*, per tal de determinar la implicació de cada una d'aquestes seqüències en l'activitat de *Trl* i en la repressió per GAGA.

D.2.3.1. Delecions

Per tal de completar i fer més precís el mapatge del promotor de *Trl*, vàrem realitzar tota una sèrie de delecions addicionals, les quals assajàrem per transfecció transitòria a cèl·lules S2 per tal de determinar el seu nivell d'activitat, tant transfectant-les soles, com cotransfectant-les amb una quantitat fixa de GAGA que dóna un nivell mitjà de repressió de *Trl*. Tots els resultats obtinguts (veure figura D.10) es van comparar amb la regió p2b, la mínima que conserva tota l'activitat del promotor i que a més també és perfectament reprimible per la sobreexpressió de GAGA.

Com es pot veure a la figura D.10, excepte les construccions amb una activitat molt baixa (“-345/-33” i “-345/+14”), tota la resta de formes delecionades del promotor de *Trl* poden ser reprimides per GAGA de forma similar, independentment de l'activitat que mostrin en absència de sobreexpressió de GAGA. És de destacar la probable presència d'un important element del promotor de *Trl*, relacionat amb l'activació d'aquest, a la regió compresa entre -345 i -298, ja que la construcció “-298/+737” té aproximadament la meitat d'activitat que p2b. També és interessant el fet que la regió entre +14 i +71 podria contenir un element implicat en la repressió per GAGA, ja que mentre “-345/+71” és reprimible per GAGA, “-345/+14” ja no ho és. Tot i amb això, cal senyalar que l'activitat d'aquest últim promotor és molt pobre, i això el podria fer insensible a l'acció repressora de GAGA, tot i tenir els llocs implicats en repressió intactes.

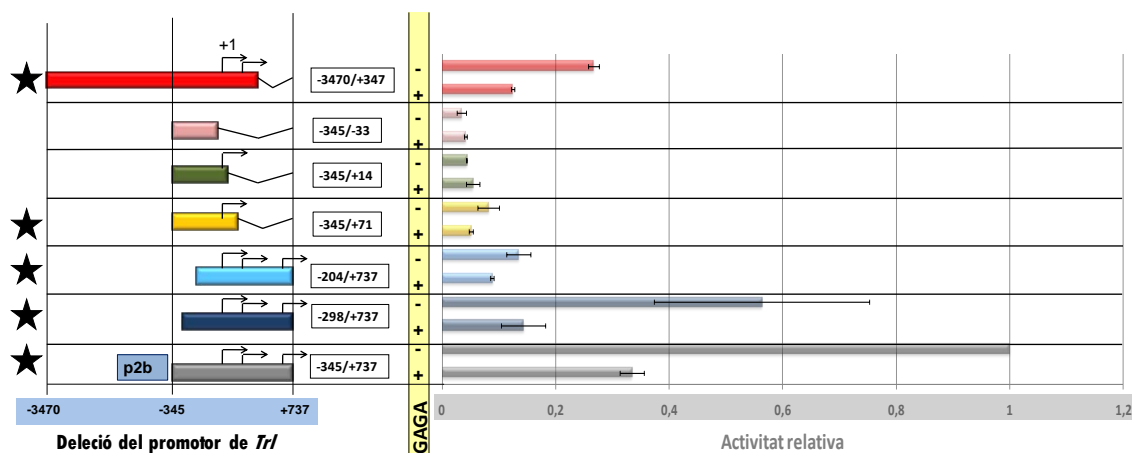


Figura D.10. Resum dels resultats de les transfeccions transitòries a S2 amb les delecions del promotor de *Trl*. A la part esquerra de la figura es representa esquemàticament cada delecció, amb la regió que comprèn i els inicis de transcripció que conserva. A cada transfecció es va utilitzar la quantitat fixa de 3 µg de cada delecció. Al centre de la figura s’indica si la transfecció es feia en absència (-) o presència (+) d’1 µg del vector d’expressió de GAGA₅₁₉. A la part dreta es representen els resultats de les diverses transfeccions a S2, indicant la seva activitat luciferasa normalitzada i relativitzada al valor de l’activitat de p2b sense GAGA. Cada valor és la mitjana d’almenys quatre experiments independents. Amb una estrella s’indiquen les delecions que continuen sent reprimibles per GAGA.

RESULTATS

D.2.3.2. Mutacions

A banda de les delecions successives del promotor de *Trl*, també vàrem realitzar tot un seguit de mutacions de certes seqüències del promotor de *Trl*, a partir de la regió p2b, basant-nos en tres criteris (veure figura D.11).

Primer, vàrem mutar aquelles seqüències que, per mitjà de *footprintings in vivo*, hem vist que esdevenen protegides al sobreexpressar GAGA a embrions (veure figura D.11.B). Aquestes seqüències eren tres (veure apartat D.2.2.2). Dues d'elles, situades al voltant de -303 i -265 respectivament, coincideixen amb llocs GAGA vistos amb anterioritat per *footprinting in vitro* amb DNAsa I (Kosoy et al. 2002). Aquestes dues les vàrem mutar tant per separat com les dues a l'hora, per si la falta d'una es podia compensar amb l'altra. La tercera seqüència, situada al voltant de -227, no coincidia amb cap lloc d'unió potencial per a GAGA. A la llum dels resultats obtinguts (veure figura D.12.A), cap d'aquestes seqüències és determinant pel mecanisme de repressió de *Trl* estudiat, ja que totes les construccions que les portaven mutades eren perfectament reprimibles al ser cotransfectades amb el vector d'expressió de GAGA. Totes les construccions mostraven una activitat inicial (sense GAGA) similar a p2b, amb l'excepció de p2b-227, que presenta aproximadament el 50% de l'activitat inicial dels altres promotors. Aquest fet indica que aquesta seqüència està implicada en la unió d'algun factor important per a l'activitat d'aquest promotor, a nivell d'activació, però no a nivell de repressió per GAGA.

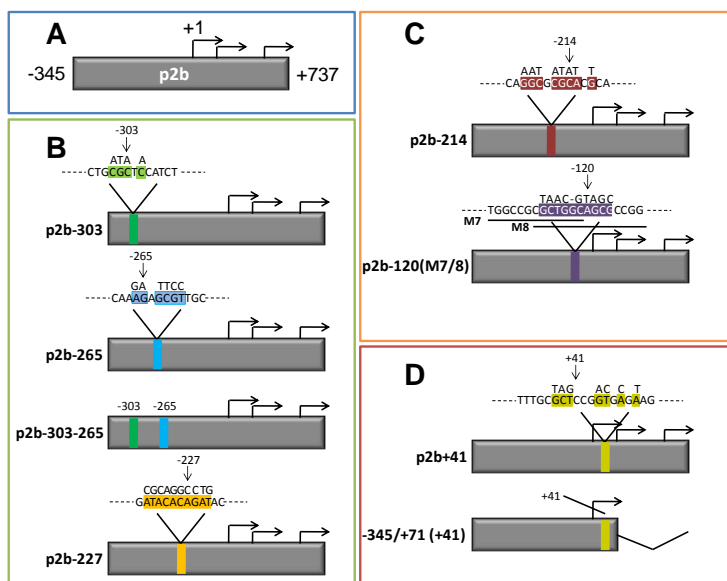


Figura D.11. Esquema de les diferents mutacions del promotor de *Trl* realitzades, agrupades segons el criteri que va portar a fer-les. S'indica la seqüència original, amb els parells de bases que es van mutar en colors i indicant a sobre la seqüència per la que es va substituir. Amb una fletxa s'indica la posició que dona nom a la construcció. **(A)** Promotor p2b original, indicant les posicions inicial i final. **(B)** Mutacions basades en els resultats dels *footprintings in vivo* amb DMS, a embrions. **(C)** Mutacions de dues seqüències molt similars, una d'elles contenint el motius del promotor basal M7 i M8. **(D)** Mutacions basades en els resultats obtinguts amb les delecions del promotor de *Trl*.

El segon criteri que vàrem seguir per a la mutació del promotor de *Trl* es va basar en la presència de dues seqüències molt semblants situades al voltant de -214 i -120 (veure figura D.11.C). D'aquestes seqüències, la que es troba a -120 està formada per dos motius consens M7 i M8 situats de forma solapada. D'aquests motius, descrits com a motius del promotor basal (Ohler et al. 2002), M7 s'ha vist que està present en un 15% de promotors on s'hi uneix TRF2, sent el segon motiu (darrera el motiu DRE) més freqüent en aquest tipus de promotors (Isogai et al. 2007a). Això resultava interessant perquè obria la possibilitat de que el promotor de *Trl* estigués format per diversos promotors, on s'hi unirien TBP o TRF2, i per tant podrien ser regulats de forma diferent i independent. De totes maneres, com es pot veure a la figura D.12.B, si bé aquestes seqüències poden estar implicades en l'activació de *Trl* (ja que tenen menys activitat inicial que p2b), no semblen determinants per la repressió donada per GAGA.

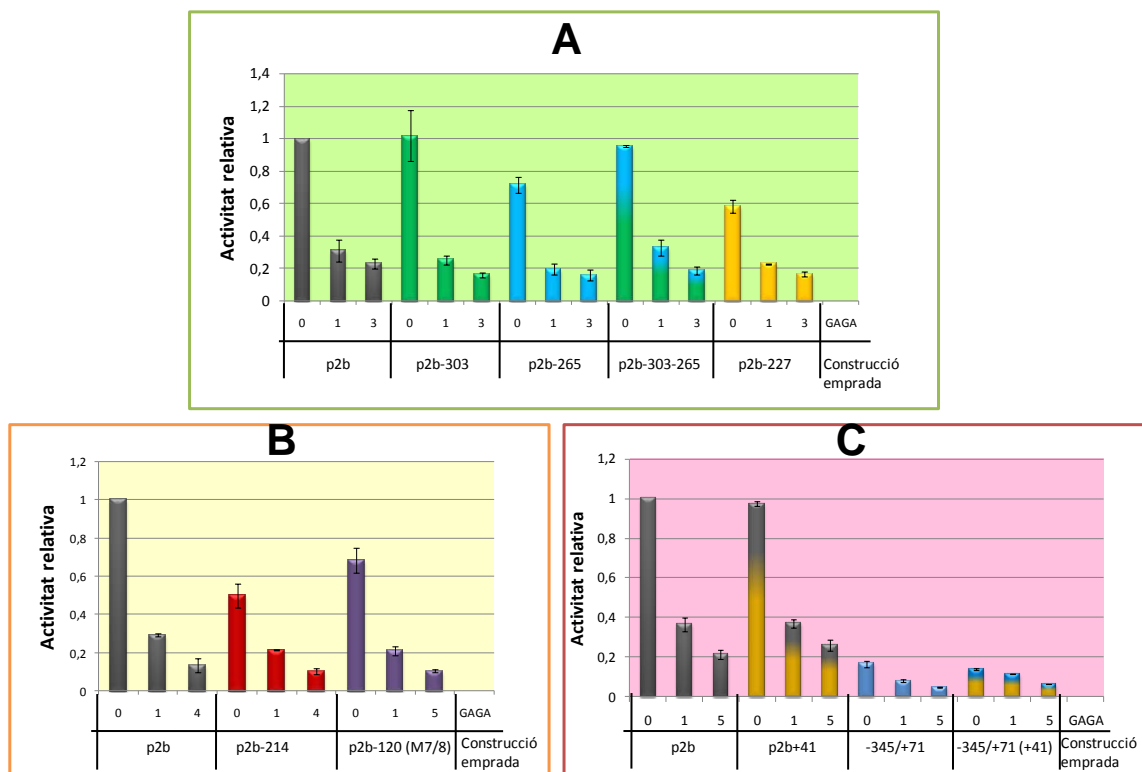


Figura D.12. Transfeccions transitòries a S2 amb les construccions corresponents de cada mutació del promotor de *Trl*. Cada valor és la mitjana de dos a sis experiments independents i representa el valor normalitzat de l'activitat luciferasa, relativitzat al valor de p2b en absència de sobreexpressió de GAGA. La quantitat de vector *reporter* amb cada promotor és sempre de 3 µg per transfecció. La quantitat de GAGA₅₁₉ cotransfectada s'indica en µg a sota de cada columna. **(A)** Resultats de les mutacions basades en els *footprintings in vivo* a embrions. **(B)** Resultats de les mutacions de dues seqüències molt similars, una d'elles contenint els motius del promotor basal M7 i M8. **(C)** Resultats de les mutacions basades en la informació obtinguda de les delecions del promotor de *Trl*.

Per últim, basant-nos en els resultats obtinguts amb les delecions del promotor (veure apartat anterior), vàrem realitzar una mutació més. Com ja hem vist, la construcció “-345/+71”

RESULTATS

continua sent reprimible per GAGA mentre que la construcció “-345/+14” no (veure figura D.10 de l'apartat anterior). L'única diferència entre aquestes dues construccions és la presència de la regió entre +14 i +71 a la construcció “-345/+71”. Anteriorment s'havia descrit, per *footprinting in vitro* amb DNAsa I, una àmplia regió d'unió per a GAGA en aquesta seqüència, entorn a +41 (veure figura D.5.A de l'apartat D.2.2.1 (Kosoy et al. 2002)). Així, vàrem mutar aquesta seqüència de forma que no contingués cap lloc potencial per a la unió de GAGA. Això ho vàrem fer tant al promotor p2b, com al “-345/+71” (veure figura D.11.D), per tal de determinar si aquest lloc és essencial o si només contribueix a la repressió. Com es pot veure a la figura D.12.C, sorprenentment aquesta seqüència sembla que no està implicada en el mecanisme de repressió, ja que fins i tot “-345/+71 (+41)” continua sent reprimible per GAGA.

D.3. Contribució de les diferents regions de GAGA a la repressió

A més del promotor de *Trl* i la seva seqüència, l'altre element que segur està implicat en la repressió de *Trl* per GAGA és, òbviament, la pròpia GAGA. L'estudi detallat de quina o quines regions de GAGA estan implicades en la repressió de *Trl* és de vital importància per tal d'entendre el mecanisme d'aquesta repressió. Així doncs, a banda de delectonar i mutar la regió promotora de *Trl*, és van fer tot un seguit de construccions amb formes truncades de la proteïna GAGA₅₁₉ per poder assajar, per transfecció transitòria a S2, la seva activitat sobre el promotor de *Trl* i així delimitar les regions de la proteïna implicades en el mecanisme de repressió de *Trl*. Estudis anteriors (Kosoy et al. 2002) ja havien definit que tant el domini POZ com el domini Q de GAGA són prescindibles per a la repressió de *Trl*. A més, una mutació puntual al DBD, que abolia la capacitat de GAGA per unir-se al DNA, també abolia completament la repressió, indicant que la unió de GAGA al promotor de *Trl* és imprescindible per a la repressió. Partint d'aquestes dades vàrem dissenyar diverses formes truncades de GAGA₅₁₉ per tal de discernir la implicació de la regió X en la repressió (veure figura D.13.A), així com definir si el domini d'unió al DNA, a més de necessari, també és suficient per a la repressió (veure figura D.13.B). Pel primer objectiu es van realitzar delectacions de la regió X tant a N-terminal (GAGA₅₁₉ΔX1) com a C-terminal (GAGA₅₁₉ΔX2) de la possible senyal de localització nuclear, així com una delectió de tot l'extrem N-terminal de GAGA, incloent l'NLS

(GAGA₅₁₉ΔPOZ₂₄₅). Pel segon objectiu, es va voler assajar una construcció consistent només en el DBD (GAGA-DBD), però l'aparent impossibilitat d'expressar-la correctament a cèl·lules S2 (veure més endavant) ens va portar a dissenyar diverses proteïnes de fusió més. La primera, GAGA-DBD-EGFP, és una fusió de la proteïna EGFP (*Enhanced Green Fluorescent Protein*) a C-terminal del DBD, construïda a fi de poder detectar més fàcilment la fusió. La segona, ECFP-Nuc-GAGA-DBD, és una fusió de la proteïna ECFP (*Enhanced Cyan Fluorescent Protein*), junt amb tres senyals de localització nuclear, a N-terminal del DBD. Amb aquesta construcció preteníem dirigir al nucli la proteïna de fusió, ja que l'anterior va demostrar ser tan nuclear com citoplasmàtica. Finalment, Lac I-GAGA-DBD, que és una fusió de la proteïna bacteriana Lac I a N-terminal del DBD, per a la qual teníem anticossos específics i a més sabem que es localitza estrictament al nucli, a S2.

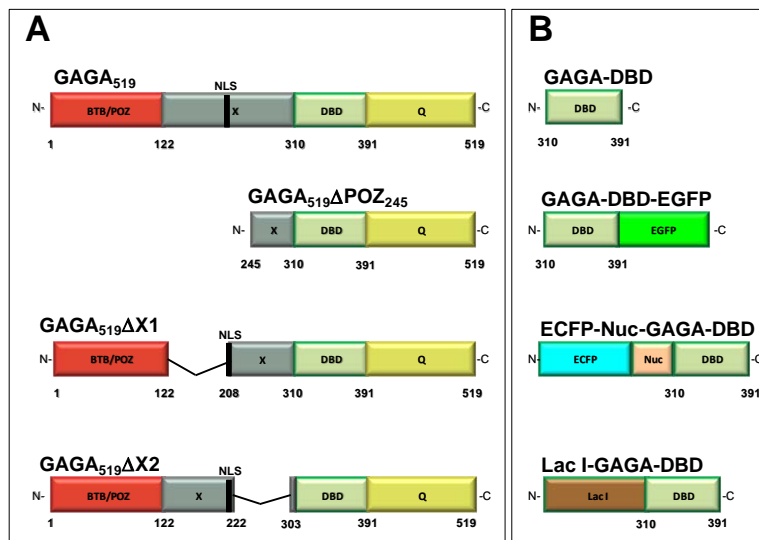


Figura D.13. Esquema resum de les diferents formes truncades de GAGA que es van realitzar i assajar. **(A)** A més de mostrar la GAGA₅₁₉, de la qual deriven totes les altres construccions, es mostren les diferents construccions realitzades a fi de determinar la implicació de la regió X en el mecanisme de repressió. **(B)** Esquema de les diferents construccions realitzades per tal de determinar el paper del DBD de GAGA en el mecanisme de repressió.

Els resultats de les transfeccions amb les formes truncades (veure figura D.14.A) indiquen que totes perden la capacitat de reprimir *Trl*, a excepció de GAGA₅₁₉ΔX2, que sembla tant competent com GAGA₅₁₉ en aquesta funció. A més, es van realitzar controls per tal de comprovar que cap de les proteïnes fusionades al DBD tingués activitat per sí mateixa, i així es pot comprovar per tots els casos, excepte per EGFP, que sembla tenir un lleuger efecte activador de *Trl*, tant per separat com fusionat al DBD. Tot i amb això, aquest efecte és molt lleu i podria ser part de l'error experimental. Addicionalment, les formes delecionades a la regió X van ser testades també per comprovar la seva capacitat d'activar el promotor d'*eve stripe 2* (veure figura D.14.B), que prèviament ja s'havia descrit que podia ser activat per GAGA (Kosoy et al. 2002; Pagans et al. 2002). Novament, només GAGA₅₁₉ΔX2 sembla competent en activació,

RESULTATS

tot i que no de forma tant efectiva com la proteïna GAGA₅₁₉. No obstant, aquesta menor activitat podria ser correlacionada amb una menor quantitat de proteïna, com mostra el control de *western blot* (veure figura D.14.C). És de destacar que, de totes aquestes construccions, l'expressió de GAGA-DBD i GAGA₅₁₉ΔX1 no va poder ser detectada ni per *western blot* ni per immunolocalització, indicant que aquestes formes podrien ser molt inestables i ràpidament degradades, fet que explicaria la seva nul·la activitat repressora. La resta de construccions que perden la capacitat de reprimir (GAGA₅₁₉ΔPOZ₂₄₅, GAGA-DBD-EGFP, ECFP-Nuc-GAGA-DBD i Lac I-GAGA-DBD) sí han estat detectades, però les immunolocalitzacions a cèl·lules S2 indiquen que la seva distribució és diferent a la de la GAGA normal (veure figura D.14.D). Així, GAGA-DBD-EGFP té una distribució similar a la d'EGFP per si sol, és a dir, tant nuclear com citoplasmàtica,

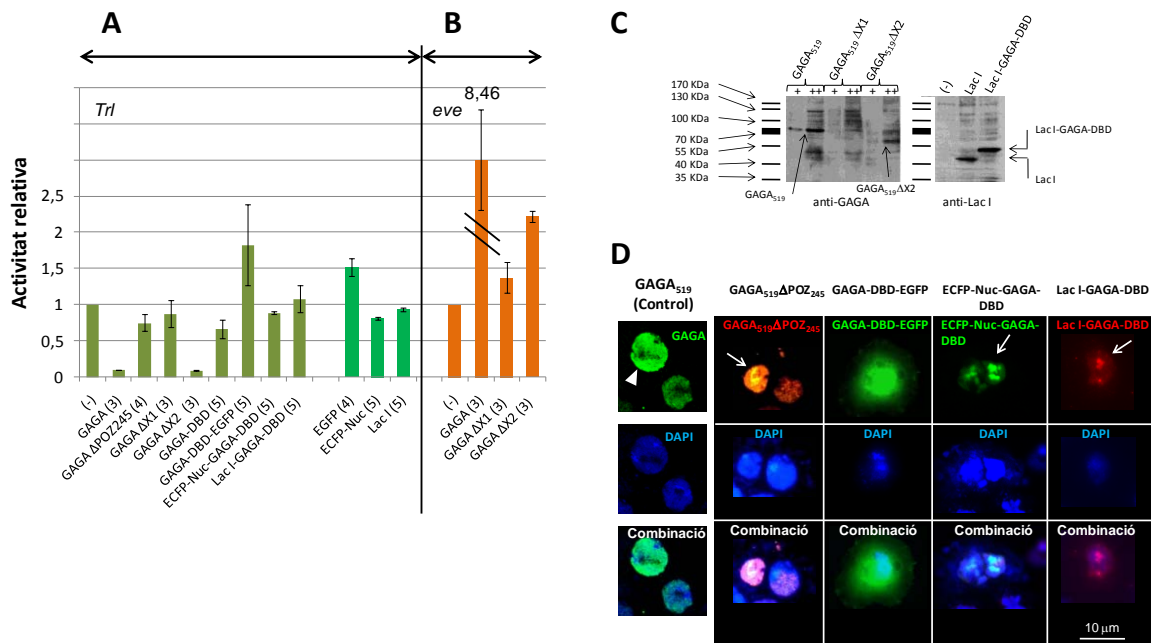


Figura D.14. Transfecció transitòria de les formes truncades de GAGA. El valor numèric que acompanya al nom de cada construcció indica els µg de vector d'expressió emprats. **(A)** Efecte sobre *Trl* (quantitat fixa de 3 µg) d'aquestes formes truncades de GAGA. Cada valor representa la mitjana dels valors luciferasa de dos a vuit experiments independents normalitzats i relativitzats al valor (-). S'inclouen els resultats control de les proteïnes EGFP, ECFP-Nuc i Lac I fusionades al DBD de GAGA. **(B)** Efecte de les delecions de la regió X sobre el promotor d'*eve stripe 2* (quantitat fixa de 7 µg). Cada valor representa la mitjana dels valors luciferasa normalitzats de quatre experiments independents i relativitzats al valor (-). **(C)** *Western blots* de les transfeccions amb algunes de les formes delecionades. A l'esquerra, *western blot* anti-GAGA on es carreguen quantitats creixents de l'extracte de cèl·lules transfectades per la construcció indicada. A la dreta, *western blot* anti-Lac I on només es carrega una quantitat de l'extracte de cèl·lules transfectades per la construcció indicada. **(D)** Immunolocalitzacions en cèl·lules S2 on es mostra el patró de localització de GAGA₅₁₉ΔPOZ₂₄₅, GAGA-DBD-EGFP, ECFP-Nuc-GAGA-DBD i Lac I-GAGA-DBD. Per a totes es varen utilitzar anticossos específics, a excepció de ECFP-Nuc-GAGA-DBD, la imatge de la qual mostra la seva pròpia emissió de fluorescència, tot i que s'ha canviat el color del blau real a verd, per evitar confusió amb el DAPI. La fletxa blanca indica l'estranya acumulació d'aquestes proteïnes, observada al nucli. A l'esquerra es mostra el patró que presenta la sobreexpressió de GAGA₅₁₉ (triangle blanc). Imatges preses amb un microscopi òptic de fluorescència Nikon Eclipse E-1000, a excepció de les d'ECFP-Nuc-GAGA-DBD i el control de GAGA₅₁₉, que s'han captat amb un microscopi confocal Leica TCS SPE.

mentre que GAGA té una distribució estrictament nuclear. Les altres tres construccions sí són nuclears, però presenten un estrany patró d'acumulació a certes zones del nucli (indicat amb una fletxa blanca a la figura), mai vist per a GAGA. Així doncs, totes aquestes formes truncades, tot i perdre l'activitat repressora, no es pot concloure que sigui degut a la pèrdua d'una regió important en aquest mecanisme, ja que podria ser un efecte de la seva defectuosa expressió i/o localització.

D.4. Contribució, a la repressió de *Trl*, de diferents factors de transcripció que interaccionen amb GAGA: TTK69 i Psq no participen

Com ja s'ha comentat a l'apartat D.2, al promotor de *Trl* ha d'existir una seqüència o seqüències que portin a GAGA a reprimir *Trl* i no a activar-lo. La nostra hipòtesi inicial era que algun factor d'unió a DNA, diferent de GAGA, estaria unint-se a aquestes seqüències i directa o indirectament interaccionant amb GAGA, fet que donaria com a resultat la repressió de *Trl*. Així doncs, dels factors de transcripció que es coneix que interaccionen amb GAGA, TTK69 i Psq semblaven els millors candidats, ja que ambdós factors disposen de múltiples llocs d'unió al promotor de *Trl*. De tota manera, s'ha de tenir en compte que la interacció de TTK69 i Psq amb GAGA s'ha descrit que es dona, almenys majoritàriament, pels dominis POZ d'aquestes proteïnes (Bardwell et al. 1994; Pagans et al. 2002; Schwendemann et al. 2002). Això contrasta amb que aquest domini sigui prescindible per a la funció repressora de GAGA.

TTK69 és un factor repressor de certs gens de desenvolupament com ara *ftz* i *eve* (Brown et al. 1993). Psq en canvi és una proteïna d'unió al DNA amb la mateixa especificitat de seqüència que GAGA i de fet s'ha vist que totes dues proteïnes colocalitzen perfectament a cromosomes politènics i a més poden actuar conjuntament, tant en l'activació com en el silenciament de multitud de gens homeòtics i d'altres tipus (Schwendemann et al. 2002).

Les estratègies que vàrem utilitzar per tal d'estudiar la implicació d'ambdues proteïnes van ser diferents. Mentre que per a TTK69 vàrem optar per la cotransfecció transitòria a cèl·lules S2 del vector d'expressió de TTK69 juntament amb el vector *reporter* del promotor de *Trl*, per Psq es va utilitzar un RNAi contra Psq per tal d'eliminar el component endògen de les

RESULTATS

cèl·lules i comprovar el seu efecte en l'activitat del promotor de *Trl*, transfectat transitòriament. L'estat inicial del promotor de *Trl* a S2 ja està parcialment reprimat per la GAGA endògena (Kosoy et al. 2002). Així, si Psq contribueix a la repressió, la seva eliminació per RNAi hauria de resultar en un augment d'activitat de *Trl*. La raó d'utilitzar l'RNAi contra Psq ve de la dificultat que va presentar per a nosaltres la sobreexpressió de Psq a cèl·lules S2.

Els resultats, tant de la sobreexpressió de TTK69 com de l'eliminació de Psq per RNAi es mostren a la figura D.15. Podem veure que tant TTK69 (figura D.15.A) com Psq (figura D.15.C) no tenen cap efecte en *Trl*, ni sobre l'activitat inicial ni sobre la repressió que exerceix GAGA sobre aquest.

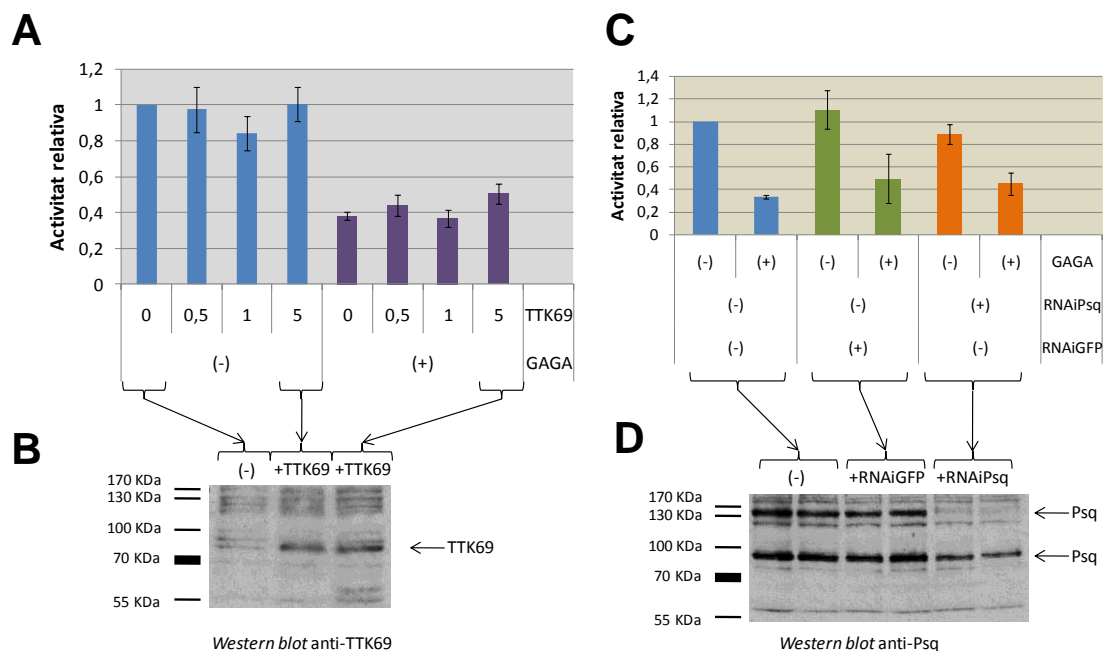


Figura D.15. TTK69 i Psq no participen en la repressió de *Trl* per GAGA. Tots els experiments estan fets transfectant transitòriament a cèl·lules S2 una quantitat fixa (3 μg) del promotor de *Trl* fusionat al gen *reporter* luciferasa. **(A)** Efecte de la sobreexpressió de TTK69 sobre l'activitat de *Trl* en presència (+) o absència (-) de 0,5 μg de vector d'expressió de GAGA₅₁₉. S'indica la quantitat de vector d'expressió de TTK69 emprada, en μg. **(B)** Comprovació per *western blot* emprant anti-TTK69 de la correcta sobreexpressió de TTK69. Les fletxes indiquen a quina transfecció correspon la mostra carregada a cada carril. **(C)** Efecte de l'addició de 100 μg per placa (directament al medi i mantenint-lo quatre dies) d'RNA d'interferència contra Psq (RNAiPsq) sobre l'activitat de *Trl*, tant en combinació amb la sobreexpressió de GAGA₅₁₉ (0,5 μg de vector d'expressió) com sense. Com a control es va afegir la mateixa quantitat d'RNAi contra GFP (RNAiGFP), que no dona cap efecte. **(D)** Control per *western blot* emprant anti-Psq de la reducció de Psq a les cèl·lules. Les fletxes indiquen a quina transfecció correspon la mostra carregada a cada carril. L'RNAiPsq elimina eficientment les diverses isoformes de Psq descrites (font: <http://flybase.org/>).

D.5. Contribució de diferents corepressors a la repressió de *Trl*

A banda de la proteïna GAGA i possiblement d'un factor d'unió específica al promotor de *Trl*, per a la funció repressora de GAGA podria ser que es necessités un corepressor específic, portat al promotor de *Trl* per GAGA, o per alguna altra proteïna implicada en aquesta repressió, que fos el responsable d'interaccionar amb la maquinària basal de transcripció. Els corepressors (i també els coactivadors) no són proteïnes d'unió específica al DNA, així doncs, la identificació d'un possible corepressor implicat en la repressió de *Trl* no aclariria el fet de per què aquest promotor és reprimat i no activat per GAGA, però sí donaria informació sobre el mecanisme de repressió.

Dels corepressors coneguts de *Drosophila*, dCtBP i Groucho destaquen per la seva importància en dos mecanismes de repressió àmpliament utilitzats: el mecanisme de repressió de llarg abast (*long-range*), mediat freqüentment per Groucho, i el de curt abast (*short-range*), mediat en molts casos per dCtBP (Courey et al. 2001). Així doncs, vàrem voler comprovar la implicació d'aquests dos factors en el mecanisme de repressió de *Trl* per GAGA.

D.5.1. Groucho no sembla participar en la repressió de *Trl*

Per determinar la implicació de Groucho es varen dur a terme assajos de transfecció transitòria a S2 sobreexpressant Groucho (amb una cua d'histidines a N-terminal) i comprovant el seu efecte sobre l'activitat del promotor de *Trl* (veure figura D.16.A). La raó de l'addició de la cua N-terminal d'histidines era facilitar la seva detecció per mitjà d'anticossos específics contra aquesta cua, ja que no disposem d'anticossos contra Groucho. Els resultats indiquen que Groucho no té cap efecte en aquest sistema. Tot i amb això, no vàrem poder comprovar la correcta sobreexpressió de Groucho degut a que l'anticòs contra les histidines a N-terminal (*anti-HIS Penta* de ratolí, produït per Qiagen) no ens va permetre detectar la sobreexpressió de Groucho. De totes maneres, hi ha indicis de que Groucho sí podria estar sobreexpressant-se a cèl·lules S2, ja que en un control d'expressió a bacteris sí que aconseguim veure una banda que podria correspondre a Groucho (veure figura D.16.B).

RESULTATS

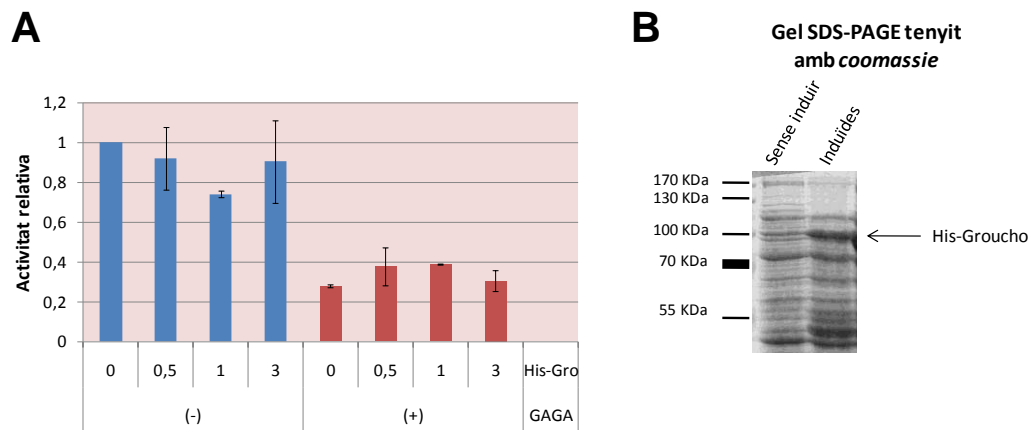


Figura D.16. Groucho sembla no estar implicat en el mecanisme de repressió de *Trl* per GAGA. **(A)** Efecte de la sobreexpressió de Groucho (His-Gro, Groucho amb cua d'histidines a N-terminal) sobre l'activitat de *Trl*, tant basal (-) com en estat reprimat per la sobreexpressió de GAGA (+) (0,3 µg del vector d'expressió per transfecció). S'indiquen els µg de His-Gro cotransfectats en cada cas. Els valors corresponen a les mitjanes de les activitats luciferasa normalitzades de quatre experiments independents, i relativitzades al primer valor. **(B)** Gel SDS-PAGE tenyit amb *coomassie* on es mostra l'expressió de His-Gro a la soca bacteriana BL21(DE3). La soca va ser transformada amb el plasmidi d'expressió (induïble per IPTG) pET28a-HisGro.

D.5.2. Possible implicació de dCtBP en la repressió de *Trl*

Per determinar la implicació de dCtBP es varen dur a terme experiments de transfecció transitòria a cèl·lules S2 sobreexpressant dCtBP, fusionat a l'epítip Flag a C-terminal per facilitar la seva detecció, i comprovant el seu efecte sobre l'activitat del promotor de *Trl* (veure figura D.17.A). Tot i que els experiments mostrats corresponen a la forma llarga de dCtBP (dCtBP-L) vàrem confirmar que la forma curta (dCtBP-S) té la mateixa activitat (dades no presentades). Ambdues isoformes semblen tenir també la mateixa activitat en altres contextos estudiats (Poortinga et al. 1998). Els resultats indiquen que dCtBP és capaç de reprimir *Trl*, tant en presència com en absència de sobreexpressió de GAGA. Això indicaria que podria estar implicat en el mecanisme de repressió de *Trl* per GAGA, tot i que també podria estar reprimint *Trl* per alguna altra via.

Per delimitar si la funció repressora de dCtBP està relacionada amb GAGA vàrem realitzar tota una sèrie d'experiments. Primer, vàrem voler comprovar l'efecte de la sobreexpressió de dCtBP a cèl·lules HeLa. Com ja s'ha comentat a l'apartat D.1.2, GAGA activa *Trl* a HeLa. Així, si dCtBP és el factor que diferencia les cèl·lules HeLa de les S2, la seva sobreexpressió (a HeLa) portaria GAGA a reprimir *Trl*. De totes maneres, a humans també existeix CtBP i l'homologia de seqüència amb la forma de *Drosophila* és molt gran (Poortinga et al. 1998), així que la possibilitat de que aquest sigui el factor diferencial entre espècies és

escassa. Els nostres resultats corroboren això (veure figura D.17.B), ja que la sobreexpressió de dCtBP a HeLa (aquest cop la forma curta, per simplicitat de construcció del plasmidi d'expressió) no té cap efecte apreciable sobre l'activitat de *Trl*. De tota manera no es pot descartar que dCtBP actuï juntament amb GAGA a S2 per reprimir *Trl*, ja que pot ser que a HeLa manquin altres factors determinants per aquesta funció.

Seguidament vàrem comprovar l'efecte que tenia l'addició al medi de TSA i Nicotinamida sobre la funció repressora de dCtBP. El TSA (Tricostatina-A) és un inhibidor de molts HDACs (complexos desacetilases d'histones) responsables de l'acció repressora de molts corepressors. Tot i que dCtBP no s'ha descrit que funcioni a través de HDACs, vàrem voler comprovar si altres corepressors sensibles a TSA podrien estar actuant en la repressió de *Trl*. Per altre banda, l'activitat corepressora de dCtBP podria estar relacionada amb el metabolisme del NAD⁺ (nicotin-adenosin dinucleòtid), ja que s'ha vist que aquesta proteïna, així com els seus homòlegs a humans i *C. elegans*, contenen regions d'alta homologia amb la família de les NAD⁺-dependents 2-hidroxiàcid deshidrogenases isomer-D específiques. La significació d'aquesta homologia no està clara ja que, al menys la forma humana de CtBP, no presenta ni activitat hidrogenasa ni uneix NAD⁺, suggerint-se que aquestes regions tindrien funcions estructurals, com ara la dimerització (Schaeper et al. 1995). De totes maneres també vàrem provar si la Nicotinamida, inhibidora del metabolisme del NAD⁺, inhibia també la funció de dCtBP. Com es pot veure als resultats (figura D.17.C), l'addició de 5mM de Nicotinamida al medi no té cap efecte en el nostre sistema, indicant que el metabolisme del NAD⁺ no té cap relació amb l'activitat de dCtBP en aquest context, ni amb el mecanisme de repressió de *Trl* per GAGA. Per contra, l'addició de 300nM de TSA al medi si sembla inhibir en part l'acció repressora de dCtBP per sí sol (punts "sense GAGA" a la figura). Quan es cotransfecta GAGA, TSA continua inhibint l'acció de dCtBP, tot i que de forma no tan clara, degut probablement a que l'efecte de GAGA per sí sol és molt més important que el de dCtBP, emmascarant així l'efecte del TSA. El que sí sembla clar és que, si bé el TSA inhibeix en part l'acció repressora donada per dCtBP, ni el TSA ni la Nicotinamida inhibeixen de cap manera la repressió donada únicament per la sobreexpressió de GAGA (punt "amb GAGA" i 0 µg de dCtBP a la figura D.17.C). Això apuntaria a que la repressió de *Trl* per GAGA no involucra dCtBP. De tota manera, cal destacar que l'activitat repressora de dCtBP no s'havia vist anteriorment que s'afectés per TSA, així que aquest podria ser un efecte indirecte.

RESULTATS

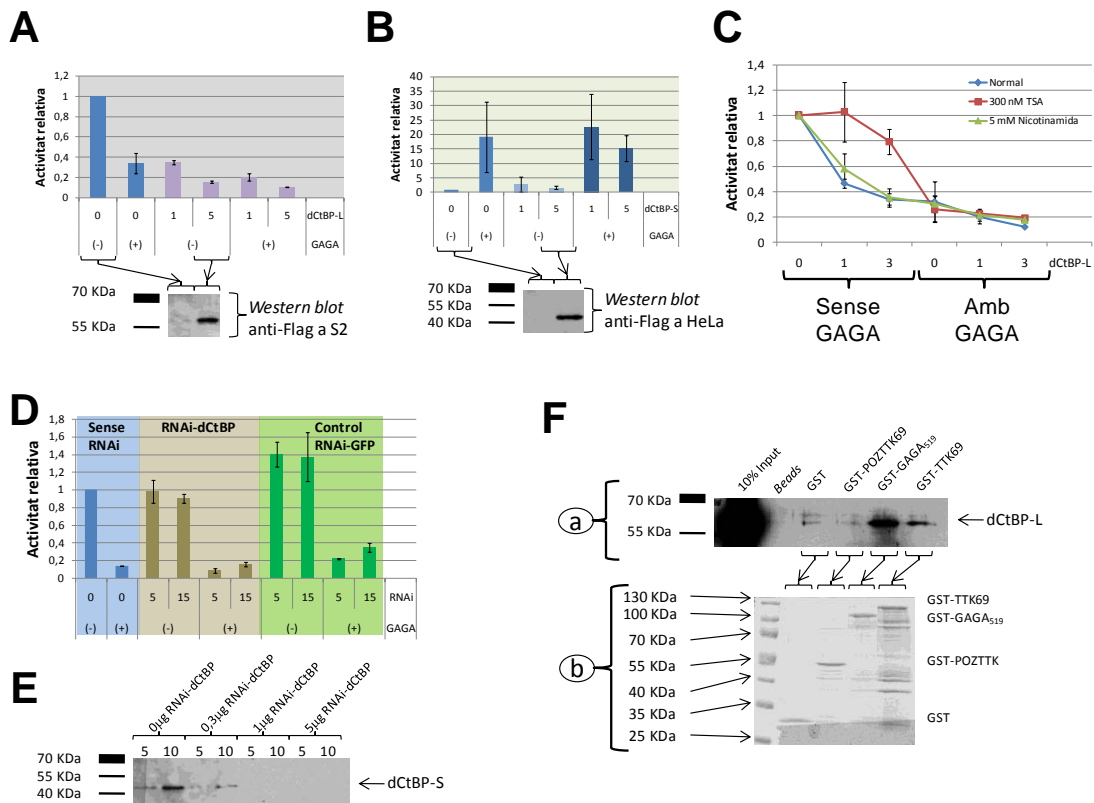


Figura D.17. Estudi de la contribució de dCtBP a l'activitat del promotor de *Trl*, així com a la seva repressió per GAGA. A tots els experiments (A, B, C i D) es transfecta sempre una quantitat fixa de 3 μ g del vector *reporter* amb el promotor de *Trl*. **(A)** Efecte a S2 de la sobreexpressió de dCtBP-L sobre l'activitat del promotor de *Trl*, en presència (+) i en absència (-) de sobreexpressió de GAGA (0,4 μ g de vector d'expressió). Els valors són les mitjanes normalitzades de quatre experiments independents, i relativitzades a l'estat inicial de *Trl* (primer valor). S'indiquen els μ g emprats del vector d'expressió de dCtBP-L. A sota, control per *western blot* anti-Flag de la correcta sobreexpressió de dCtBP-L. Les fletxes indiquen a quina mostra correspon cada carril. **(B)** Experiment similar a l'anterior però dut a terme a cèl·lules HeLa i sobreexpressant la forma curta de dCtBP (dCtBP-S). Als punts indicats amb (+) s'han cotransfectat 3 μ g de vector d'expressió de GAGA. Els valors són les mitjanes normalitzades de vuit experiments independents, i relativitzades a l'estat inicial de *Trl* (primer valor). A sota, control per *western blot* anti-Flag de la correcta sobreexpressió de dCtBP-L. Les fletxes indiquen a quina mostra correspon cada carril. **(C)** Efecte del TSA i la Nicotinamida, afegides al medi de cultiu, sobre la repressió exercida per dCtBP-L (s'indica a les abscisses la quantitat transfectada en μ g), tant en presència (amb GAGA) com en absència (sense GAGA) de 0,5 μ g del vector d'expressió de GAGA cotransfectat. Els valors són les mitjanes normalitzades de sis experiments independents, i relativitzades a l'estat inicial de *Trl* (primer valor). **(D)** Efecte de l'RNAi contra dCtBP, així com el d'un control d'RNAi contra GFP, afegits al medi durant tres dies, sobre l'activitat del promotor de *Trl* i la repressió d'aquest per 1 μ g de vector d'expressió de GAGA (+). Els valors són les mitjanes normalitzades de dos experiments independents, i relativitzades a l'estat inicial de *Trl* (primer valor). A les abscisses s'indica la quantitat d'RNAi afegit al medi i si s'ha cotransfectat (+) o no (-) vector d'expressió de GAGA. **(E)** Comprovació indirecta per *western blot* anti-Flag de l'efectivitat de l'RNAi contra dCtBP. Totes les mostres provenen de la transfecció de 3 μ g de dCtBP-S amb la quantitat d'RNAi-dCtBP indicada, afegida durant tres dies al medi. Es van carregar dues quantitats de cada mostra (5 i 10 μ g). **(F)** (a) Resultats dels GST-Pulldowns revelats per *western blot* utilitzant l'anticòs anti-Flag. A baix (b), gel SDS-PAGE tenyit amb *coomassie* fet com a control de càrrega, on s'ha carregat la mateixa quantitat de cada fusió de GST, que l'emprada en l'assaig de Pulldown.

Posteriorment vàrem comprovar si l'eliminació del component endògen de dCtBP a S2, per mitjà d'RNAi contra dCtBP, tenia algun efecte (veure figura D.17.D). Com es pot veure, l'addició d'RNAi contra dCtBP al medi no té cap efecte en l'activitat de *Trl*. Així doncs, això va en contra d'una acció conjunta dCtBP-GAGA en la repressió de *Trl*. S'ha d'indicar però, que l'eliminació del component endògen de dCtBP no va poder ser comprovada directament, ja que

no disposem d'anticossos contra dCtBP. No obstant, si es va poder comprovar indirectament, ja que les formes que sobreexpressen de dCtBP porten un epítot Flag a C-terminal (detectable amb anticossos anti-Flag), fet que va fer possible veure com l'RNAi emprat contra dCtBP era efectiu eliminant aquest dCtBP sobreexpressat (veure figura D.17.E).

Per últim, vàrem comprovar, per assajos de *GST-Pulldown*, com dCtBP és capaç d'interaccionar amb la proteïna de fusió GST-GAGA (veure figura D.17.F.a). Això aporta més evidències a favor d'una possible acció conjunta GAGA-dCtBP en la repressió de *Trl*. Com a control es pot comprovar com interacciona amb la proteïna TTK69 sencera, però no pel domini POZ d'aquesta. Això entra en contradicció amb experiments anteriors, on s'observa que TTK69 interacciona pel seu domini POZ amb dCtBP, tot i que a la part C-terminal de TTK69 hi ha seqüències consens per a la unió de dCtBP (Wen et al. 2000). A més, també es pot comprovar que no interacciona amb GST, ni amb la matriu de *Glutathione Sepharose 4B (beads)*.

D.6. Efectes de GAGA a mosques transgèniques

Amb l'objectiu de comprovar que la repressió de *Trl* per GAGA, observada a cèl·lules S2, també es dona a la mosca, vàrem dissenyar i produir diverses línies transgèniques de *Drosophila melanogaster*. Per una banda, vàrem generar mosques transgèniques que incorporen una còpia ectòpica del promotor de *Trl*, en la seva versió més llarga, fusionat al cDNA de GFP (veure figura D.18.A). Per altra banda, vàrem generar mosques transgèniques que incorporen el cDNA de GAGA₅₁₉, GAGA₅₈₁, GAGAΔQ o GAGA₅₁₉DBDKO, sota el control d'un promotor induïble que conté una UAS (*upstream activating sequence*) que presenta 5 motius d'unió per a la proteïna de llevat Gal4. A més, al laboratori disposem de línies transgèniques que sobreexpressen aquesta proteïna de llevat (Gal 4) sota el control de diversos promotors, amb patró d'expressió coneguda. Aquests promotors que expressen Gal 4 els anomenem *drivers* i gràcies a ells s'aconsegueix dirigir la sobreexpressió del transgèn d'interès (en aquest cas GAGA₅₁₉, GAGA₅₈₁, GAGAΔQ o GAGA₅₁₉DBDKO) a teixits específics, seguint un patró espacial i temporal determinat i identificable. De forma similar, vam generar mosques transgèniques que incorporen una seqüència de la regió codificant de GAGA duplicada i invertida sota el control del mateix promotor induïble que conté la UAS que, en ser activat, expressa un RNAi de doble cadena que pot silenciar l'expressió de GAGA (*knockdown*).

RESULTATS

Així doncs, mitjançant els creuaments adients d'aquestes mosques transgèniques, vàrem obtenir els individus necessaris per a poder determinar, per assajos de coimmunolocalització amb anticossos fluorescents, l'efecte sobre *Trl* de la sobreexpressió de les proteïnes GAGA₅₁₉, GAGA₅₈₁, GAGAΔQ i GAGA₅₁₉DBDKO, així com del seu silenciament (*knockdown*) per RNAi. Addicionalment, vam caracteritzar els efectes fenotípics, tant de la sobreexpressió de les proteïnes GAGA com del seu *knockdown*, sota el control de diversos *drivers*.

D.6.1. GAGA₅₁₉, GAGA₅₈₁ i GAGAΔQ reprimeixen *Trl*.

GAGA₅₁₉BDKO no

Com a part d'aquest ampli estudi dels efectes de GAGA sobre el promotor de *Trl* a mosca, el nostre grup va començar per mostrar que l'expressió de la construcció del promotor de *Trl* transgènic (*Trl* full-GFP) era ubiqua i que es podia detectar a tots els teixits de la mosca amb una intensitat semblant, de forma molt similar a com es detecta la proteïna GAGA endògena (no es mostra). Seguidament, mitjançant coimmunolocalitzacions amb anticossos específics fluorescents, es va mostrar la correcta sobreexpressió de la isoforma GAGA₅₁₉, així com l'efecte de l'expressió de l'RNAi contra GAGA (que elimina les dues isoformes indistintament) a les diferents mosques transgèniques. Aquests dos experiments, tot i haver estat realitzats pel Dr. Jordi Bernués i no formar part del treball d'aquesta Tesi Doctoral, s'han inclòs a la memòria per facilitar l'anàlisi de tots els resultats de forma global. Utilitzant diversos *drivers* aconseguim dirigir la sobreexpressió dels transgens a zones determinades, que apareixen tenyides d'un vermell més intens a les imatges, front un vermell més fluix, que correspon a la detecció de la GAGA endògena. Així, com es pot veure a la figura D.18, GAGA₅₁₉ reprimeix *Trl*, tant a larves (discos imaginals) com a embrions. A més, també s'ha pogut comprovar que el promotor de *Trl* a larves (figura D.18.E), en absència de sobreexpressió de GAGA, ja es troba parcialment reprimat pel component endogen de GAGA, com demostra el fet que la seva eliminació per mitjà d'RNAi faci augmentar el nivell d'expressió del transgèn *Trl* full-GFP.

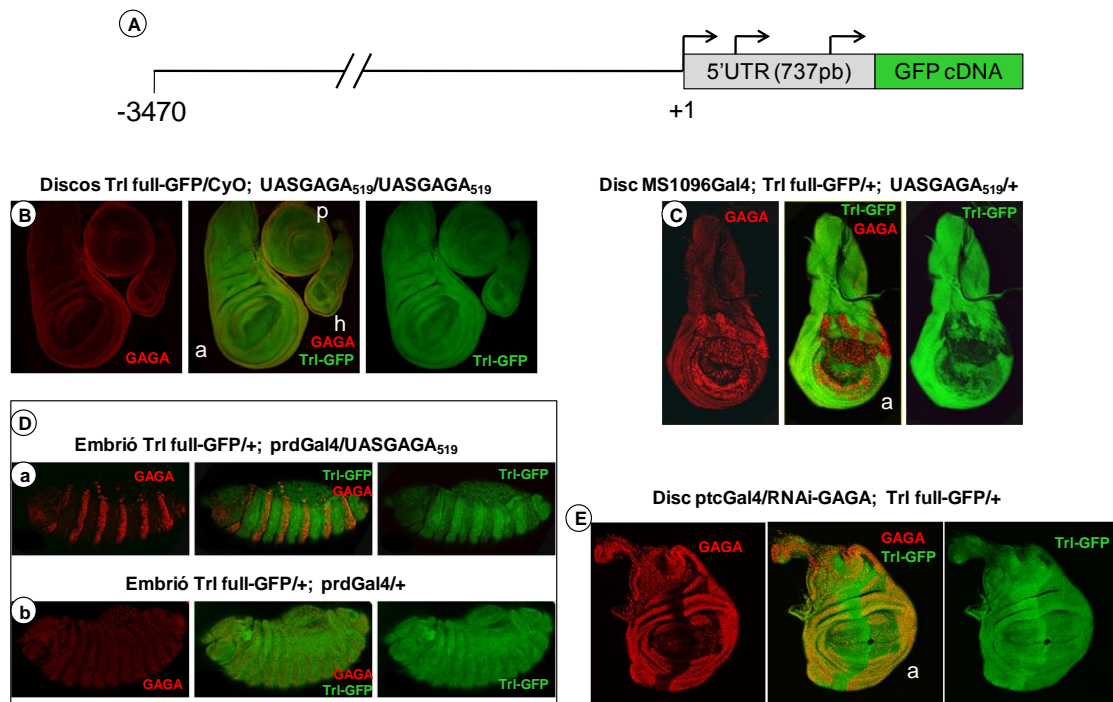


Figura D.18. $GAGA_{519}$ reprimeix *Trl* tant a embrió com a discos imaginals de larva. **(A)** Diagrama del promotor *Trl* full-GFP emprat per a generar mosques transgèniques. **(B, C i E)** Imatges obtingudes amb un microscopi confocal Leica TCS SP5, de discos imaginals de larva en 3r estadi tenyits amb anticossos anti-GAGA (vermell) i anti-GFP (verd). **(B)** Discos imaginals control d'ala (a), halteri (h) i pota (p), portant el transgèn *Trl* full-GFP, però no sobreexpressant $GAGA_{519}$, al no portar cap *driver*. **(C)** Disc imaginal d'ala que porta el transgèn *Trl* full-GFP i que sobreexpressa $GAGA_{519}$ sota el control del *driver* MS1096Gal4. **(D)** Imatges obtingudes amb un microscopi confocal Leica TCS SP5 d'embrions tenyits amb anticossos anti-GAGA (vermell) i anti-GFP (verd). (D.a) Repressió de *Trl* full-GFP per la sobreexpressió de $GAGA_{519}$ sota el control del *driver* *prdGal4*. (D.b) Embrions control portant els transgens *Trl* full-GFP i *prdGal4*, però que no sobreexpressen $GAGA_{519}$. **(E)** Disc imaginal d'ala que, portant el transgèn *Trl* full-GFP, sobreexpressa l'RNAi contra GAGA, sota el control del *driver* *ptcGal4*. Totes les mostres s'han obtingut d'individus creixent a 18°C, a excepció de (E), on els individus creixien a 29°C. Resultats, ja publicats, obtinguts al nostre laboratori pel Dr. Jordi Bernués (Bernues et al. 2007).

Seguidament, ja formant part del treball d'aquesta Tesi Doctoral, ens vàrem proposar veure l'efecte de la sobreexpressió de $GAGA\Delta Q$ sobre *Trl* full-GFP, tant a discos imaginals de larves en estadi 3 (veure figura D.19.A) com a embrions (veure figura D.19.C). Prèviament ja s'havia caracteritzat que aquesta forma truncada de la proteïna GAGA era perfectament capaç de reprimir *Trl*, en assajos de transfecció transitòria a cèl·lules S2 (Kosoy et al. 2002). Això mateix s'observa a embrions i discos imaginals de mosques transgèniques, com es pot comprovar a la figura D.19. A les zones on s'expressa el transgèn $UASGAGA\Delta Q$ s'observa una forta davallada del senyal de GFP, produït per l'expressió de *Trl* full-GFP, indicant que $GAGA\Delta Q$ és capaç de reprimir *Trl* també a embrions i a discos imaginals.

RESULTATS

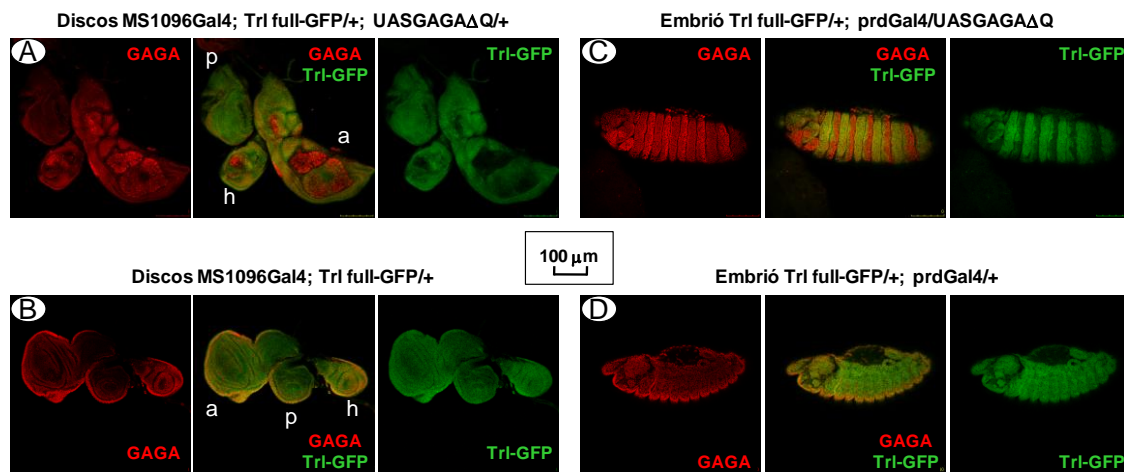


Figura D.19. GAGA Δ Q reprimeix *Trl* tant a discos imaginals de larva com a embrió. Imatges obtingudes amb un microscopi confocal Leica TCS SPE de discos imaginals de larva en 3r estadi (**A i B**) i d'embrions (**C i D**) tenyits amb anticossos anti-GAGA (vermell) i anti-GFP (verd). (**A**) Discos imaginals d'ala (a), halteri (h) i pota (p) que porten el transgèn *Trl* full-GFP i que sobreexpressen GAGA Δ Q sota el control del *driver* MS1096Gal4 (expressió a disc d'ala i halteri). (**B**) Discos imaginals control d'ala (a), halteri (h) i pota (p), portant els transgens *Trl* full-GFP i MS1096Gal4, però que no sobreexpressen cap forma de GAGA. (**C**) Repressió de *Trl* full-GFP per la sobreexpressió de GAGA Δ Q sota el control del *driver* prdGal4. (**D**) Embrions control portant els transgens *Trl* full-GFP i prdGal4, però que no sobreexpressen cap forma de GAGA. Totes les mostres s'han obtingut d'individus creixent a 18°C.

Finalment, també vàrem comprovar l'efecte de la sobreexpressió de GAGA₅₁₉DBDKO i la isoforma GAGA₅₈₁, sobre l'activitat del promotor *Trl* full-GFP. GAGA₅₁₉DBDKO és idèntica a GAGA₅₁₉, excepte per una mutació a l'aminoàcid 361, que passa d'histidina a prolina. Aquest aminoàcid és un dels encarregats de coordinar el Zn²⁺ al dit de Zinc de GAGA. Així, aquesta mutació desmunta el dit de Zinc, el que aboleix la unió de GAGA al DNA. Prèviament, per transfecció transitòria a cèl·lules S2, s'havia descrit que GAGA₅₁₉DBDKO era incapaç de reprimir *Trl*, mentre que la isoforma GAGA₅₈₁ era perfectament capaç de reprimir (Kosoy et al. 2002). Això mateix observem a discos imaginals de larva en estadi 3 (veure figura D.20). Amb cap dels *drivers* emprats (MS1096Gal4 i ptcGal4) la sobreexpressió del transgèn UASGAGA₅₁₉DBDKO provoca cap efecte en el nivell d'expressió del transgèn *Trl* full-GFP (veure figures D.20.A i D.20.B). En canvi, la sobreexpressió del transgèn UASGAGA₅₈₁, dirigida pels mateixos *drivers* que el cas anterior, dóna com a resultat, a tots els casos, una davallada en el nivell d'expressió de *Trl* full-GFP, indicant que reprimeix *Trl* (veure figures D.20.C i D.20.D).

Amb tot, confirmem que la repressió de *Trl* per GAGA es dóna també a mosques transgèniques i és independent de teixit perquè es dóna a tots els teixits estudiats. A més, al igual que succeeix a cèl·lules S2, totes dues isoformes de GAGA (GAGA₅₁₉ i GAGA₅₈₁) són capaces de reprimir, el domini Q és prescindible per a la repressió i aquesta requereix la unió al DNA de GAGA.

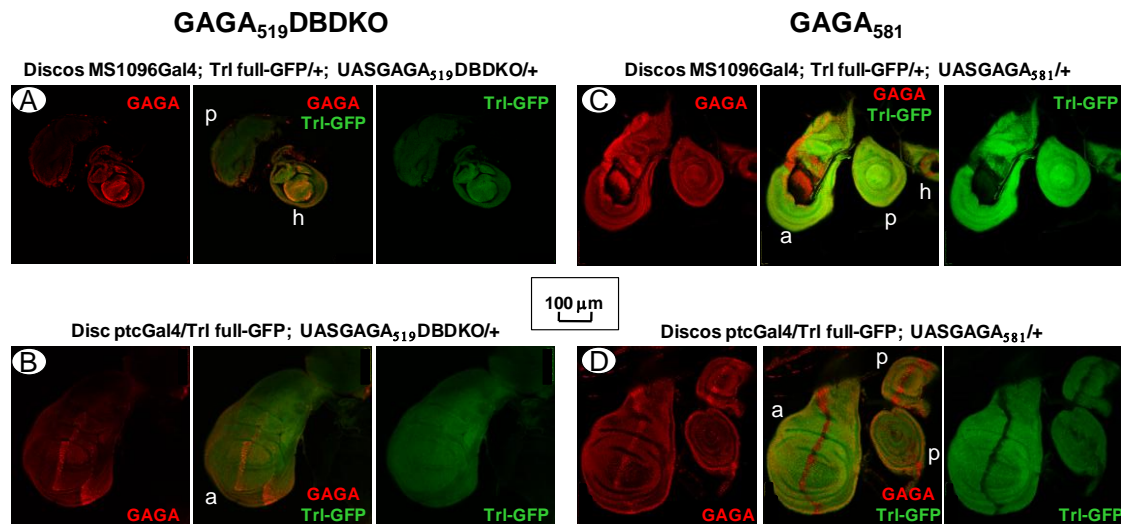


Figura D.20: $GAGA_{519}DBDKO$ no reprimeix *Trl* mentre que $GAGA_{581}$ sí. Imatges obtingudes amb un microscopi confocal Leica TCS SP5 de discos imaginals de larva en 3r estadi tenyits amb anticossos anti-GAGA (vermell) i anti-GFP (verd). **(A i B)** Discos imaginals d'ala (a), halteri (h) i pota (p) que porten el transgèn *Trl* full-GFP i que sobreexpressen $GAGA_{519}DBDKO$ sota el control dels *drivers* MS1096Gal4 (A) i *ptcGal4* (B). **(C i D)** Discos imaginals d'ala (a), halteri (h) i pota (p) que porten el transgèn *Trl* full-GFP i que sobreexpressen $GAGA_{581}$ sota el control dels *drivers* MS1096Gal4 (C) i *ptcGal4* (D). Totes les mostres s'han obtingut d'individus creixent a 18°C.

D.6.2. La sobreexpressió de $GAGA_{581}$ a mosques transgèniques produeix efectes fenotípics similars als de $GAGA_{519}$, però menys severos

Englobats dins d'aquest ampli estudi dels efectes de GAGA a mosques transgèniques, també vàrem analitzar els efectes fenotípics de la sobreexpressió de totes les construccions utilitzades a l'apartat anterior: $GAGA_{519}$, RNAi-GAGA, $GAGA\Delta Q$, $GAGA_{519}DBDKO$ i $GAGA_{581}$. De totes, l'única que forma part del treball d'aquesta Tesi Doctoral és $GAGA_{581}$. Els resultats de les altres construccions, ja publicats, han estat obtinguts pel Dr. Jordi Bernués (Bernues et al. 2007).

A la taula D.1 es resumeix la letalitat provocada per la sobreexpressió de cada isoforma de GAGA amb els diferents *drivers*. Com es pot veure, la sobreexpressió de $GAGA_{519}$ amb qualsevol dels *drivers* emprats comporta una alta letalitat. En canvi, la sobreexpressió de $GAGA_{581}$ amb qualsevol dels *drivers* provats és molt més viable, a excepció del *driver* que expressa Gal4 sota el control del promotor d'actina (*ActGal4*), que presenta una letalitat similar per ambdues isoformes. S'ha de destacar que tots aquests fenotips de letalitat han estat

RESULTATS

observats a, com a mínim, dues línies de mosques transgèniques independents per a cada isoforma, per descartar possibles efectes de posició del transgèn.

Taula D.1. Letalitat observada per la sobreexpressió de GAGA ₅₁₉ i GAGA ₅₈₁ a mosques transgèniques.		
<i>driver</i> emprat per a la sobreexpressió	Construcció sobreexpressada	Letalitat
ptcGal4	UASGAGA ₅₁₉	Moren a larva el 100% d'individus, tant a 18°C com a 25°C.
	UASGAGA ₅₈₁	Alta letalitat, sobretot a 25°C, on la majoria moren a fase pupa. Alguns aconseguen arribar a adult.
MS1096Gal4	UASGAGA ₅₁₉	Moren a fase pupa, alguns aconseguen arribar a adult a 18°C.
	UASGAGA ₅₈₁	No hi ha letalitat ni a 18°C, ni a 25°C.
dppGal4	UASGAGA ₅₁₉	Moren just abans d'eclosionar la pupa. Uns pocs aconseguen arribar a adult, tant a 18°C com a 25°C.
	UASGAGA ₅₈₁	Viabilitat normal a 18°C. A 25°C uns pocs moren a fase pupa.
ActGal4	UASGAGA ₅₁₉	Letal embrionari el 100 % dels casos.
	UASGAGA ₅₈₁	Letal embrionari el 100 % dels casos.

A més, amb els individus que arribaven a adult dels creuaments anteriors, vàrem procedir a un estudi comparatiu a nivell fenotípic, parant atenció sobretot als fenotips d'ales (veure figura D.21), ja que tots els *drivers* analitzats s'expressen als discos imaginals que donaran lloc a aquest teixit. Així, mentre la sobreexpressió de GAGA₅₁₉ sota el control de ptcGal4 és letal en el 100% dels casos, tant a 18°C com a 25°C (el que s'indica amb RIP a la figura D.21.A), la sobreexpressió de GAGA₅₈₁ amb el mateix *driver* només és clarament perniciosa per a la mosca a 25°C. Tant els pocs adults obtinguts a 25°C com els obtinguts a 18°C, sempre presentaven únicament una severa malformació a la vena creuada anterior (indicat amb un triangle negre a la figura D.21.C), la posició de la qual es troba dins del domini d'expressió de *ptc* a l'ala. Pel que fa als resultats obtinguts amb el *driver* MS1096Gal4, tot i que la simple expressió de la proteïna Gal4 que condueix aquest *driver* ja dóna un lleuger fenotip de "deltas" a certes venes (indicat amb un triangle negre a la figura D.21.B), la sobreexpressió de qualsevol de les dues isoformes de GAGA resulta sempre en una completa malformació de l'estructura interna de l'ala, amb un fenotip més sever en el cas de la sobreexpressió de GAGA₅₁₉. La resta del cos sembla normal en aquestes mosques (dades no mostrades). Els resultats obtinguts amb aquest *driver* anirien en concordança amb el fet que el domini d'expressió de MS1096Gal4 al disc imaginal d'ala és molt ampli, a diferència de, per exemple, el de ptcGal4 (veure figura D.18.C i D.20.D, on el senyal de GAGA correspon al domini d'expressió

de MS1096Gal4 i *ptcGal4*, respectivament). Finalment, la sobreexpressió de $GAGA_{519}$ sota el control de *dppGal4* a més de donar una alta mortalitat just al final de la fase pupa, als pocs individus que aconseguen arribar a adult s'observen malformacions a les potes, de severitat variable, arribant inclús a trobar-se alguns individus sense potes (dades no mostrades). Tot això no s'observa en cap cas per la sobreexpressió de $GAGA_{581}$ sota el control d'aquest *driver* (dades no mostrades). Per a qualsevol de les dues isoformes, amb aquest *driver*, no s'obté cap fenotip aparent a l'ala, tot i que l'ala no arriba a desplegar-se (al no sortir l'individu de la pupa) per a les mosques que sobreexpressen $GAGA_{519}$ (veure figura D.21).

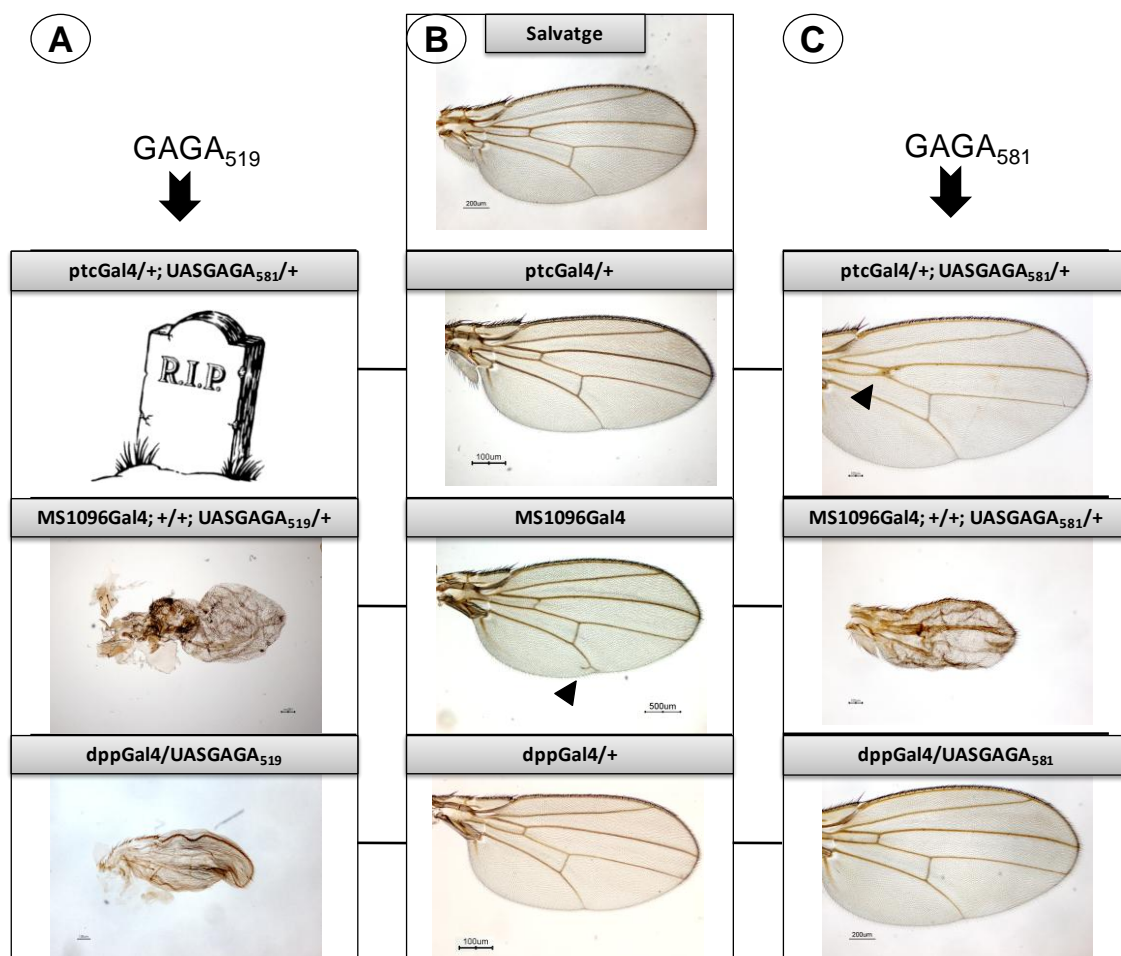


Figura D.21. Comparació dels diferents fenotips d'ala, derivats de la sobreexpressió de $GAGA_{519}$ i $GAGA_{581}$, amb els *drivers* *ptcGal4*, *MS1096Gal4* i *dppGal4*. Els triangles negres indiquen les diferents malformacions a les venes observades tant per la sobreexpressió de $GAGA_{581}$ amb el *driver* *ptcGal4* (C, primera imatge), com per la simple presència del *driver* *MS1096Gal4* (B, tercera imatge). (A) Fenotips que s'obtenen de la sobreexpressió de $GAGA_{519}$ (mostres obtingudes pel Dr. Jordi Bernués). Les mosques que sobreexpressen $GAGA_{519}$ sota el control de *ptcGal4* no arriben a adults. (B) Ales de mosques control. A dalt, ala d'una mosca salvatge que no sobreexpressa cap transgèn. Les tres restants, ales de mosques transgèniques que només porten el *driver*, mostrant així l'efecte de la sobreexpressió de Gal4 conduïda per cada *driver*. (C) Fenotips que s'obtenen de la sobreexpressió de $GAGA_{581}$. Per facilitar la comparació entre isoformes es mostren només mostres obtingudes de mascles a 18°C. Les femelles tenen un fenotip similar. A 25°C tots els fenotips són més severos.

RESULTATS

Tant les dades de letalitat, com els fenotips observats a adults, estarien d'acord amb unes funcions solapants però no idèntiques de les dues isoformes de GAGA, a més també estan en concordança amb certs estudis que indiquen que la isoforma GAGA₅₁₉ seria la més activa per la seva major capacitat de rescat d'alguns fenotips dels mutants de *Trl* (Greenberg et al. 2001).

Pel que fa als fenotips resultants de la sobreexpressió de les construccions RNAi-GAGAi, GAGAΔQ i GAGA₅₁₉DBDKO, analitzats pel Dr. Jordi Bernués (Bernues et al. 2007), cal destacar que GAGA₅₁₉DBDKO no dóna cap fenotip aparent amb cap dels *drivers* emprats, ni resulta en cap tipus de letalitat, el que indica que la unió de GAGA al DNA és vital per a dur a terme les seves funcions. Per la seva banda, la sobreexpressió de GAGAΔQ resulta en una letalitat similar a la que dóna la sobreexpressió de GAGA₅₁₉. No obstant, la sobreexpressió de GAGAΔQ sota el control del *driver* prdGal4 no presenta letalitat, mentre que aquest *driver* resulta letal a fase larvària quan condueix la sobreexpressió de GAGA₅₁₉, suggerint que l'activitat transactivadora de GAGA pot ser prescindible a certes cèl·lules. Finalment, la sobreexpressió de l'RNAi contra GAGA, tot i portar associada tota una sèrie de fenotips, és en general molt més ben tolerada que no pas la sobreexpressió de qualsevol de les dues isoformes de GAGA, com demostra la seva baixa letalitat.

D.7. GAGA reprimeix la còpia endògena de *Trl*

Els experiments presentats fins ara mostren la repressió del que s'ha definit experimentalment (Kosoy et al. 2002) com a promotor de *Trl*, fusionat als gens reporter luciferasa i GFP, tant a transfeccions transitòries a cèl·lules en cultiu com inserit al genoma de mosques transgèniques. No obstant, poc es coneix sobre el comportament de la còpia endògena de *Trl* en resposta a la sobreexpressió de GAGA. Tot i que és d'esperar un comportament similar, per tal de demostrar que la còpia endògena de *Trl* és igualment reprimible per GAGA vàrem dur a terme experiments d'RT-PCR i *Real Time* RT-PCR, tant a cèl·lules S2, com a mosques transgèniques que sobreexpressen GAGA. L'estudi es va fer a nivell d'RNA ja que una comparació directa dels nivells de proteïna és tècnicament difícil, per la impossibilitat de distingir entre la GAGA endògena i la sobreexpressada, amb els anticossos de que disposem.

D.7.1. RT-PCRs i *Real Time* RT-PCRs de cèl·lules S2 establement transfectades que sobreexpressen GAGA

Utilitzant la línia de cèl·lules S2 establement transfectades que sobreexpressen GAGA₅₁₉ sota el control del promotor de la metal·lotioneïna, i encebadors específics de la regió 5'UTR de *Trl* (no present al cDNA utilitzat per crear la línia estable d'S2), vàrem procedir a comparar per RT-PCR i posteriorment per *Real Time* RT-PCR els nivells d'mRNA endogen de *Trl* (tots els transcrits) de cèl·lules estables sense induir i induïdes afegint 500 µM de CuSO₄ al medi 24h abans de l'extracció de l'RNA total del cultiu. Aquesta inducció dona com a resultat la sobreexpressió de GAGA₅₁₉ a les cèl·lules establement transfectades (veure figura D.8 de l'apartat D.2.2.3). A més, per descartar possibles efectes de la inducció amb CuSO₄, vàrem comparar també els nivells d'mRNA endògen de *Trl* de cèl·lules S2 normals i cèl·lules S2 tractades amb CuSO₄ de la mateixa manera que les estables. Com a control de càrrega vàrem utilitzar encebadors que amplifiquen una regió de l'mRNA d'U6, un snRNA transcrit per l'RNA Pol III, que en teoria no hauria de veure's afectat per la sobreexpressió de GAGA₅₁₉.

Els experiments de RT-PCR (veure figura D.22.A) ja indiquen una davallada prou important de la quantitat d'mRNA de *Trl* al induir amb CuSO₄ les cèl·lules establement transfectades (figura D.22.A, carrils 7 i 8), respecte al mateix tipus de cèl·lules sense induir (figura D.22.A, carrils 5 i 6). Aquesta diferència no s'observa a les cèl·lules S2 normals (figura D.22.A, carrils 3 i 4 respecte carrils 1 i 2), indicant que la davallada de l'mRNA de *Trl* observada a les cèl·lules establement transfectades és deguda a la sobreexpressió de GAGA, i no a un efecte indirecte de l'addició de CuSO₄ al medi.

Per quantificar en quina mesura cauen els nivells d'mRNA de *Trl* al sobreexpressar GAGA₅₁₉ a les cèl·lules estables, vàrem procedir a realitzar experiments de *Real Time* RT-PCR amb mostres de les mateixes característiques que les emprades per a les RT-PCRs. Els resultats obtinguts (veure figura D.22.B) indiquen una baixada d'aproximadament un 75% en la quantitat d'mRNA de *Trl* al induir amb CuSO₄ la sobreexpressió de GAGA₅₁₉ a les cèl·lules establement transfectades. S'ha de destacar el fet que el simple tractament de les cèl·lules amb CuSO₄ ja produeix una certa davallada de la quantitat d'mRNA de *Trl*, com es pot veure a les dues primeres columnes de la figura D.22.B, corresponents a cèl·lules S2 normals sense tractar (-) i tractades (+). Aquesta davallada del 45% aproximadament a cèl·lules S2 estaria indicant que,

RESULTATS

del 75% de davallada que observem a les cèl·lules establement transfectades, gran part seria degut a efectes indirectes i no pròpiament a la sobreexpressió de GAGA₅₁₉. No obstant, sembla clar que, tot i que probablement no sigui una davallada dràstica, la sobreexpressió de GAGA redueix de manera significativa els nivells d'mRNA de *Trl*.

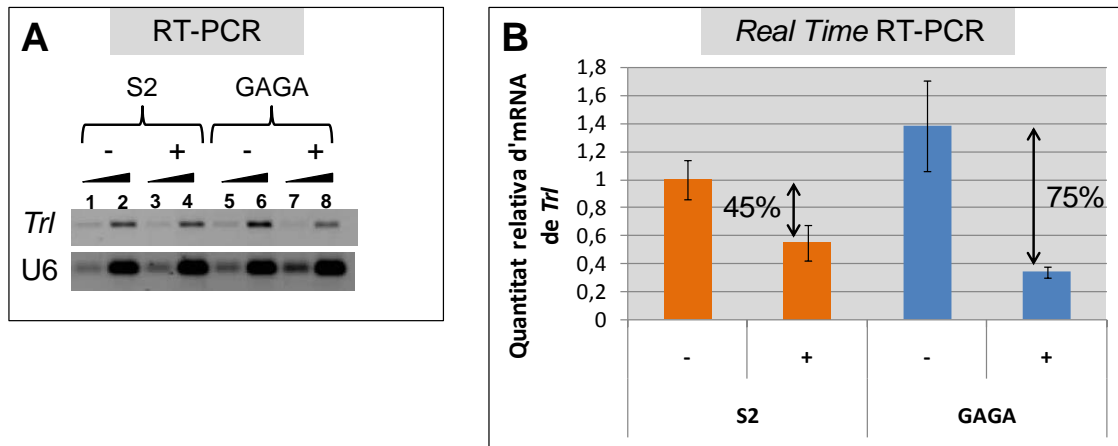


Figura D.22. GAGA reprimeix la còpia endògena de *Trl* a cèl·lules S2 establement transfectades. "S2" són les cèl·lules normals i "GAGA" són les cèl·lules S2 establement transfectades que sobreexpressen GAGA₅₁₉ sota el control del promotor de la metal·lotioneïna. El signe (-) indica que les cèl·lules no varen ser tractades, mentre que el signe (+) indica que les cèl·lules varen ser tractades afegint 500 μ M de CuSO₄ al medi 24h abans de realitzar els experiments. (A) RT-PCRs per detectar l'mRNA de la còpia endògena de *Trl* i el del control de càrrega U6. Cada reacció es va realitzar amb dues concentracions d'RNA total de partida diferents (5 ng i 50 ng, indicat pel triangle a la figura). (B) *Real Time* RT-PCR per quantificar els nivells de transcrits del gen *Trl* endògen. Els valors han estat normalitzats amb el control U6 i relativitzats al valor de les cèl·lules S2 sense tractar amb CuSO₄ (S2 -). S'indiquen també les diferències entre cèl·lules tractades i no tractades, en tant per cent.

D.7.2. RT-PCRs i *Real Time* RT-PCRs de discos imaginals de mosques transgèniques que sobreexpressen GAGA

De manera anàloga a com es mostra a l'apartat anterior, vàrem voler comprovar que a la mosca també es dona la repressió de la còpia endògena de *Trl* per la sobreexpressió de GAGA. Per fer-ho vàrem comparar, tant per RT-PCR com posteriorment per *Real Time* RT-PCR, els nivells d'mRNA de *Trl* endògen, a discos imaginals d'ala de larves de 3r estadi de dos tipus de mosques transgèniques. Per una banda mosques MS1096Gal4; +/+; UASGAGA₅₁₉/+ (que anomenarem MS1096Gal4-GAGA per simplificar); és a dir, mosques que porten el *driver* MS1096Gal4 (al cromosoma X) que dirigeix la sobreexpressió de GAGA₅₁₉ (al cromosoma 3). Per altra banda mosques MS1096Gal4; +/+; +/+ (que anomenarem MS1096Gal4 per simplificar); és a dir, mosques que únicament porten el *driver* MS1096Gal4 (al cromosoma X) i que per tant no sobreexpressen GAGA₅₁₉. Vàrem escollir el *driver* MS1096Gal4 per a sobreexpressar GAGA₅₁₉

perquè, dels *drivers* de que disposàvem al laboratori, és el d'expressió més general a disc imaginal d'ala (veure figura D.23.C) i que permet obtenir larves de 3^{er} estadi quan l'utilitzem per sobreexpressar GAGA₅₁₉ (veure taula D.1). Tots els individus emprats en aquests experiments es van fer créixer a 18°C, ja que és la temperatura a la qual la viabilitat del transgènic MS1096Gal4-GAGA és major, en concordança amb un menor nivell de sobreexpressió de GAGA₅₁₉.

Les RT-PCRs, dutes a terme amb els mateixos encebadors que els experiments de l'apartat anterior, mostren una apreciable menor quantitat d'mRNA de *Trl* endògen a discos de larves MS1096Gal4-GAGA, indicant que, als discos imaginals d'ala, la sobreexpressió de GAGA₅₁₉ reprimeix l'expressió de la còpia endògena de *Trl* (veure figura D.23.A).

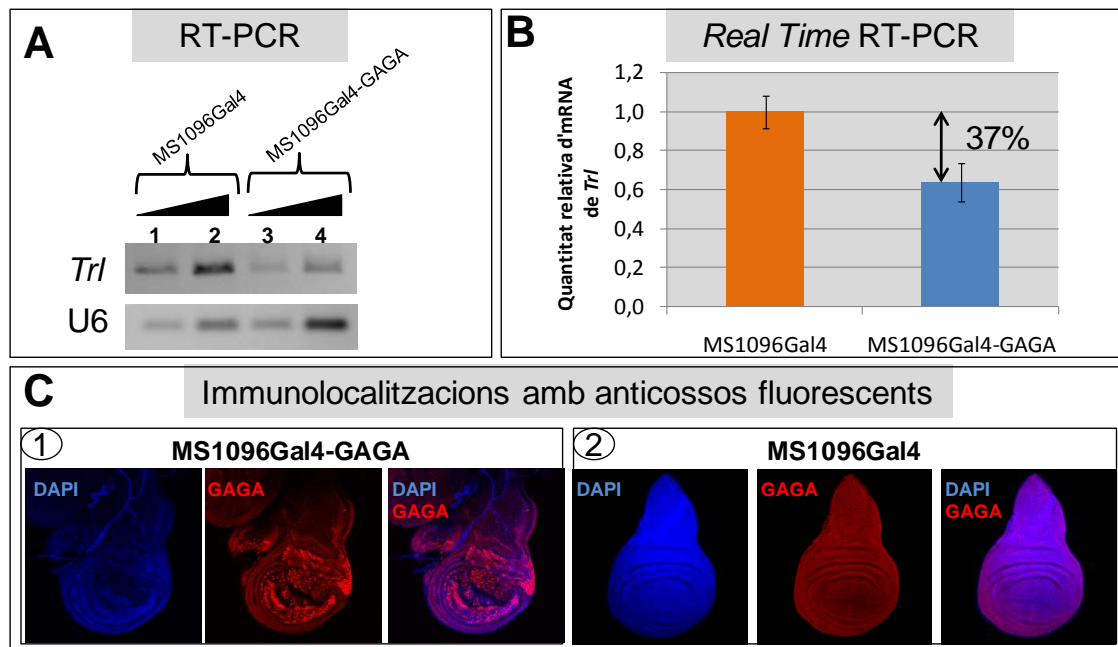


Figura D.23. GAGA reprimeix la còpia endògena de *Trl* a disc imaginal d'ala de larves de 3^r estadi. **(A)** RT-PCRs per detectar l'mRNA de la còpia endògena de *Trl* i el del control de càrrega U6. Cada reacció es va realitzar amb dues concentracions d'RNA total de partida diferents (5 ng i 20 ng, indicat pel triangle). **(B)** *Real Time* RT-PCR per quantificar els nivells de transcrits del gen *Trl* endògen. Els valors han estat normalitzats amb el control U6 i relativitzats al valor de la mostra MS1096Gal4. **(C)** Imatges de microscopia confocal (Leica TCS SPE) d'immunolocalitzacions amb anticòs anti-GAGA realitzades a discos imaginals d'ala de larves de 3^r estadi d'individus MS1096Gal4-GAGA (C.1) i MS1096Gal4 (C.2). Totes les mostres han estat obtingudes a 18°C.

Per quantificar amb més precisió els nivells d'mRNA de *Trl* endògen a cada tipus de transgènic, vàrem realitzar experiments de *Real Time* RT-PCR amb mostres de les mateixes característiques que les emprades per a les RT-PCRs. Els resultats obtinguts (veure figura D.23.B) mostren com els discos d'ala d'individus MS1096Gal4-GAGA tenen uns nivells un 37% més baixos d'mRNA de *Trl*, corroborant els resultats obtinguts per RT-PCR. Tenint en compte

RESULTATS

que la sobreexpressió de GAGA₅₁₉ amb el *driver* MS1096Gal4 escassament afecta un 40% del disc d'ala, la reducció real de l'expressió de *Trl* és força considerable i podríem estimar-la al menys en un ~70%.

D.8. Estudi funcional de la sobreexpressió i de la reducció de la quantitat de GAGA, en cèl·lules S2, per mitjà de *microarrays*

Com ja s'ha comentat a la introducció, GAGA és una proteïna molt polifacètica d'expressió aparentment ubíqua. S'ha vist implicada en la regulació de multitud de gens, actuant a múltiples nivells, resultant tant en activació com en repressió. Així doncs, per tal de dur a terme un estudi a nivell global de les funcions de GAGA a *Drosophila*, vàrem realitzar experiments de *microarray* a cèl·lules S2. Vam utilitzar cèl·lules S2 normals, així com dues línies d'S2 establiment transfectades: la línia GAGA, ja utilitzada anteriorment (veure apartats D.2.2.3 i D.7.1), que sobreexpressa GAGA₅₁₉ sota el control del promotor de la metal·lotioneïna, induïble per la presència de metalls pesants al medi; i la línia RNAiGAGA, que sota el control del mateix promotor, expressa un *hairpin* d'RNA que cobreix part dels dominis POZ i X de GAGA, i que, actuant com a RNA de doble cadena, s'incorporarà a la maquinària de l'RNA d'interferència, resultant en la degradació activa dels mRNAs d'ambdues isoformes de GAGA. Amb aquestes línies cel·lulars vàrem realitzar dos tipus d'experiments de *microarray* complementaris.

➤ **Experiment de sobreexpressió:** comparem els resultats dels *microarrays* de cèl·lules control amb els de cèl·lules de la línia GAGA induïda amb 500µM de CuSO₄ 24h. La inducció es realitza durant 24h perquè hem comprovat que és un temps suficient per a obtenir una bona sobreexpressió de GAGA en aquesta línia cel·lular (veure figura D.24.A).

- Control: és la mitjana conjunta dels resultats obtinguts de dues mostres independents de cèl·lules S2 normals induïdes 24h amb 500µM de CuSO₄ i de dues mostres independents de cèl·lules de la línia GAGA sense induir. Es va decidir ajuntar aquests dos tipus de control per tal d'augmentar la significació estadística de les dades, ja que aquests dos tipus de mostra, tot i no ser idèntiques, eren molt més

semblants entre elles que entre les que sobreexpressen GAGA, indicant que la sobreexpressió de GAGA té molt més efecte sobre les cèl·lules que no pas la inducció amb CuSO_4 durant 24h.

- GAGA: és la mitjana conjunta dels resultats obtinguts de quatre mostres independents de cèl·lules de la línia GAGA induïdes 24h amb $500\mu\text{M}$ de CuSO_4 i que, per tant, sobreexpressen GAGA_{519} .

➤ **Experiment d'RNAi:** comparem els resultats dels *microarrays* de cèl·lules control amb els de cèl·lules de la línia RNAiGAGA induïdes amb $500\mu\text{M}$ de CuSO_4 durant 96h. La inducció es realitza durant 96h ja que hem comprovat que la vida mitjana de la proteïna GAGA és molt llarga, d'uns dos o tres dies (dades no mostrades), i 96h és el temps mínim per veure una baixada significativa en la quantitat de proteïna (veure figura D.24.C).

- Control: és la mitjana conjunta dels resultats obtinguts de dues mostres independents de cèl·lules S2 normals induïdes 96h amb $500\mu\text{M}$ de CuSO_4 .
- RNAiGAGA: és la mitjana conjunta dels resultats obtinguts de dues mostres independents de cèl·lules de la línia RNAiGAGA induïdes 96h amb $500\mu\text{M}$ de CuSO_4 i que per tant expressen el *hairpin* de GAGA que actuarà com a RNAi contra aquesta.

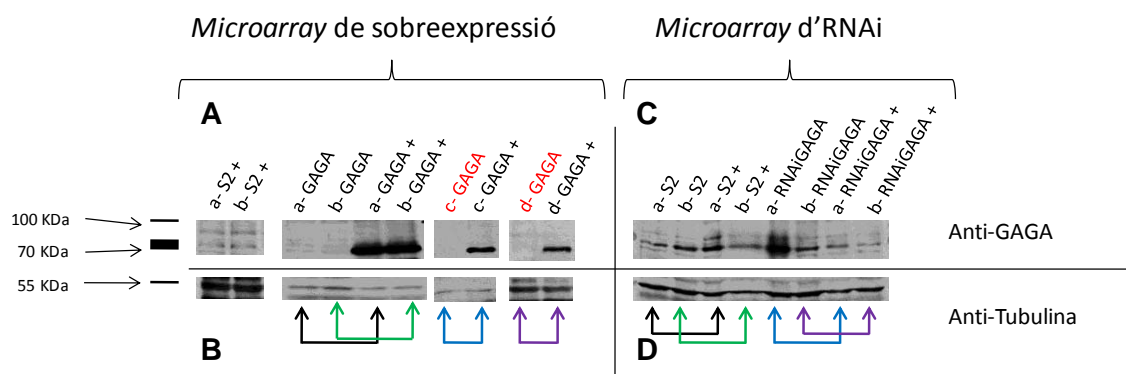


Figura D.24. Comprovació per *western blot* de la correcta inducció de les cèl·lules emprades pels experiments de *microarray*. El signe (+) indica que les mostres han estat tractades amb $500\mu\text{M}$ de CuSO_4 durant 24h en el cas del *microarray* de sobreexpressió i 96h en el d'RNAi. La lletra davant el nom de les mostres, així com les fletxes sota els carrils, indiquen les mostres que es corresponen, ja que van ser separades del cultiu inicial 24h abans del tractament amb CuSO_4 . A dalt, figures (A) i (C), *western blot* anti-GAGA de les mostres indicades. En vermell, dues mostres GAGA sense induir que no vàrem emprar als *microarrays*, però que s'inclouen a la imatge per comprovar la correcta sobreexpressió de la seva corresponent mostra induïda. A baix, figures (B) i (D) *western blot* anti-tubulina de les mateixes membranes, com a control de càrrega.

Després d'analitzar estadísticament les dades (tal com es descriu a Mètodes, apartat C.2.5.5), es varen obtenir tot un conjunt de gens diferencialment expressats tant a l'experiment de sobreexpressió com al d'RNAi. Per l'experiment de sobreexpressió vàrem obtenir un total de

RESULTATS

127 gens diferencialment expressats, a més del propi gen *Trl*, que com és d'esperar, surt augmentat a les mostres GAGA, ja que l'estem sobreexpressant amb la inducció per CuSO_4 . Per l'experiment d'RNAi obtenim 917 gens diferencialment expressats. Aquests resultats són consistents amb el fet que a l'experiment de sobreexpressió la inducció és molt més curta (24h) i per tant tenim major probabilitat de veure gens afectats directament per GAGA, mentre que la llarga inducció necessària per a reduir significativament la quantitat de GAGA per RNAi (96h) resulta en una major proporció de gens afectats de manera indirecta. En qualsevol cas, tot i que per *western blot* detectem la reducció de GAGA a la línia RNAiGAGA, els resultats dels *microarrays* no mostren *Trl* com un gen reprimat. Aquest fet, junt amb una explicació més detallada de la disposició de totes les sondes per a *Trl* emprades al *microarray*, es discutirà amb més detall a l'apartat E.2 de la discussió.

Com exemple dels resultats obtinguts, a la figura D.25.A es mostren els 20 gens més estadísticament significatius que es troben activats, així com els únics 10 gens que es troben reprimats, a l'experiment de sobreexpressió. A la figura D.25.B es mostren els 15 gens més estadísticament significatius que es troben activats i reprimats, a l'experiment d'RNAi. Per últim, a la figura D.25.C es representa esquemàticament la proporció de gens activats i reprimats a cada experiment. Tenint en compte els resultats obtinguts, es fa evident que GAGA actua clarament com a proteïna activadora de l'expressió gènica, fet que es fa més evident en l'experiment de sobreexpressió, on només es troben 10 gens regulats negativament per GAGA, front de 117 regulats positivament. A l'experiment d'RNAi també s'observa la tendència activadora de GAGA, tot i que no de forma tan clara com a l'experiment de sobreexpressió, degut probablement al major temps de tractament, que potencia els efectes indirectes.

A

SOBREEXPRESSIÓ DE GAGA

Nº	Sonda Affymetrix	Símbol del Gen	Nom del Gen	Control	GAGA	Fold Change	Fold Ci	P
1	1635665_at	Tdc1	Tyrosine decarboxylase 1	3,0078469	5,49342739	5,600596493	(4.92, 6.37)	0,99999951
2	1631701_a_at	Cpr49Ac	Cuticular protein 49Ac	4,65377554	8,42100621	13,16599651	(10.4, 17.82)	0,99999391
3	1629754_s_at	Trl	Trithorax-like	8,6309686	11,7104247	8,452957238	(5.87, 12.18)	0,99973033
4	1637513_at	CG11275	CG11275	3,38582011	5,02948792	3,124591981	(2.57, 3.8)	0,99971786
5	1629085_at	CG14227	CG14227	4,68801854	6,86216073	4,513173342	(3.33, 6.11)	0,99919206
6	1632923_at	CG14661	CG14661	3,15552012	4,63015858	2,779139886	(2.25, 3.43)	0,99905853
7	1630223_at	skl	sickle	2,54163039	4,40432279	3,636857516	(2.76, 4.78)	0,99881814
8	1638956_at	Fas2	Fasciclin 2	3,4171935	6,13029932	6,557317817	(4.39, 9.79)	0,99878489
9	1632377_at	CG8399	CG8399	5,04177417	6,12922851	2,12498749	(1.81, 2.5)	0,99873639
10	1628146_at	crb	crumbs	5,51817677	6,94545766	2,689393555	(2.15, 3.36)	0,99828399
11	1639896_at	ND	ND	2,69064946	4,01080832	2,496936026	(2.03, 3.08)	0,99812944
12	1623238_at	CG5618	CG5618	6,11659228	7,42603585	2,478459302	(1.98, 3.1)	0,99701234
13	1641167_s_at	CG7047	CG7047	3,66855869	4,94223689	2,417771994	(1.95, 3.01)	0,99689912
14	1628867_s_at	CG33334	CG33334	4,22696313	5,38350531	2,29224914	(1.82, 2.72)	0,99658113
15	1641120_at	CG13506	CG13506	2,68233645	4,49318354	3,508482338	(2.56, 4.81)	0,99645253
16	1632591_at	Phk-3	Pherokine 3	3,54729262	6,08519422	5,80743699	(3.7, 9.11)	0,9960472
17	1633935_at	skf	skiff	3,29416812	4,59857297	2,469818193	(1.95, 3.12)	0,99565608
18	1632299_at	alphaP54	alphaP54	4,73012632	6,24621875	2,860153205	(2.17, 3.76)	0,99556364
19	1641530_s_at	CG2264	CG2264	3,859303	6,34255533	5,591565761	(3.54, 8.84)	0,99494331
20	1639592_at	CG3624	CG3624	5,43054243	6,85079094	2,676316084	(2.06, 3.48)	0,99493971
...								
118								
1	1639654_at	CG3603	CG3603	5,3487099	4,60951763	1,669241013	(1.42, 1.96)	0,97923821
2	1625443_a_at	CG33054	CG33054	8,99924191	8,15829818	1,791221481	(1.47, 2.18)	0,96467498
3	1636931_at	br	broad	6,0931413	5,36587377	1,655500592	(1.39, 1.98)	0,95581282
4	1639645_at	CG9740	CG9740	9,24963325	8,75988663	1,404198236	(1.22, 1.61)	0,90528519
5	1633545_at	PGRP-SD	PGRP-SD	9,09282053	7,97230042	2,174253428	(1.58, 2.98)	0,90506393
6	1627842_at	CG14898	CG14898	7,27067159	6,72536636	1,45932906	(1.24, 1.71)	0,88902859
7	1624946_at	CG6790	CG6790	4,40737171	3,69639284	1,636914391	(1.33, 2.02)	0,88240409
8	1631626_at	aub	aubergine	8,91843079	7,91392558	2,006255317	(1.49, 2.7)	0,88017014
9	1633645_at	CG13512	CG13512	6,08909682	5,30929105	1,716899714	(1.36, 2.18)	0,86658799
10	1636304_at	egg	CG12196 gene product from transcript CG12196-RA	6,72427708	5,72267372	2,00223973	(1.48, 2.72)	0,86395515



B

REDUCCIÓ DE GAGA AMB RNAi

Nº	Identificador d'Affymetrix	Símbol del Gen	Nom del Gen	Control	RNAiGAGA	Fold Change	Fold Ci	P
1	1631016_at	CG32822	CG32822	2,823616986	6,139918974	9,961078739	(8.91, 11.14)	0,999996
2	1635481_a_at	CG6659	CG6659	3,815550987	6,983838744	8,989792135	(7.61, 10.62)	0,99996425
3	1625623_at	Nrx-1	Neurexin 1	6,592157814	9,322629144	6,636724236	(5.74, 7.67)	0,99996327
4	1628503_at	CG30424	CG30424	4,599791124	6,340209641	3,341320835	(2.99, 3.74)	0,99990845
5	1629702_a_at	ab	abrupt	2,797661329	5,085686371	4,883870826	(4.2, 5.68)	0,99989299
6	1634707_s_at	Gfat	Glucosamine-Fructose-6-Phosphate Aminotransferase	7,35525436	8,938830914	2,997119389	(2.67, 3.37)	0,99982574
7	1632582_at	CG9270	CG9270	4,297991439	6,297093409	3,997510908	(3.43, 4.65)	0,99980557
8	1639729_s_at	ND	ND	7,731760786	9,169524673	2,709006559	(2.42, 3.03)	0,99977686
9	1627930_s_at	CG32280	CG32280	7,111065684	8,446735921	2,523927093	(2.26, 2.82)	0,99972218
10	1627263_at	CG31601	CG31601	3,399609467	5,255419111	3,619548238	(3.1, 4.23)	0,99970972
11	1631474_s_at	l(2)k05713	lethal (2) k05713	7,599887565	8,926074448	2,50739082	(2.24, 2.8)	0,9997001
12	1629786_a_at	CG31116	CG31116	6,488951229	7,875281056	2,614128078	(2.32, 2.94)	0,99967321
13	1623529_s_at	Or19b	Odorant receptor 19b	3,954243507	5,532069859	2,985197436	(2.6, 3.42)	0,99965861
14	1630366_at	CG31149	CG31149	5,192697457	6,742973247	2,928731202	(2.56, 3.35)	0,9996398
15	1632491_at	CG32985	CG32985	7,328663857	8,809075577	2,790283517	(2.45, 3.18)	0,99963335
...								
392								
1	1623950_s_at	ama	Amalgam	9,977335528	6,264821593	13,10925627	(11.66, 14.74)	0,99999733
2	1634959_at	TFIIA-S-2	TFIIA-S-2	6,114271479	3,015913718	8,564433138	(7.48, 9.81)	0,99998772
3	1637421_at	Chit2	Chitinase 2	9,510334475	7,004268174	5,686904991	(5.03, 6.41)	0,99998106
4	1640729_s_at	nrv3	nervana 3	7,4433584	4,865948768	5,968670583	(5.26, 6.77)	0,9999801
5	1638892_s_at	CG5397	CG5397	8,80648086	6,418545515	5,234077715	(4.6, 5.95)	0,99997001
6	1635894_at	CG30148	CG30148	9,645110408	7,405480461	4,722759098	(4.18, 5.33)	0,99996867
7	1629857_at	CG10359	CG10359	6,998961785	4,968705058	4,084775324	(3.64, 4.58)	0,99996282
8	1639147_s_at	CG30438	CG30438	8,372512564	6,191792673	4,533797298	(4.5, 14)	0,99995929
9	1633989_at	CG1124	CG1124	6,206941097	4,062739119	4,420476764	(3.89, 5.02)	0,99995076
10	1637975_at	Dhc36C	Dynein heavy chain at 36C	7,289551635	5,248651292	4,115022566	(3.64, 4.65)	0,99994936
11	1628174_at	CG33119	CG33119	8,133045087	6,004872056	4,371635241	(3.83, 4.98)	0,99994303
12	1628690_at	CG1503	CG1503	7,092814002	4,266017991	7,094967181	(5.9, 8.53)	0,99992658
13	1623103_at	CG14946	CG14946	8,985851416	6,418171666	5,928551869	(5, 7.03)	0,99992058
14	1632213_s_at	CG31292	CG31292	4,922846706	3,156439181	3,4020575	(3.02, 3.83)	0,99991285
15	1625611_at	CG33115	CG33115	8,803818309	7,133946256	3,181863737	(2.84, 3.57)	0,99990717

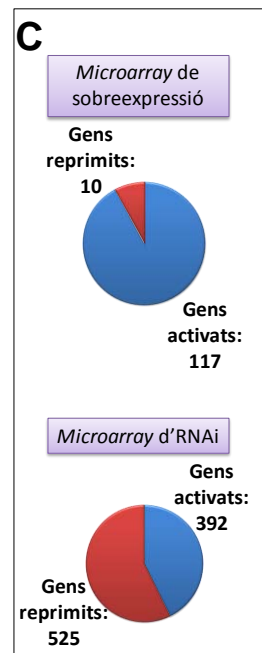


Figura D.25. Resultats més estadísticament significatius dels dos experiments de *microarray*. A la columna “control” i “GAGA” es mostra el valor mitjà en escala logarítmica. La columna “Fold Change” indica el nivell de canvi entre les mostres “control” i les “GAGA”, en escala lineal. La columna “Fold Ci” indica l’interval de confiança del “Fold Change”, en escala lineal. El valor “P” és una estimació estadística de la probabilitat de que el gen en qüestió estigui realment expressat de forma diferencial entre les mostres “control” i “GAGA”, sent 1 el 100% de probabilitat. Tots els gens estan ordenats de forma decreixent tenint en compte la seva significació estadística (valor “P”). **(A)** Experiment de sobreexpressió. Es mostren els 20 gens activats més significatius i els 10 gens reprimits. **(B)** Experiment d’RNAi. Es mostren els 15 gens més significatius, tant dels que es troben activats com dels que es troben reprimits. **(C)** Representació esquemàtica de la proporció de gens activats i reprimits a cada experiment de *microarray*. **Nota:** a l’experiment de sobreexpressió descomptem *Trl* de la llista de gens sobreexpressats.

RESULTATS

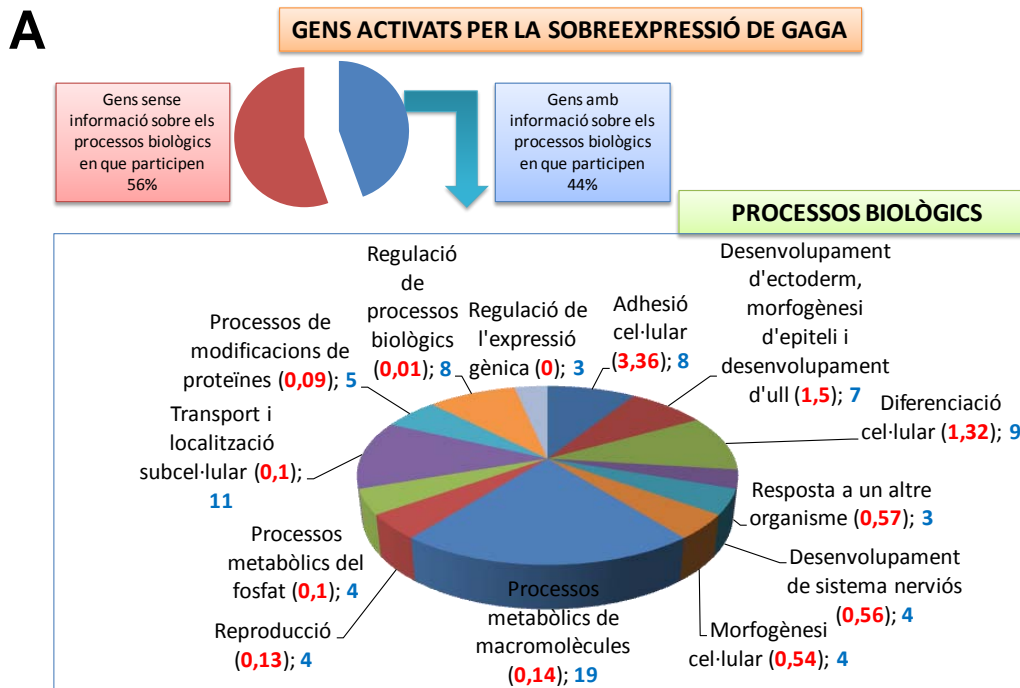
A fi de destacar els diferents processos biològics en que GAGA pot estar participant, vàrem analitzar quins són els més representats a la llista completa de gens diferencialment expressats de cada experiment de *microarray*, agrupant-los utilitzant l'eina bioinformàtica DAVID (*Database for Annotation, Visualization and Integrated Discovery*) disponible *online* a l'adreça <http://david.abcc.ncifcrf.gov/home.jsp>.

A la figura D.26.A es mostra la proporció de gens que s'activen amb la sobreexpressió de GAGA dels que es té informació sobre els processos biològics en que poden participar. Aquesta informació pot ser que estigui demostrada experimentalment o que s'hagi inferit per homologia. També es representen els grups que es poden formar, segons processos biològics, amb els gens dels que es disposa informació. Per augmentar el sentit biològic de cada grup, aquests tenen una sèrie de requisits (imposats pel programa DAVID) que fan que no tots els gens pugin estar formant part d'un grup. A més, degut a que alguns gens estan implicats en més d'un procés, poden estar formant part de més d'un grup. Els grups s'han ordenat de més estadísticament significatius a menys, seguint el sentit de les agulles del rellotge. Aquesta significació estadística es representa pel valor *enrichment score* que s'obté amb l'eina DAVID i és una estimació de quant enriquida està la llista de gens en una funció biològica concreta, respecte del que es podria donar per l'atzar. L'àrea que ocupa cada procés biològic representat a la figura ve donada pel número de gens que formen cada grup. Com es pot observar a la figura, el grup més significatiu (*enrichment score* més alt, de 3,36 en aquest cas) és el que correspon a la funció biològica de gens relacionats amb l'adhesió cel·lular.

Pel tractament de les dades dels gens que apareixen reprimits per la sobreexpressió de GAGA vàrem seguir una estratègia diferent ja que, degut a que només eren 10 gens, resulta complicat agrupar-los en funcions biològiques. Així, a la figura D.26.B es mostren aquests gens amb una breu descripció de les característiques conegudes.

De manera anàloga, pel tractament de les dades de l'experiment d'RNAi, vàrem agrupar tant els gens activats (veure figura D.27.A) com reprimits (veure figura D.27.B) per la reducció de GAGA. En aquests cassos però, només es representen els grups amb un *enrichment score* més alt (major o igual a 0,3) per facilitar la representació, ja que el número de gens és major i per tant també ho és el número de grups, segons funció biològica, que es poden obtenir. Cal destacar que, en favor de la fiabilitat dels resultats obtinguts, el grup més significatiu obtingut

dels gens que es reprimeixen en reduir GAGA per RNAi, adhesió cel·lular, és precisament el més significatiu obtingut dels gens que s'activen per sobreexpressió de GAGA (veure figura D.26.A).



B

GENS REPRIMITS PER LA SOBREEXPRESSIÓ DE GAGA

Nº	Identificador d'Affymetrix	Símbol del Gen	Nom del Gen	Descripció
1	1639654_at	CG3603	CG3603	Homologia amb la família SDR (<i>short-chain dehydrogenases/reductases</i>). Possible relació amb metabolisme de lípids, transducció de senyals i comunicació cel·lular
2	1625443_a_at	CG33054	CG33054	Homologia amb proteïnes de processament de l'Appr-1" (ADP-ribose-1"monophosphate)
3	1636931_at	br	broad	Proteïna amb domini POZ requerida pel "puffing" i la transcripció de gens tardans de glàndula salival durant la metamorfosi
4	1639645_at	CG9740	CG9740	Possible interacció amb certes proteïnes com mRpL51(<i>mitochondrial 39S ribosomal protein L51</i>)
5	1633545_at	PGRP-SD	Peptidoglycan-recognition protein-SD precursor	Implicat en la resposta immune unint peptidoglicans de bacteris gram positius i activant la via de Toll
6	1627842_at	CG14898	CG14898	Possible inteacció amb la proteïna mtacp1 (<i>mitochondrial acyl carrier protein 1</i>)
7	1624946_at	CG6790	CG6790	Possible implicació en la biosíntesi de glicosilfosfatidilinositol (GPI)
8	1631626_at	aub	aubergine	Implicat en la via del RNA d'interferència
9	1633645_at	CG13512	CG13512	Funció molecular d'unió a proteïnes predita de la seva seqüència
10	1636304_at	egg	CG12196 gene product from transcript CG12196-RA	Proteïna amb domini SET (<i>Su(var)3-9, Enhancer-of-zeste, Trithorax</i>). Sembla necessari a l'ovari per la trimetilació de la lisina 9 de la H3

Figura D.26. Agrupació, segons la seva funció biològica, dels gens diferencialment expressats a l'experiment de sobreexpressió, utilitzant l'eina DAVID. **(A)** A dalt, proporció de gens activats per la sobreexpressió de GAGA dels que es coneix la seva implicació en algun procés biològic. A baix, agrupació d'aquests gens en diverses funcions biològiques. Els diferents grups estan ordenats seguint el sentit de les agulles del rellotge, segons la seva significació estadística, donada per l'enrichment score (indicat entre parèntesi en vermell). També s'indica el número de gens a cada grup (en blau). Es representen tots els grups possibles, seguint criteris de crivellatge donats pel programa DAVID. **(B)** Relació dels 10 gens que es troben reprimits per la sobreexpressió de GAGA, amb una breu descripció de les característiques conegudes de cadascun.

RESULTATS

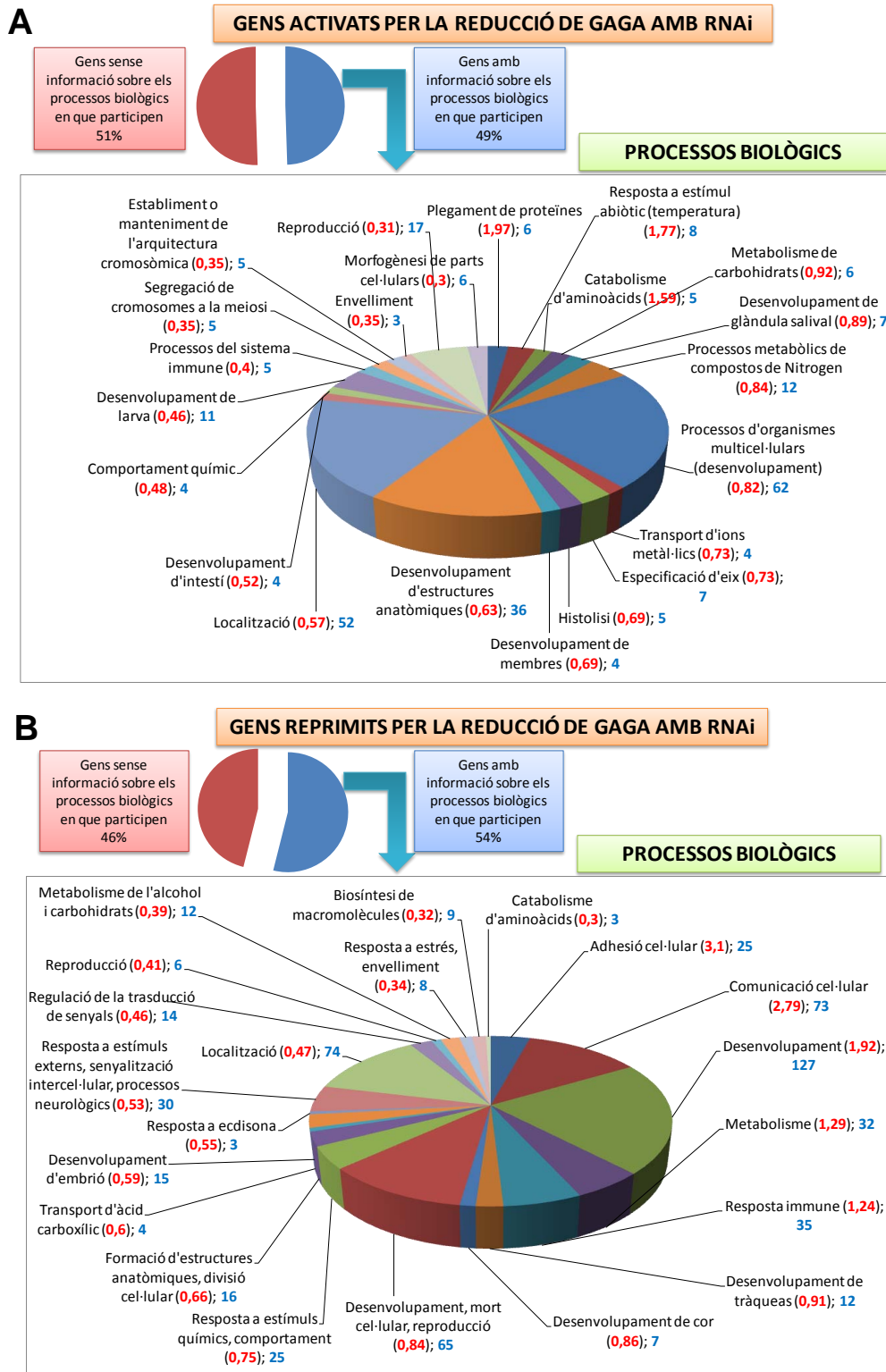


Figura D.27. Agrupació, segons la seva funció biològica, dels gens diferencialment expressats a l'experiment d'RNAi, utilitzant l'eina DAVID. Els diferents grups estan ordenats seguint el sentit de les agulles del rellotge, segons la seva significació estadística, donada per l'*enrichment score* (indicat entre parèntesi en vermell). També s'indica el número de gens a cada grup (en blau). Es representen només els grups més significatius (*enrichment score* $\geq 0,3$), seguint criteris de crivellatge donats pel programa DAVID. **(A)** A dalt, proporció de gens activats per la reducció de GAGA amb RNAi, dels quals es coneix la seva implicació en algun procés biològic. A baix, agrupació d'aquests gens en diverses funcions biològiques. **(B)** A dalt, proporció de gens reprimits per la reducció de GAGA amb RNAi, dels quals es coneix la seva implicació en algun procés biològic. A baix, agrupació d'aquests gens en diverses funcions biològiques.

Finalment, vàrem creuar les dades obtingudes en els dos tipus de *microarray*, a fi de trobar els gens regulats per GAGA de forma més consistent, ja que els que s'activen per la sobreexpressió de GAGA és probable que es reprimeixin per la reducció d'aquesta, i viceversa. Hi ha 25 gens que estan diferencialment expressats en els dos tipus de *microarray* (veure figura D.28, diagrama de Venn). A la figura també es representa el sentit d'aquesta expressió diferencial, a cada tipus de *microarray*. Així, com és d'esperar, la majoria d'aquests 25 gens es troben activats al sobreexpressar GAGA i reprimits al reduir la quantitat d'aquesta per RNAi. Tot i amb això, existeixen uns pocs gens que segueixen la mateixa tendència en tots dos experiments de *microarray*, és a dir, que s'activen o es reprimeixen en tots dos experiments, cosa que no seria esperable si estiguessin regulats directament per GAGA. És de destacar que només un gen, PGRP-SD (*peptidoglycan-recognition protein-sd precursor*), es troba reprimit per la sobreexpressió de GAGA i activat en reduir la quantitat d'aquesta per RNAi. Aquest gen està implicat en resposta immune contra bacteris grampositius i com es pot veure a la figura D.27.A, la funció de resposta immune es troba significativament sobrerrepresentada (*enrichment score* de 0,4) als gens que s'activen per reducció de GAGA amb RNAi, obrint la possibilitat de que GAGA estigui regulant negativament aquesta funció. Tot i amb això el panorama és més complex, ja que la resposta immune també es troba sobrerrepresentada (*enrichment score* d'1,24) als gens que es reprimeixen al reduir GAGA per RNAi (veure figura D.27.B), encara que aquest grup està format per altres gens.

Així doncs, tenint en compte la gran coherència dels resultats obtinguts en aquests dos tipus *microarrays* realitzats a cèl·lules S2, podem dir que la funció predominant de GAGA probablement sigui l'activació transcripcional. A més, a l'espera de ser confirmats en altres sistemes diferents de línies cel·lulars, els nostres resultats obren la possibilitat de que GAGA estigui implicada en multitud de processos a més dels ja coneguts, com poden ser l'adhesió cel·lular i la resposta immune.

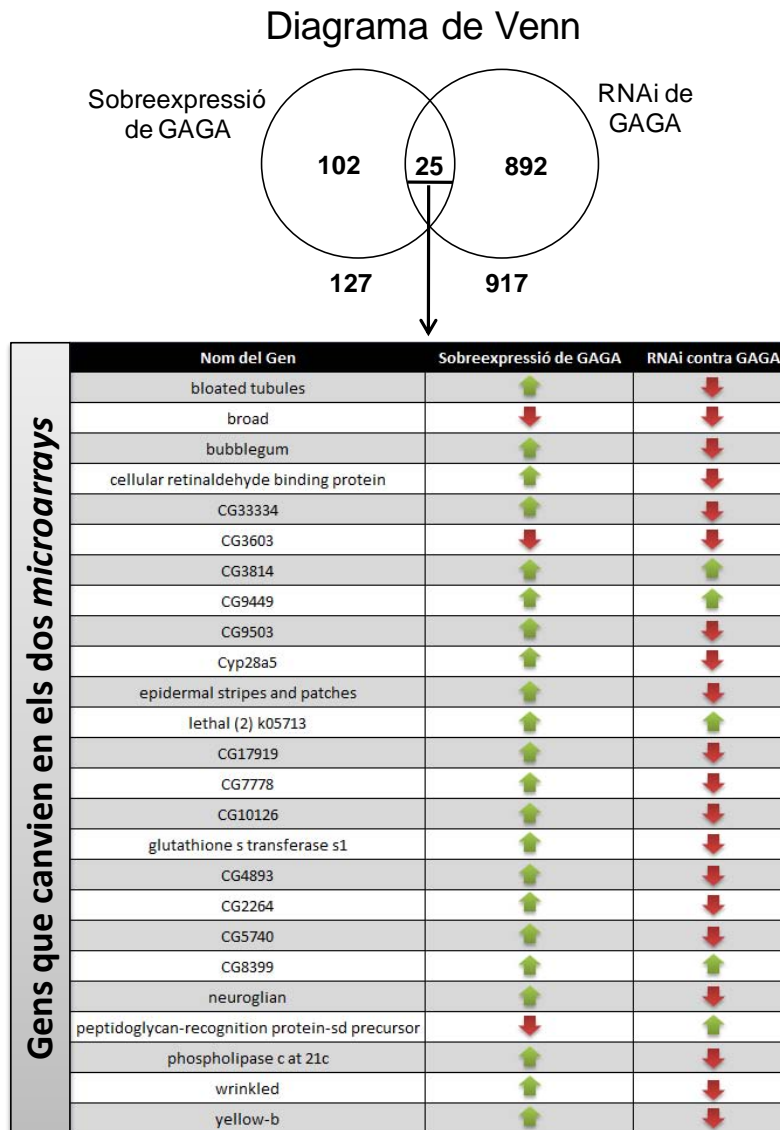


Figura D.28. Creuament de les dades obtingudes amb els dos tipus d'experiment de *microarray*. A dalt, diagrama de Venn que representa el solapament de gens entre els dos experiments. A baix, relació dels 25 gens diferencialment expressats que comparteixen els dos *microarrays*, indicant si s'activa (fletxa verda cap amunt) o si es reprimeix (fletxa vermella cap a baix) en cada cas.

D.8.1. *Sickle* s'activa per la sobreexpressió de GAGA

Els experiments de *microarray* aquí presentats formen part d'un projecte més ampli, que pretén l'anàlisi funcional *in vivo* de GAGA, no només a la línia cel·lular S2, sinó també a múltiples sistemes amb més rellevància biològica com ara certs teixits de larva, on es pensa que GAGA realitza funcions vitals per a la mosca. Per aquesta raó, tenint només les dades de cèl·lules S2, no vàrem entrar en una comprovació detallada dels possibles gens controlats per GAGA. No obstant això, un dels gens positius ens va cridar l'atenció i el vàrem corroborar per

altres tècniques. La baixa viabilitat de les mosques transgèniques que sobreexpressen GAGA, junt amb alguns dels fenotips obtinguts amb la sobreexpressió de GAGA conduïda per *dppGal4*, on la pèrdua o malformació a les potes suggereix un mecanisme de mort cel·lular (Bernues et al. 2007), ens van portar a pensar que GAGA, en determinats llinatges cel·lulars, podria estar controlant alguna de les vies que porten a l'apoptosi. Per aquesta raó, que el gen *sickle*, implicat en aquest procés (Srinivasula et al. 2002; Wing et al. 2002), aparegui com activat al sobreexpressar GAGA (veure figura D.25.A), ens va portar a corroborar aquest positiu del *microarray* per altres tècniques. Així, vàrem realitzar immunolocalitzacions amb anticossos anti-*sickle* a la línia cel·lular d'S2 establiment transfectada emprada al *microarray* de sobreexpressió per determinar si la inducció de la sobreexpressió de GAGA es correlacionava amb un augment de la quantitat de *sickle* a aquestes cèl·lules. Efectivament, es pot comprovar que al induir la sobreexpressió de GAGA, podem detectar *sickle* a les cèl·lules que sobreexpressen GAGA, tot i que, o bé no a totes elles o bé no sempre amb la mateixa intensitat, ja que es poden veure cèl·lules que sobreexpressen GAGA on no es detecta *sickle* (veure figura D.29.A, triangles blancs). El que sí sembla clar però, és que les cèl·lules que més intensament sobreexpressen GAGA són en les que més fiablement es detecta *sickle* (veure figura D.29.A, fletxes blanques). Aquestes dades, sumades a resultats obtinguts pel Dr. Jordi Bernués, consistents en experiments d'immunolocalització amb anti-*sickle* a discos d'ala de larves de 3r estadi, que sobreexpressen *GAGA₅₁₉* sota el control del *driver* *dppGal4* (veure figura D.30), confirmen que al sobreexpressar GAGA, la quantitat de *sickle* augmenta. Aquests resultats apunten a que *sickle*, un gen implicat en apoptosi, pot estar sota el control de GAGA.

RESULTATS

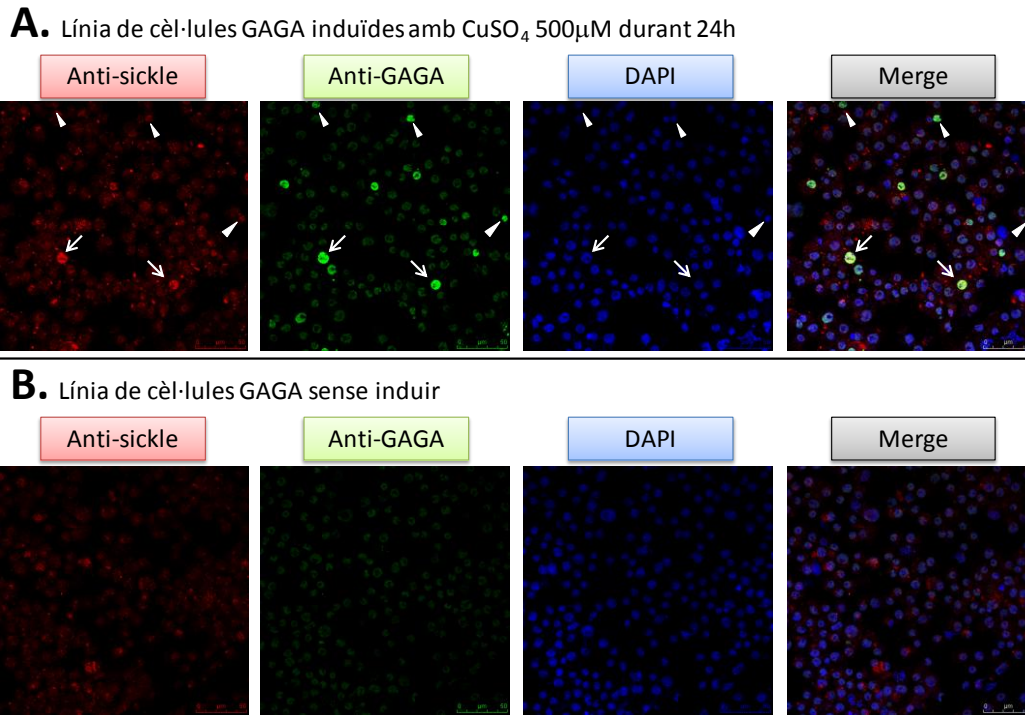


Figura D.29. La sobreexpressió de GAGA_{519} fa augmentar sickle a S2, al menys a un cert nombre de cèl·lules. Les imatges han estat captades amb un microscopi confocal Leica TCS SP5. **(A)** Cèl·lules de la línia GAGA que la sobreexpressen per mitjà de la inducció amb CuSO_4 . Les fletxes blanques indiquen les cèl·lules que més intensament sobreexpressen GAGA (en verd), a les quals és pot detectar clarament sickle (en vermell). Els triangles blancs indiquen cèl·lules on, tot i sobreexpressar GAGA de manera clara, no es pot assegurar que sickle estigui augmentat. **(B)** Control de les mateixes cèl·lules on no s'indueix la sobreexpressió de GAGA.

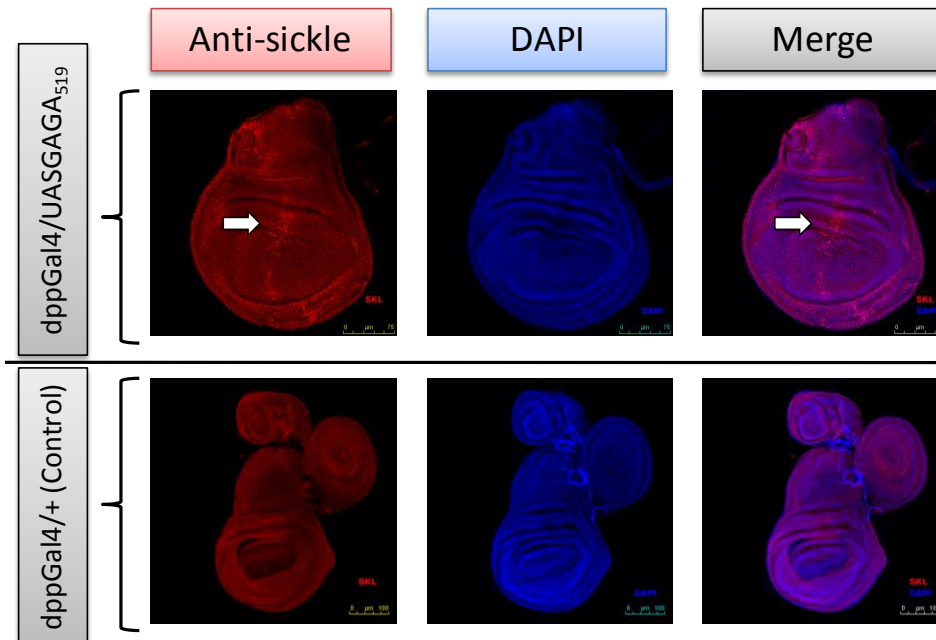


Figura D.30. La sobreexpressió de GAGA_{519} fa augmentar sickle a discos d'ala de larves de 3r estadi. Imatges obtingudes amb un microscopi confocal Leica TCS SP5. En vermell es pot veure sickle, detectat amb l'anticòs anti-sickle. En blau apareix el DAPI. A dalt, immunolocalitzacions de discos d'ala de larves que sobreexpressen GAGA_{519} sota el control del *driver* *dppGal4*. La fletxa blanca indica com la detecció de sickle segueix el patró d'expressió del *driver* *dppGal4*. A baix, immunolocalitzacions control de discos d'ala de larves que només porten el *driver* i per tant, no sobreexpressen GAGA_{519} .

E. DISCUSSIÓ

E.1. Estudi de la repressió de *Trl* per GAGA

En aquest treball ens hem proposat aprofundir en l'estudi del factor de transcripció GAGA, de *Drosophila melanogaster*. GAGA, és un factor essencial per a la mosca amb múltiples funcions. Descrita inicialment com a proteïna del TrxG (Farkas et al. 1994), per la seva capacitat de mantenir l'expressió de certs gens homeòtics, també s'ha vist implicada en el seu silenciament en algunes circumstàncies, tenint així característiques de proteïna del PcG (Hagstrom et al. 1997; Horard et al. 2000; Hodgson et al. 2001). A mesura que el coneixement sobre aquesta proteïna augmentava, es va veure que juga un paper important en la transcripció de molts gens, no només d'alguns homeòtics, ja sigui funcionant com a factor antirepressor, alleujant la repressió que l'estructura de la cromatina imposa sobre el DNA (Wilkins et al. 1997), ja sigui funcionant com a factor activador de la transcripció, estimulants-la activament (Vaquero et al. 2000; Vaquero et al. 2008). A més, també sembla jugar un paper important en l'estructura de la cromatina i en la funció d'alguns elements *insulator* (Lehmann 2004). Així doncs, tenint en compte tots els processos en que GAGA està implicada i la seva funció, aquesta proteïna ha estat vista generalment com a un factor activador de la transcripció, que pot actuar a múltiples nivells. De fet, l'estructura de la proteïna està d'acord amb aquesta idea, al disposar d'un domini d'activació de la transcripció (el domini Q, a C-terminal de la proteïna), que li confereix la capacitat transactivadora, suposadament per interacció amb la maquinària basal de la transcripció (Vaquero et al. 2000; Vaquero et al. 2008). Per tot això, va sorprendre trobar que GAGA regula l'activitat del seu propi promotor, reprimint la seva expressió (Kosoy et al. 2002). Això justifica un dels objectius principals d'aquest treball: estudiar el mecanisme pel qual GAGA, un activador transcripcional, actua com a repressor en el context del promotor del gen que la codifica.

E.1.1. GAGA a sistemes heteròlegs

La repressió de GAGA sobre *Trl* va ser descrita inicialment a les cèl·lules S2, de *Drosophila melanogaster* (Kosoy et al. 2002). Com a primer pas en l'estudi de GAGA com a repressor del seu propi promotor vàrem comprovar si això també es donava a cèl·lules en cultiu d'altres organismes. Per fer-ho, vam escollir un organisme que també tingués *Trl* i un altre que no. Així, vàrem emprar cèl·lules WR DV1 de *Drosophila virilis* perquè, tot i pertànyer al mateix

DISCUSSIÓ

gènere que *Drosophila melanogaster*, les dues espècies estan evolutivament molt separades, uns 60 milions d'anys segons alguns estudis (Beverley et al. 1984). D'altra banda, també vàrem emprar cèl·lules HeLa, d'origen humà. Així, mentre que a *D. virilis* sí que existeix un homòleg per a la proteïna GAGA (Lintermann et al. 1998), a mamífers en general i humans en particular no s'ha descrit cap proteïna amb homologia de seqüència amb GAGA, tot i que alguns estudis senyalen la importància de les seqüències d(GA)_n en la regulació de certs gens a mamífers (Kim et al. 2002; Volpi et al. 2002), i inclús s'ha trobat un factor a ratolí que podria tenir certa relació estructural amb GAGA (Bevilacqua et al. 2000).

GAGA està present a les 12 espècies de *Drosophila* seqüenciades fins ara i, alineant les seves seqüències aminoacídiques, resulta molt significatiu l'elevada conservació que existeix (veure Annexos, apartat H.1.1), sobretot tenint en compte que, per exemple, una proteïna com Cid (*Centromere identifier*), la variant centromèrica de la histona H3 de *Drosophila*, no sembla estar tan conservada a nivell de seqüència (veure Annexos, apartat H.1.2). De fet, la conservació de *Trl* entre espècies de *Drosophila* és tal, que fins i tot es troba conservada l'estructura gènica, mantenint el mateix nombre d'introns i exons, amb mides molt similars (veure Annexos, apartat H.1.4). Tenint això en compte, ens vàrem preguntar si les proteïnes GAGA de dues espècies de *Drosophila* molt allunyades evolutivament, com ho són *D. melanogaster* i *D. virilis*, eren funcionalment equivalents pel que fa al mecanisme de repressió. Els resultats obtinguts per transfecció transitòria mostren inequívocament que les proteïnes GAGA d'ambdues espècies són, pel que fa al mecanisme de repressió de *Trl*, funcionalment equivalents, i encara més, tant el factor GAGA de *D. melanogaster* com el de *D. virilis* reprimeixen el promotor de *Trl* de les dues espècies, i ho fan tant a cèl·lules S2 com a cèl·lules WR DV1. Encara que a primera vista aquest resultat pugui semblar l'esperat, no s'ha d'oblidar que la conservació de seqüència dels promotors de *Trl* de les dues espècies és baixa (veure Annexos, apartat H.1.3), tot i que ambdós tenen molts llocs potencials d'unió per a GAGA.

Pel que fa a les cèl·lules HeLa, el resultat obtingut és molt sorprenent. GAGA, en lloc de reprimir *Trl*, l'activa de manera dosi-dependent. Sembla que l'activació que es dona és de característiques similars a la que s'observa sobre altres promotors normalment activables per GAGA a S2, com és el cas del promotor d'*eve stripe 2*. A més, aquesta activació requereix del domini de transactivació Q i és força ineficient quan manca el domini d'oligomerització POZ, com ja estava descrit pels casos d'activació per GAGA a cèl·lules S2 de mosca (Espinass et al.

1999; Vaquero et al. 2000). Una de les possibles explicacions per aquest fenomen (hipòtesi 1 a la figura E.1) és que a HeLa manqués el factor o factors que converteixen GAGA en repressor de *Trl*, a *Drosophila*. Els antecedents que teníem sobre la repressió de *Trl* per GAGA a *Drosophila* suggerien que havia d'existir algun element de la seqüència de *Trl*, al qual presumiblement s'hi uniria un altre factor que, actuant conjuntament d'alguna manera amb GAGA unida al promotor, donés la repressió de *Trl* observada a *Drosophila*. Tenint aquesta possibilitat en compte, podria ser que aquest factor fos precisament el que manqués a les cèl·lules HeLa, o que no estigués prou conservat com per reconèixer les seqüències d'unió presents al promotor de *Trl*. També és possible que el mecanisme de repressió de *Trl* per GAGA a *Drosophila* impliqui algun altre factor més que no té per què unir-se al DNA necessàriament, és a dir, un cofactor. Així, també podria ser que a HeLa manqués aquest hipotètic cofactor, o no estigués prou conservat.

A banda de tot això, s'ha de recalcar un fet important, i és que l'activitat basal del promotor de *Trl* a HeLa és molt baixa, en comparació amb la que observem a S2. Així, tot i que l'activació sembla molt important tenint en compte els valors relatius que obtenim (veure figura D.3 de l'apartat D.1.2 de Resultats), en realitat, el màxim nivell d'activació es correspon a uns nivells d'activitat luciferasa inferiors als nivells basals de *Trl* a cèl·lules S2 de *Drosophila*. Tot i que els dos sistemes no són comparables en aquests termes, aquest fet suggereix la possibilitat de que en realitat a HeLa el que manqui és l'autèntic activador del promotor de *Trl* (hipòtesi 2 a la figura E.1). Així, segons això, GAGA a S2 reprimiria *Trl* perquè desplaçaria a un activador molt més fort. Al mancar aquest activador a HeLa, el que observem al afegir GAGA és una activació del promotor, ja que GAGA és un activador transcripcional, que també es donaria a S2, però que en aquesta última línia cel·lular es mostraria en forma de repressió degut a que GAGA, amb menor potència activadora, desplaçaria a aquest fort activador de *Trl*. En favor d'això també està el fet que, a cèl·lules S2, la forma de GAGA sense el domini de transactivació Q (GAGA Δ Q) reprimeix *Trl* de forma més eficient que GAGA (Kosoy et al. 2002). Això podria ser degut a que aquesta forma truncada de GAGA conserva la capacitat de desplaçar a l'activador de *Trl* però no la d'aportar una mínima activació al promotor.

DISCUSSIÓ

De totes maneres, no s'ha d'oblidar que HeLa, pel que fa al mecanisme de regulació de *Trl*, és un sistema completament diferent a les cèl·lules de *Drosophila* i que per tant, tots els resultats obtinguts a aquesta línia cel·lular són molt difícils d'interpretar, degut a que en cap cas estem observant una situació natural.

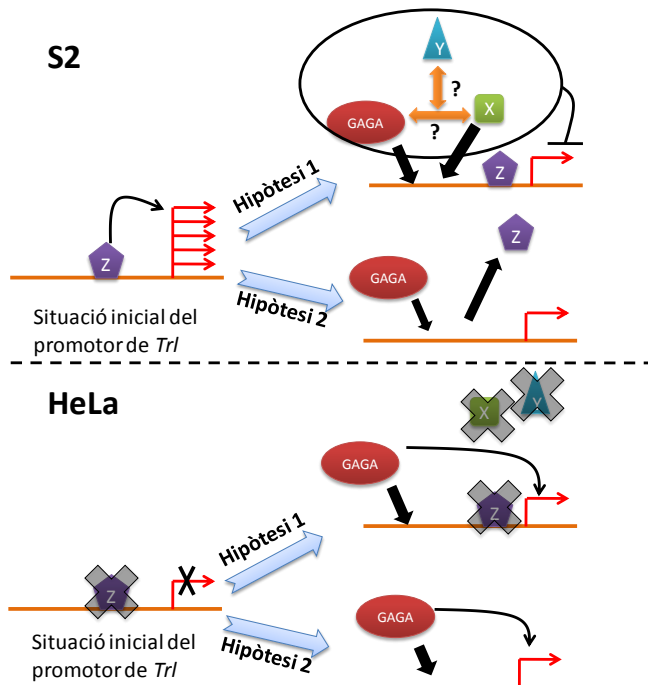


Figura E.1. Dues possibles hipòtesis per explicar l'activació de *Trl* per GAGA a HeLa.

Hipòtesi 1.- A S2, partint d'una situació inicial en la que el promotor de *Trl* té una certa activitat deguda a factors activadors desconeguts (Z), la unió de GAGA al promotor de *Trl*, juntament amb la unió específica al promotor d'algun factor repressor (X), i potser amb la col·laboració d'algun cofactor (Y), portarien a la repressió activa del promotor de *Trl*. A HeLa, la situació inicial seria de pràctica falta d'expressió deguda a la manca d'aquests factors activadors (Z). Així, també mancarien els factors repressors (X) i/o corepressors (Y), fent que la unió de GAGA al promotor resultés en activació.

Hipòtesi 2.- A S2, la unió de GAGA al promotor de *Trl* provocaria el desplaçament d'un o varis factors activadors (Z), resultant en una disminució de l'activitat de *Trl*, deguda a la menor capacitat activadora de GAGA. A HeLa en canvi, al partir d'una situació inicial on manquen aquests factors activadors (Z), al unir-se GAGA al promotor observariem activació d'aquest.

E.1.2. Contribució del promotor de *Trl* a la repressió

El fet que el promotor de *Trl*, a diferència de tots els altres estudiats fins ara, respongués reprimint-se a la presència de GAGA ens va fer pensar que aquesta activitat diferencial hauria de dependre de la seqüència del promotor de *Trl*. La nostra idea inicial era trobar algun element o elements de la seqüència que servissin com a llocs d'unió per a un o varis factors que actuarien juntament amb GAGA per a donar la repressió de *Trl*. Així, la mutació o eliminació d'aquests llocs d'unió hauria d'implicar la pèrdua de la capacitat repressora de GAGA. Per tal de trobar aquestes seqüències del promotor de *Trl* implicades en la repressió, vàrem seguir diverses aproximacions. Primer vàrem realitzar una cerca de factors de transcripció que poguessin estar unint-se al promotor de *Trl*, per mitjà dels programes TFSEARCH i TESS, a més de definir els elements del promotor basal o *core* presents al promotor de *Trl*. Aquesta primera aproximació ha donat resultats força inconsistents, segons el programa utilitzat. Com a mostra de la poca fiabilitat dels algorismes emprats per aquests programes està el fet que TFSEARCH

no troba llocs d'unió per a GAGA a la seqüència del promotor i TESS en troba només quatre quan, a simple vista, ja es poden veure multitud més de seqüències consens GAGAG o CTCTC, a les quals GAGA ha demostrat poder unir-se per *footprinting in vitro* (veure Resultats, apartat D.2.2.1, figura D.5.A, (Kosoy et al. 2002)). A més, en experiments recents de *ChIP on chip*, realitzats tant a cèl·lules S2 (Lee et al. 2008) com a embrions (Schuettengruber et al. 2009), es mostra amb claredat que GAGA està abundantment unida a la seva regió promotora.

Seguidament, vàrem voler comprovar l'estat del promotor de *Trl* *in vivo*, en diverses situacions. Així, vàrem realitzar un estudi del promotor de *Trl* per mitjà de *footprinting in vivo* amb DMS, centrant-nos en les regions FII i FIII, ja que s'havia vist anteriorment que aquestes dues regions, quan es fusionaven al promotor de Citomegalovirus (CMV), tant per separat com juntament (amb major eficàcia), eren suficients per transferir parcialment la capacitat de ser reprimat per GAGA al promotor de CMV, per si sol insensible a GAGA (Kosoy et al. 2002). D'aquesta manera, vam definir tant les seqüències del promotor que estan ocupades a les cèl·lules S2 en estat normal, com les seqüències del promotor que canvien el seu estat d'ocupació per l'augment de la quantitat de GAGA a embrions de mosques transgèniques, comparant els embrions control amb els que sobreexpressen GAGA. Com a resultat d'aquests experiments vàrem obtenir tres seqüències (que hem anomenat -303, -265 i -227 per la posició aproximada que ocupen, respecte al lloc d'inici de la transcripció que anomenem +1) que esdevenien protegides al sobreexpressar GAGA als embrions. D'aquestes tres seqüències, les dues primeres corresponen a llocs GAGA, segons *footprinting in vitro* (Kosoy et al. 2002). A més, la primera està molt propera a una hipersensibilitat, i les altres dues coincideixen amb dues zones protegides de les que vam obtenir als *footprintings in vivo* amb DMS de cèl·lules S2 normals, el que podria venir a indicar que la sobreocupació d'aquests llocs pot portar a la repressió de *Trl*. Per això vàrem mutar aquestes seqüències, una a una i també mutant a l'hora -303 i -265 (perquè al ser dos llocs potencials d'unió per a GAGA, la falta d'un podria ser que es compensés amb l'altre), per comprovar si continuaven sent reprimibles per GAGA utilitzant el nostre sistema de transfecció transitòria a S2. Malauradament, cap d'aquestes mutacions del promotor de *Trl* eliminava la capacitat de ser reprimat per GAGA, indicant que cap d'aquestes seqüències era essencial per a que es donés la repressió. És destacable que -227, que no coincideix ni per seqüència ni per *footprinting in vitro* amb cap lloc d'unió per a GAGA, pot ser una seqüència implicada en l'activitat d'aquest promotor, ja que la seva mutació fa baixar

DISCUSSIÓ

l'activitat inicial al voltant del 50%. També és destacable que cap d'aquestes seqüències ha estat reconeguda pels programes informàtics com a llocs d'unió per a cap factor de transcripció, si bé -303 està força propera a un possible lloc d'unió pel factor TTK69.

Una de les possibles hipòtesis per explicar la repressió de *Trl* per GAGA és que, atenent a la gran quantitat de llocs al promotor als que GAGA demostra poder unir-se per *footprinting in vitro* (veure Resultats, apartat D.2.2.1, figura D.5.A, (Kosoy et al. 2002)), podria ser que la repressió vingués del fet que al sobreexpressar GAGA, l'excés de proteïna ocupés massivament els llocs GAGA al promotor interferint amb la maquinària de transcripció. Tanmateix, els resultats dels nostres *footprintings in vivo* no mostren aquesta massiva ocupació del promotor, en contra d'aquesta hipòtesi. D'altra banda, els resultats obtinguts a HeLa, on GAGA activa la transcripció de *Trl*, tampoc li donen suport.

Paral·lelament als experiments de *footprinting in vivo* amb DMS, també vàrem realitzar tot un seguit de delecions del promotor de *Trl*, així com algunes mutacions seguint diversos criteris. Ja anteriorment, per mitjà de delecions successives de la regió promotora, s'havia arribat a acotar la regió mínima del promotor de *Trl*, que conserva tota la seva activitat i que a més continua sent reprimible per GAGA, en el que es va anomenar fragment p2b, que va de -345 a +737, prenent com a +1 l'inici de transcripció més distal a la regió codificant (Tesi doctoral d'Ana Kosoy, 2002; (Kosoy et al. 2002)). Partint d'aquests resultats, vàrem voler acotar encara més la regió mínima realitzant tot un seguit de delecions, pels dos extrems de la seqüència. Vàrem testar, per transfecció transitòria a cèl·lules S2, si aquestes formes delecionades del promotor continuaven sent reprimibles per GAGA. Els resultats indiquen que totes les delecions que vàrem realitzar continuen sent reprimibles per GAGA, a excepció de dues delecions severes, els fragments -345/-33 i -345/+14. La primera perd tot el 5' UTR, perdent així els tres inicis de transcripció caracteritzats anteriorment (Kosoy et al. 2002). La segona només conté 14 pb de 5' UTR, i només un inici de transcripció (el +1). Aquestes dues formes delecionades, si bé perden la capacitat de ser reprimides per GAGA, també perden de manera molt important la seva activitat basal o inicial (al no afegir GAGA), quedant aquesta molt per sota de la mínima observada quan GAGA reprimeix, que és d'un 10% de l'activitat inicial (per exemple, veure Resultats, figura D.2.A). Així doncs, podria ser que simplement estem davant de dos fragments del promotor no funcionals que, tot i conservant les regions importants per a la repressió, aquesta no pogués ser observada degut a la falta d'activitat del promotor. De totes maneres,

s'ha de destacar que la següent deleció analitzada, que conté una regió més gran del 5'UTR (el fragment -345/+71), sí que pot ser reprimida per GAGA, fins a uns nivells d'activitat semblants als de -345/-33 i -345/+14. Aquest fragment només es diferencia de -345/+14 en contenir 57 pb més de la regió 5' UTR. Així, podria ser que en aquesta regió de 57 pb existís algun element de seqüència important pel mecanisme de repressió. Anteriorment s'havia descrit, per *footprinting in vitro* amb DNasa I, una àmplia regió d'unió per a GAGA a aquesta seqüència, entorn a la posició +41 (veure Resultats, apartat D.2.2.1, figura D.5.A, (Kosoy et al. 2002)). Per això, vàrem mutar aquesta seqüència de forma que no contingués cap lloc per a la unió de GAGA, tant en el context del promotor p2b sencer (obtenint la construcció p2b+41), com al fragment -345/+71 (obtenint la construcció 345/+71(+41)), per tal de determinar si aquest lloc és essencial o només contribueix a la repressió. Sorprenentment, al provar aquestes formes mutades del promotor de *Trl* en el nostre sistema de transfecció transitòria a cèl·lules S2, cap de les dues perd la capacitat de ser reprimida, conservant més o menys intacta la seva activitat inicial, respecte al fragment equivalent sense mutar a +41. Aquests resultats estan en concordança amb la idea que la pèrdua de la capacitat de ser reprimides per GAGA, de les dues formes delecionades del promotor -345/-33 i -345/+14, en realitat és deguda a una falta de funcionalitat.

Per tant, per mitjà de deleccions successives no vàrem ser capaços de trobar cap regió concreta necessària per a la repressió per GAGA. El que sí vàrem trobar són regions del promotor implicades en la seva activitat, ja que amb la simple deleció dels 47 pb que diferencien p2b (-345/+737) del fragment -298/+737 ja observem una pèrdua d'aproximadament el 50% de l'activitat inicial. A més, també és destacable la probable existència d'elements importants del promotor a la regió 5' UTR, ja que les deleccions d'aquesta regió tenen un efecte molt evident sobre l'activitat del promotor.

Finalment, vàrem fer un parell de mutacions més del promotor de *Trl* basant-nos en la presència de dues seqüències molt semblants situades al voltant de -214 i -120. D'aquestes seqüències, la que es troba a -120 està formada per dos motius consens M7 i M8 situats de forma solapada. D'aquests motius, descrits com a motius del promotor basal (Ohler et al. 2002), M7 s'ha vist que està present en un 15% de promotors on s'hi uneix TRF2, sent el segon motiu (darrera del motiu DRE) més freqüent en aquest tipus de promotors (Isogai et al. 2007a). Això resultava interessant perquè obria la possibilitat de que el promotor de *Trl* estigués format

DISCUSSIÓ

per diversos promotors, on s'hi unrien TBP o TRF2, i per tant podrien ser regulats de forma diferent i independent. No obstant, cap d'aquestes seqüències sembla ser imprescindible per a que es doni la repressió de *Trl* per GAGA, ja que al mutar-les i provar-les en el nostre sistema de transfecció transitòria a S2 no perden la capacitat de ser reprimides. Aquestes seqüències semblen, però, estar implicades en l'activació de *Trl*, ja que la seva mutació redueix un 50% l'activitat inicial que presenta p2b.

Agafant totes les dades en conjunt (veure resum esquemàtic a la figura E.2), la impressió que ens queda és que el promotor de *Trl* podria contenir múltiples regions implicades en la seva activitat, tant a la regió *upstream* de l'inici de transcripció més distal (+1), com al 5' UTR, i que mentre conservi part d'aquesta activitat podrà ser reprimat per GAGA. Aquesta impressió ve del fet que la majoria de delecions i/o mutacions provades afecten l'activitat del promotor i, només eliminant aquesta activitat, hem aconseguit afectar la repressió per GAGA, indicant que la capacitat de repressió va íntimament lligada a la capacitat d'expressió.

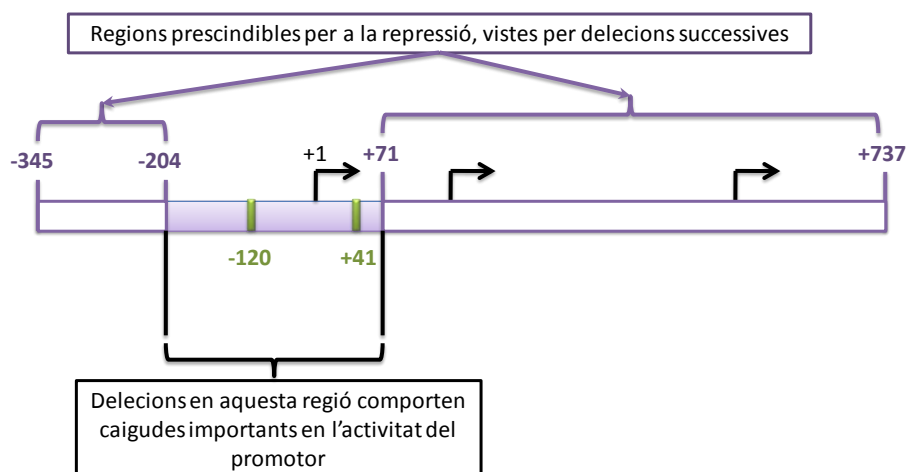


Figura E.2. Resum esquemàtic del promotor de *Trl* (fragment p2b), mostrant les dades que hem obtingut de tots els experiments amb les formes delecionades i/o mutades d'aquest. Les regions de -345 a -204 i de +71 a +737 són prescindibles en repressió, ja que eliminar-les no aboleix la repressió per GAGA. Delecions a la regió entre -204 i +71 comporten una caiguda molt important en l'activitat del promotor, el que no permet estudiar, per mitjà de delecions, la seva implicació en el mecanisme de repressió. Mutacions a certs nucleòtids a -120 i a +41 (en verd) no alteren tampoc la capacitat del promotor de ser reprimat per GAGA. Amb fletxes negres s'indiquen els orígens de transcripció.

Tenint en compte la gran quantitat de possibles llocs GAGA a la seqüència del promotor, i que sembla clar que GAGA ha d'unir-se al promotor per poder reprimir (Kosoy et al. 2002), podria ser que l'eliminació d'uns llocs quedés compensada per la presència dels altres. Al mateix temps, l'hipotètic factor o factors que han d'unir-se al promotor per portar a la repressió de *Trl* també haurien de tenir múltiples llocs d'unió possibles i l'eliminació d'aquests

també hauria de poder compensar-se amb la resta. Tanmateix, la possibilitat de que l'excés de GAGA estigués "folrant" el promotor i d'aquesta manera dificultant l'acció de la maquinària de transcripció ja ha estat descartada.

E.1.3. Contribució de les diferents regions de GAGA a la repressió

Estudis anteriors (Kosoy et al. 2002) ja havien definit que tant el domini POZ com el domini Q de GAGA són prescindibles per a la repressió de *Trl*. A més, una mutació puntual al DBD, que aboleix la capacitat de GAGA per unir-se al DNA, també aboleix completament la repressió, indicant que la unió de GAGA al promotor de *Trl* és imprescindible per a la repressió. Així doncs, a fi d'esclarir la implicació de la regió X i del DBD en el mecanisme de repressió de *Trl*, es varen preparar tot un seguit de construccions per sobreexpressar diverses formes truncades de la proteïna GAGA i comprovar la seva activitat repressora de *Trl*, a cèl·lules S2.

De totes les construccions provades (veure Resultats, apartat D.3, figura D.13), només l'anomenada GAGA₅₁₉ΔX2 conserva intacta la capacitat de reprimir *Trl*, així com d'activar el promotor *eve stripe 2*, tot i que no de forma tan eficient com GAGA₅₁₉. La resta de formes truncades de GAGA podrien ser candidates a haver perdut una regió imprescindible per a la repressió de *Trl*, però diversos problemes d'expressió i de localització subcel·lular dificulten la interpretació dels resultats. GAGA₅₁₉ΔX1 i GAGA-DBD sembla ser que, o no s'expressen correctament, o es degraden molt ràpidament, ja que no les detectem. La fusió del DBD de GAGA amb la proteïna EGFP (GAGA-DBD-EGFP) sí s'expressa, però es localitza tant nuclear com citoplasmàticament, fet que podria explicar que no sigui capaç de reprimir. La resta de construccions (GAGA₅₁₉ΔPOZ₂₄₅, ECFP-Nuc-GAGA-DBD i Lac I-GAGA-DBD) també les detectem, però per immunolocalització a cèl·lules S2 veiem com s'acumulen de forma anormal al nucli, el que també podria estar afectant a la seva activitat.

Aquests resultats, tot i no resultar conclouents per a definir regions importants en la repressió de *Trl*, sí que aporten altres dades interessants. La construcció GAGA-DBD-EGFP, perd la localització estrictament nuclear que mostra GAGA, indicant que al DBD de GAGA no existeixen elements que serveixin de senyal de localització nuclear, malgrat que inclou les 3 regions bàsiques (BR1-BR3). D'altra banda, les fusions del DBD a ECFP amb tres NLS (ECFP-Nuc-

DISCUSSIÓ

GAGA-DBD) i a Lac I (Lac I-GAGA-DBD) sí que el dirigeixen al nucli, però la seva distribució allà és força estranya, indicant que la correcta distribució intranuclear no depèn únicament de la seva capacitat d'unir-se al DNA a través del DBD. Sorprenentment, la construcció GAGA₅₁₉ΔPOZ₂₄₅, tot i no tenir l'NLS, també és nuclear. Això podria deure's a que altres regions de la proteïna poden estar implicades en la localització nuclear de GAGA, o que almenys hi poden contribuir, probablement a través de la interacció amb altres proteïnes que la cotransporten al nucli. Alternativament, potser la seva entrada al nucli és per difusió i, un cop allà, és retinguda a través de la seva interacció amb el DNA, però acumulant-se de forma anormal. En contra d'aquesta possibilitat està el fet que aquest fenomen seria d'esperar que també succeís amb la fusió del DBD a EGFP, i no ho hem observat. En qualsevol cas, l'anormal acumulació del mutant GAGA₅₁₉ΔPOZ₂₄₅ suggereix que, o bé el propi domini POZ, o la part més N-terminal de la regió X, poden col·laborar en la correcta unió de GAGA als seus llocs dins del nucli.

Experiments de fusió de cada domini per separat a EGFP han indicat que només la regió X és capaç de dirigir EGFP al nucli de forma selectiva (Laura Regué i Jordi Bernués, no publicat). Per tant, malgrat que el domini Q s'ha vist que pot actuar com a domini d'interacció proteïna-proteïna (Wilkins et al. 1999), la possibilitat de que aquest domini, ja sigui per mitjà de la interacció amb el domini Q de la GAGA endògena, o per mitjà de la interacció amb alguna altra proteïna nuclear, sigui el responsable de portar GAGA₅₁₉ΔPOZ₂₄₅ al nucli és força improbable.

En conjunt, els resultats de localització i activitat dels diversos mutants (veure resum, a la figura E.3) suggereixen la possibilitat que la zona (a la regió X) entre els aminoàcids 122 i 222 (que inclouen l'NLS) sigui essencial per a la correcta disposició de GAGA dins del nucli i per tant, també per a la seva activitat, tant en la repressió de *Trl* com en l'activació d'altres promotors. Pel que està descrit, els aminoàcids del 122 al 222 de GAGA, podrien estar implicats en la interacció de GAGA amb el factor dSAP18, una proteïna que estaria associada al complex corepressor, amb activitat desacetilasa d'histones, Sin3-HDAC (Espinás et al. 2000); i amb els complexos remodeladors de cromatina NURF i FACT, per les seves subunitats NURF301 (Xiao et al. 2001) i dSSRP1 (Shimojima et al. 2003), respectivament. Així, encara que no és evident, és possible que l'associació de GAGA amb algun d'aquests factors sigui crucial per a la seva correcta disposició dins del nucli, però també és possible que hi hagi altres interaccions encara no descrites que expliquin la seva correcta distribució.

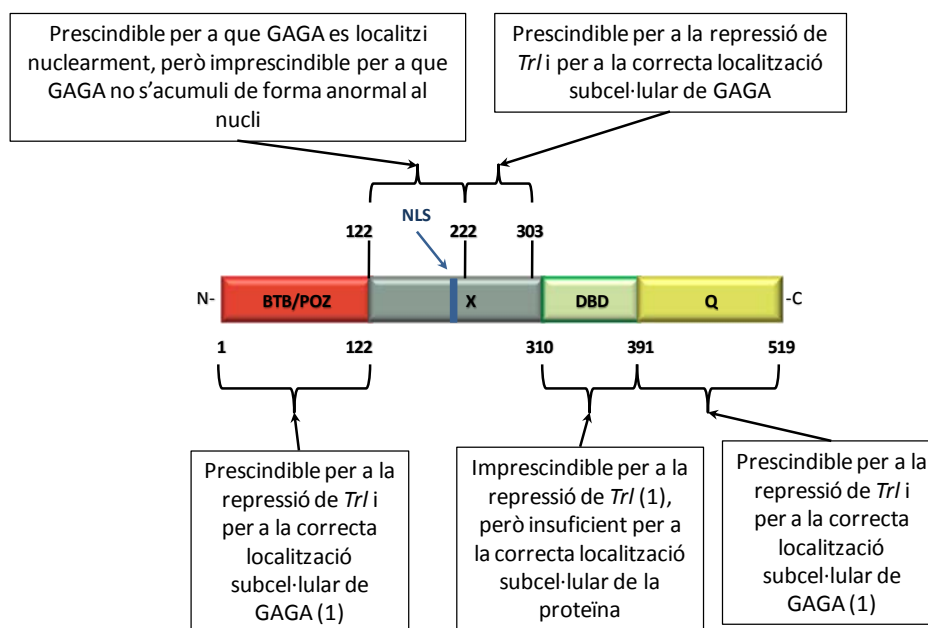


Figura E.3. Esquema de la proteïna GAGA₅₁₉ i els seus dominis, on s'indica la informació de que disposem, respecte al paper que juga cada regió en el mecanisme de repressió de *Trl* i la localització subcel·lular de la proteïna. (1)(Kosoy et al. 2002).

E.1.4. Contribució de diversos factors al mecanisme de repressió de *Trl*

Una de les possibles hipòtesis per explicar el mecanisme de repressió de *Trl* per GAGA (veure figura E.1, primera hipòtesi, cas d'S2; i figura E.4 per a més detall), invoca la presència d'un factor o factors repressors (X a la figura E.4), d'unió específica al promotor de *Trl*, que actuarien juntament amb GAGA i potser amb algun factor o factors que no s'uneixin específicament al DNA (corepressor Y a la figura E.4), per a dur a terme la repressió de *Trl*. Aquesta hipòtesi, es basa en la premissa que a la seqüència del promotor de *Trl* ha d'existir un o diversos elements que el diferenciïn de la resta de promotors on GAGA s'hi uneix i que determinin la repressió. Aquests elements s'esperava que fossin seqüències d'unió per a factors de transcripció, amb funció repressora. Donat que la estratègia de buscar directament aquesta seqüència determinant per a la repressió no va donar pistes sobre quin podria ser aquest factor repressor X, en el cas de que efectivament existís (veure apartat E.1.2), vàrem provar directament dos dels factors de transcripció que ja sabem que interaccionen amb GAGA i que a més tenen llocs d'unió possibles a la seqüència de *Trl*: Psq i TTK69. Totes tres proteïnes, GAGA, Psq i TTK69 tenen dominis POZ, i és mitjançant aquests per on es suposa que interaccionen majoritàriament (Bardwell et al. 1994; Pagans et al. 2002; Schwendemmann et al. 2002).

DISCUSSIÓ

Encara que el domini POZ de GAGA és completament prescindible per a que es doni la repressió, altres dades ens van fer considerar la possible contribució de Psq i TTK69 a la repressió de *Trl*, com ara que tots dos són essencialment repressors (Brown et al. 1993; Huang et al. 2002), que Psq reconeix les mateixes seqüències al DNA que GAGA (Schwendemann et al. 2002) i que per a TTK69 vàrem trobar, per mitjà dels programes informàtics, que podria unir-se a tres llocs al promotor de *Trl* (veure Resultats, apartat D.2.1, figura D.4), el primer situat entorn a -600, el segon entorn a -303 i el tercer entorn a -32. És de destacar que aquest tercer lloc està situat a una zona de les que, per deleccions del promotor, hem trobat que eren vitals per a l'activitat (veure apartat E.1.2, figura E.2). Tanmateix, els resultats obtinguts amb aquests dos factors de transcripció indiquen que ni TTK69 ni Psq estan implicats en el mecanisme de repressió de *Trl* per GAGA.

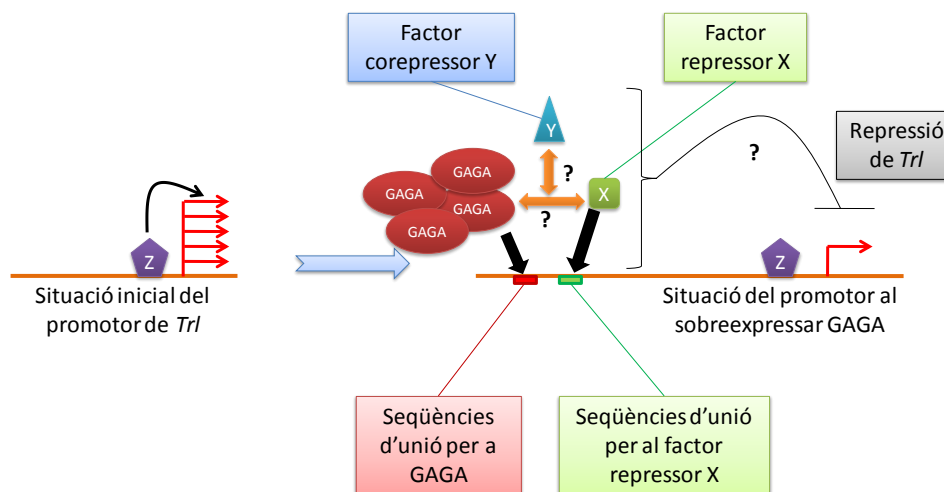


Figura E.4. Una de les hipòtesis possibles per explicar el mecanisme de repressió de *Trl* per GAGA. El promotor de *Trl*, en absència de GAGA, presenta una certa activitat, deguda probablement a factors activadors desconeguts (Z). Al afegir GAGA, aquesta s'hi uneix a les seves seqüències presents al promotor i juntament amb un factor repressor desconegut (X), el qual també reconeix seqüències específiques presents al promotor i amb la possible ajuda d'un o diversos factors corepressors (Y), que podrien estar interaccionant amb GAGA, amb el factor repressor (X) o amb tots dos, és disminueix l'activitat transcripcional de *Trl*, afectant a la maquinària basal, al conjunt d'activadors de *Trl* o per algun altre mecanisme.

A banda d'aquest hipotètic factor repressor d'unió específica al promotor de *Trl*, la hipòtesi presentada a la figura E.4 admet també l'existència d'un o diversos factors corepressors (anomenat Y a la figura) que, sense unir-se específicament al DNA, estiguin implicats en el mecanisme de repressió de *Trl* per GAGA i ser igualment necessaris. Aquest hipotètic corepressor podria interaccionar amb GAGA, amb el factor repressor X, o amb tots dos. Vàrem decidir estudiar dos dels corepressors més generals i coneguts a *Drosophila*: Groucho, freqüentment implicat en el mecanisme de repressió de llarg abast o *long-range*; i

dCtBP, freqüentment implicat en el mecanisme de repressió de curt abast o *short-range* (Courey et al. 2001).

Els resultats obtinguts amb Groucho suggereixen que aquesta proteïna no està implicada en el mecanisme de repressió de *Trl*. Per contra, els resultats obtinguts amb dCtBP sí apunten a una implicació d'aquest corepressor en la repressió de *Trl*. La sobreexpressió de dCtBP a S2, per transfecció transitòria, reprimeix el promotor de *Trl*, tant si es cotransfecta el vector d'expressió de GAGA, com sí no. Així, sembla clar que dCtBP està implicat en la repressió de *Trl*, però no queda clar si està implicat en el mecanisme de repressió dependent de GAGA, o si ho fa per altres vies. En aquest sentit, altres experiments realitzats per intentar esclarir aquest punt van aportar algunes evidències a favor, i d'altres aparentment en contra, de l'acció conjunta GAGA-dCtBP. A favor està el fet que per *GST-Pulldown* veiem que GAGA i dCtBP poden interaccionar directament. D'altra banda, no veiem que la sobreexpressió conjunta de GAGA i dCtBP, per transfecció transitòria a cèl·lules HeLa, resulti en la repressió de *Trl* en aquestes cèl·lules, com esperaríem si dCtBP fos l'element que diferencia a les cèl·lules HeLa de les S2 en el mecanisme d'acció de GAGA sobre *Trl*. Tanmateix, aquesta és una feble evidència en contra d'una acció conjunta GAGA-dCtBP a S2, ja que els sistemes de transcripció poden haver divergit suficientment com per a no poder ser complementats d'una manera tan simple, i podria ser un altre factor o un conjunt d'aquests el que diferenciés HeLa d'S2. Una altra evidència en contra és el fet que la depleció de dCtBP amb RNAi no té cap efecte ni en l'activitat de *Trl*, ni en la seva repressió per GAGA. Si dCtBP fos el corepressor implicat, esperaríem que la seva eliminació primer, augmentés l'activitat basal del promotor de *Trl* que, recordem, ja està en certa mesura reprimida per GAGA a les cèl·lules S2 (Kosoy et al. 2002), i segon dificultés en alguna mesura l'acció repressora de la sobreexpressió de GAGA. De totes maneres, s'ha de fer notar que l'RNAi podria no haver sigut prou eficient eliminant dCtBP, ja que, tot i que podem comprovar indirectament l'efectivitat del RNAi, per com elimina el dCtBP fusionat a l'epítot flag que sobreexpressem, no hem pogut comprovar la reducció de la forma endògena de dCtBP. Amb aquesta prevenció, podem dir que GAGA i dCtBP, tot i poder interaccionar, pel que fa a la repressió de *Trl* probablement actuen per mecanismes diferents.

Donat que s'havia descrit que dCtBP té regions d'homologia amb algunes deshidrogenases NAD-dependents, s'ha proposat que el metabolisme del NAD⁺ pot tenir algun efecte sobre l'activitat de dCtBP (Schaeper et al. 1995). En aquesta direcció, vàrem assajar com

DISCUSSIÓ

podria afectar al mecanisme de repressió de *Trl* l'addició al medi de TSA (un inhibidor general de desacetilases no NAD-depenents) i de nicotinamida (un inhibidor del metabolisme del NAD⁺, que pot inhibir les desacetilases NAD-depenents). En els nostres experiments, l'addició de nicotinamida no afecta ni a l'activitat de dCtBP sobre *Trl* ni a la repressió de *Trl* per GAGA, suggerint que el metabolisme del NAD⁺ no està implicat ni en l'acció de dCtBP ni en el mecanisme de repressió de *Trl* per GAGA. En concordança amb això, Schaeper i col·laboradors ja indiquen que la forma humana de dCtBP no sembla tenir activitat deshidrogenasa, ni unir NAD⁺ (Schaeper et al. 1995). El TSA és un inhibidor de molts HDACs (complexos desacetilases d'histones) responsables de l'acció repressora de molts corepressors. Tot i que dCtBP no s'ha descrit que funcioni a través de HDACs, vàrem voler comprovar si altres corepressors sensibles a TSA podrien estar actuant en la repressió de *Trl*. Sorprenentment, l'addició de TSA afecta l'acció repressora de dCtBP, però no sembla afectar la capacitat repressora de GAGA. Això va a favor de la idea de que GAGA i dCtBP reprimeixen *Trl* per mecanismes independents. Tot i amb això, s'ha de tenir en compte que no s'havia descrit anteriorment que el TSA afectés la capacitat corepressora de dCtBP, així que no es pot descartar la possibilitat de que estiguem veient un efecte indirecte de l'addició de TSA al medi.

E.1.5. Repressió de *Trl* per GAGA, a la mosca

Tota la informació *in vivo* de que disposàvem, a l'inici d'aquesta Tesi Doctoral, sobre la repressió de *Trl* per GAGA provenia d'experiments a la línia cel·lular S2 (Kosoy et al. 2002). Així, a fi de demostrar que aquesta inusual activitat de GAGA sobre el seu propi promotor no és deguda a una particularitat d'aquesta línia cel·lular i constitueix un mecanisme genuí de regulació de *Trl*, vàrem comprovar que la repressió de *Trl* per GAGA també es donés a mosca i que aquesta repressió requerís de les mateixes regions de GAGA, fet indicatiu de que el mecanisme funciona de manera anàloga a cèl·lules S2 i a mosca.

Al nostre grup, vàrem demostrar que la sobreexpressió de GAGA₅₁₉ a mosques transgèniques era capaç de reprimir l'expressió d'un transgèn consistent en el promotor de *Trl* (tant en la seva versió més llarga com en la mínima) fusionat al cDNA de GFP com a *reporter*. Al mateix temps, també vàrem demostrar que l'expressió a mosques transgèniques d'un RNAi contra GAGA (no específic d'isoforma) augmentava l'expressió del transgèn consistent en el promotor de *Trl* fusionat al cDNA de GFP, indicant que, al igual que a cèl·lules S2, a la mosca

també existeix un cert grau de repressió constitutiva d'aquest promotor donat pel component endogen de GAGA (Bernues et al. 2007).

Així doncs, ja dins del treball corresponent a aquesta Tesi Doctoral, ens vàrem proposar veure quin seria l'efecte de la sobreexpressió de GAGA₅₈₁, GAGAΔQ i GAGA₅₁₉DBDKO sobre aquest *reporter* consistent en el promotor de *Trl* fusionat al cDNA de GFP, a mosques transgèniques. Tots els resultats obtinguts van reproduir el que observem a cèl·lules S2, és a dir, GAGA₅₈₁ és perfectament capaç de reprimir *Trl*, al igual que GAGAΔQ, indicant que el domini Q de GAGA no està implicat en la repressió de *Trl*. Al mateix temps, la mutació al DBD de GAGA que impedeix la seva unió al DNA, present a la construcció GAGA₅₁₉DBDKO, provoca que es perdi completament la funció repressora de *Trl*, al igual que succeeix a S2.

A *Drosophila melanogaster*, l'mRNA de *Trl* és dipositat per la mare als oòcits. L'expressió zigòtica comença unes 4h després de ser dipositats els ous, té un pic sobre les 8h i decreix remarcablement, arribant a un estat basal d'expressió després de les 16-20h (Bhat et al. 1996). Sorprenentment, la proteïna GAGA és present a uns nivells relativament similars des d'embrions de 0 a 3h fins a adults, mostrant només una reducció entre les fases de larva 2 i larva 3 tardana (Benyajati et al. 1997). Així, *Trl* pertany al grup de gens que mantenen un nivell constant d'expressió al llarg de tot el desenvolupament. Per tant, és probable que el mecanisme de repressió de *Trl* per GAGA jugui un important paper en mantenir un nivell constant i controlat de la proteïna al llarg de tota la vida de la mosca.

E.1.6. Efectes fenotípics de la sobreexpressió de GAGA a mosques transgèniques

Generar les línies de mosques transgèniques que permeten la sobreexpressió dirigida de GAGA₅₁₉, GAGA₅₈₁, RNAi-GAGA, GAGAΔQ i GAGA₅₁₉DBDKO ens ha permès descriure els efectes fenotípics observables a mosca, derivats de la sobreexpressió d'aquestes formes de GAGA. En aquest sentit, el treball realitzat per a aquesta Tesi Doctoral es centra en la isoforma GAGA₅₈₁, ja que els resultats obtinguts amb les altres formes de GAGA ja han estat descrits (Bernues et al. 2007).

Així, comparant els efectes que produeix la sobreexpressió de GAGA₅₁₉ amb els que produeix la sobreexpressió de GAGA₅₈₁, és de destacar que els efectes són similars però no

DISCUSSIÓ

idèntics. El que s'observa és que, en general, la sobreexpressió de GAGA₅₈₁ és millor tolerada per la mosca, i els fenotips que dona a les ales dels individus adults estan en la mateixa línia, encara que són menys severs, que els que obtenim amb la sobreexpressió de GAGA₅₁₉. Altres resultats obtinguts recentment al nostre laboratori també van en aquesta direcció. Així, s'han aconseguit mosques transgèniques que sobreexpressen GAGA₅₁₉ i GAGA₅₈₁ a la vegada, sota el control del mateix *driverGal4*. Els resultats obtinguts, pel que fa a letalitat i efectes fenotípics, tot i que preliminars, semblen estar a mig camí entre la forta letalitat de la sobreexpressió de GAGA₅₁₉ i la no tan acusada letalitat de GAGA₅₈₁. Tenint en compte que la sobreexpressió simultània de les dues isoformes de GAGA depèn de la mateixa quantitat d'activador Gal4 produït pel *driver*, és d'esperar una quantitat de proteïna total induïda similar a la que es produeix quan només es sobreexpressa una única isoforma, però repartida entre les dues isoformes. Per tant, aquests resultats també estan a favor de la idea de que les dues proteïnes realitzen funcions solapants, essent GAGA₅₈₁ menys activa. Això mateix ja havia estat descrit anteriorment a certs estudis, on els autors proposen que la isoforma GAGA₅₁₉ seria la més activa per la seva major capacitat de rescatar alguns fenotips dels mutants de *Trl* (Greenberg et al. 2001).

De l'anàlisi dels fenotips resultants de la sobreexpressió de les construccions RNAi-GAGAi, GAGAΔQ i GAGA₅₁₉DBDKO (Bernues et al. 2007), és de destacar que la construcció GAGA₅₁₉DBDKO, consistent en la forma de GAGA₅₁₉ amb una mutació puntual al DBD que impedeix la seva unió al DNA (Kosoy et al. 2002), no presenta cap tipus de letalitat ni defecte fenotípic, indicant que GAGA, per a dur a terme les seves funcions a la mosca, necessita unir-se al DNA. També és sorprenent el fet que la sobreexpressió de GAGAΔQ, que en principi seria una forma de GAGA incapaç de dur a terme les funcions transactivadores de la transcripció que requereixen del domini Q, però que seria perfectament capaç de, per exemple, reprimir *Trl*, presenti una letalitat molt similar a la observada per a GAGA₅₁₉, amb la majoria de *drivers* emprats. Això podria suggerir que, a la mosca, les funcions de GAGA que no requereixen del domini transactivador Q són de vital importància. En aquesta línia, certs estudis indiquen que el domini Q de GAGA seria prescindible *in vitro* per a funcions com ara la remodelació de cromatina (Agianian et al. 1999). Així, importants funcions com aquesta podrien ser les que portessin a la letalitat tant de la sobreexpressió de GAGA₅₁₉ com de GAGAΔQ. Una explicació alternativa seria que GAGAΔQ es pugui comportar com a un dominant negatiu que reduís

dràsticament l'expressió de gens normalment activats per GAGA i essencials per a la mosca. De tota manera, aquesta similar letalitat per la sobreexpressió de GAGA₅₁₉ i GAGAΔQ no s'observa amb tots els *drivers* emprats, i fins i tot desapareix completament al sobreexpressar GAGAΔQ amb el *driver* prdGal4. Això suggereix que a cada tipus cel·lular podrien ser més importants unes o altres funcions de GAGA, el que pot voler dir que GAGA, tot i ser d'expressió ubiqua (Soeller et al. 1993), i com a proteïna polifacètica que és, podria dur a terme funcions diferents i concretes a cada tipus cel·lular. Finalment, cal destacar que la sobreexpressió de l'RNAi contra GAGA a mosques transgèniques, amb qualsevol dels *drivers* emprats, tret dels d'expressió ubiqua ActGal4 i armadilloGal4 (Bernues et al. 2007), és en general més ben tolerada, provocant molta menys letalitat que la sobreexpressió de GAGA₅₁₉. Això suggereix la possibilitat que, a la mosca, existeixi un nivell màxim tolerable de GAGA que no ha de ser excedit, per sobre del qual els efectes negatius són molt importants. Aquesta idea és molt interessant ja que podria justificar de forma molt clara l'existència d'un mecanisme d'autocontrol de la dosi de GAGA, per mitjà de la repressió de *Trl* pel seu propi producte, GAGA.

E.1.7. Repressió de GAGA sobre la còpia endògena del gen *Trl*

Tots els experiments fets anteriorment on es descriu la repressió de *Trl* per GAGA estan fets, tant a cèl·lules S2 com a mosques transgèniques, sobre còpies ectòpiques del promotor de *Trl* fusionades a gens *reporter* com la luciferasa, en el cas d'S2, i GFP en el cas de les mosques transgèniques. Així, poc es coneixia de la resposta del gen *Trl* situat en el seu context genòmic natural, a la sobreexpressió de GAGA. Per aquesta raó, mitjançant RT-PCR i *Real Time* RT-PCR, tant a cèl·lules S2 com a discos imaginals d'ala de larves de 3r estadi, vàrem comprovar quin és l'efecte de la sobreexpressió de GAGA₅₁₉ sobre el nivell de transcrits provinents de la còpia endògena de *Trl*. Donat que l'expressió de *Trl* genera diversos transcrits (Soeller et al. 1993; Benyajati et al. 1997; Kosoy et al. 2002), els encebadors emprats en els nostres experiments d'RT-PCR i de *Real Time* RT-PCR estan dissenyats de tal manera que amplifiquin indistintament tots els transcrits possibles de la còpia endògena de *Trl*, sense amplificar els transcrits provinents de les còpies ectòpiques inserides tant a les cèl·lules S2 establement transfectades, com a les mosques transgèniques.

DISCUSSIÓ

Els resultats obtinguts indiquen que, tant a cèl·lules S2, com a discos d'ala de larves de 3r estadi, la sobreexpressió de GAGA reprimeix l'expressió de la còpia endògena del gen *Trl*. Això confirma que, ni la posició cromosòmica, ni altres possibles elements reguladors de *Trl*, que podrien estar presents tan a 3' de la regió codificant, com a la pròpia regió codificant, en regions intròniques, o fins i tot molt allunyats de la posició del gen, modifiquen l'activitat repressora de GAGA sobre el gen que la codifica. A la vegada, aquests resultats aporten validesa als obtinguts amb la regió utilitzada del promotor de *Trl* fusionada als gens *reporter*, ja que aquesta conserva tots els elements necessaris per a que es doni el mecanisme de repressió de *Trl* per GAGA.

Una altra dada important que es pot deduir d'aquests resultats sobre la còpia endògena de *Trl* és que la repressió donada per la sobreexpressió de GAGA és probablement a nivell transcripcional, ja que per mitjà de les RT-PCRs i *Real Time* RT-PCRs observem la baixada en els nivells d'mRNA de *Trl*. Això descarta possibles mecanismes de regulació a nivell post-transcripcional, com ara afectar a la traducció, per exemple, bloquejant el transport nucli-citoplasma dels mRNAs de *Trl*, possiblement per unió directa de la proteïna GAGA a les seves seqüències presents a la regió 5'UTR. No hem d'oblidar però que no es pot descartar completament que la sobreexpressió de GAGA redueixi els nivells d'mRNA independentment de la transcripció, per exemple, afectant a la seva estabilitat i promovent la seva degradació. De totes maneres, no observar repressió de *Trl* per GAGA a cèl·lules HeLa indica que, sí es donés aquesta repressió post-transcripcional, no seria simplement per la unió de GAGA a la regió 5'UTR.

Finalment, també hem de fer referència a un fet important i és que en recents experiments de *CHIP on chip* a cèl·lules S2 (Lee et al. 2008) i a embrions (Schuettengruber et al. 2009), GAGA apareix abundantment unida a la regió promotora i a l'inici de la regió codificant de *Trl* (incloent els primers introns). Que estigui disposada sobre el promotor és en certa manera esperable i va en concordança amb els nostres resultats que indiquen que el promotor de *Trl* ja està en certa mesura reprimat pel component endogen de GAGA. El sorprenent és que també es disposi sobre l'inici de la regió gènica. No obstant això, dels nostres experiments es pot deduir que aquesta acumulació de GAGA a la regió gènica no és determinant pel mecanisme de repressió, ja que les nostres construccions amb el promotor de *Trl* fusionat a gens *reporter* no inclouen la regió gènica i són perfectament reprimibles per GAGA.

E.1.8. Models per explicar el mecanisme de repressió de *Trl* per GAGA

Amb tots els experiments realitzats hem generat una sèrie de dades que, si bé no resolen completament el mecanisme de repressió de *Trl* per GAGA, sí que aporten una valuosa informació que ens permet donar suport o descartar possibles mecanismes de repressió. D'entre la multitud de possibilitats per explicar la repressió de *Trl* per GAGA, a l'inici d'aquest treball, i per les dades anteriors de que disposàvem, vàrem tenir en compte tres possibles models, que no necessàriament havien de ser mútuament excloents. Aquests models són els que es resumeixen a la figura E.5.

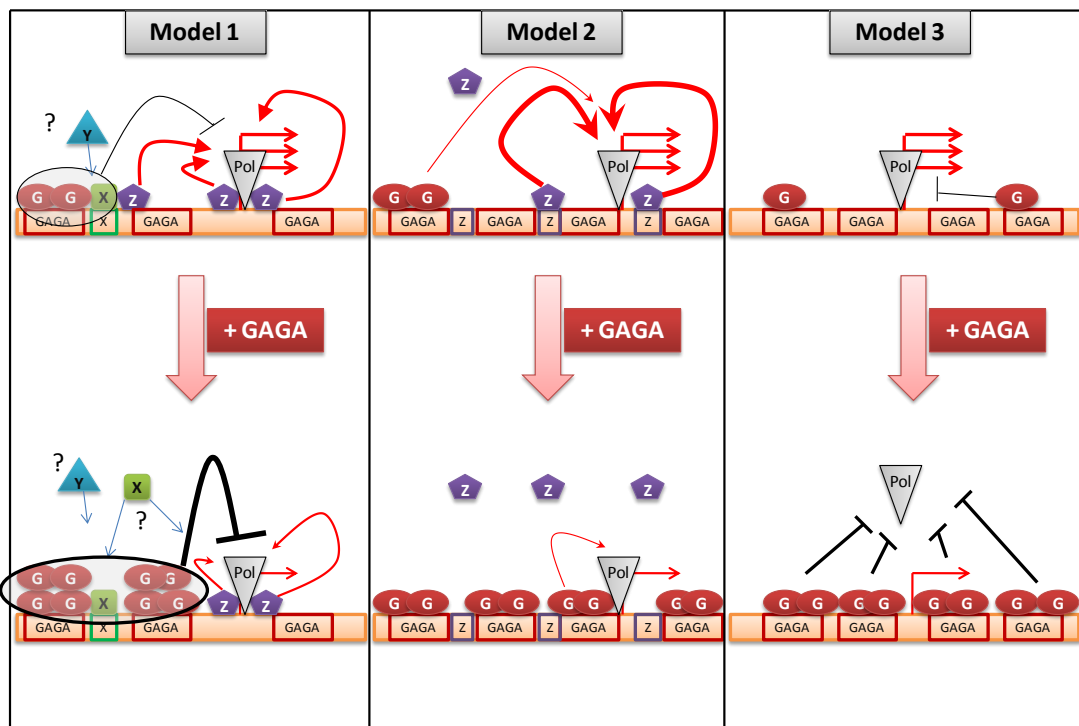


Figura E.5. Representacions esquemàtiques dels diferents models hipotètics proposats per explicar la repressió de *Trl* per GAGA. La barra taronja representa el promotor de *Trl*. Les fletxes vermelles en angle recte representen el nivell de transcripció del promotor. Les fletxes vermelles corbades representen activació transcripcional i les línies negres, repressió. Les figures amb les lletres G, X, Y, Z i Pol representen GAGA, un hipotètic repressor, un hipotètic corepressor, un hipotètic factor activador i la maquinària basal de transcripció, respectivament. A la part superior de la figura es representa l'estat basal del promotor, i a la part inferior l'estat al sobreexpressar GAGA.

El primer model suggereix l'existència, a la seqüència del promotor de *Trl*, de llocs d'unió per a algun factor de caràcter repressor (X a la figura E.5) que, al unir-se al promotor, col·laboraria amb GAGA d'alguna manera i, potser amb l'ajuda d'algun factor corepressor (Y a la figura E.5), donaria la repressió activa del promotor de *Trl*, per sobre de l'acció activadora que

DISCUSSIÓ

puguin tenir altres factors activadors al promotor (Z a la figura E.5). Aquest model té un precedent en un model actualment acceptat per explicar com Dorsal, un conegut activador transcripcional a *Drosophila*, es converteix en un repressor al unir-se a certes regions promotores. Així, la unió de Dorsal a una regió silenciadora del gen *zernknült* (*zen*) provoca el reclutament del corepressor Groucho. Això no passa si només hi ha llocs d'unió de Dorsal, ja que es requereix la presència d'un element de DNA anomenat AT2, al qual s'uneixen certs factors que col·laboren amb Dorsal en el reclutament de Groucho, provocant això la conversió de Dorsal en repressor (Dubnicoff et al. 1997; Valentine et al. 1998). Pel que fa al promotor de *Trl*, segons aquest primer model, en una situació basal, on la GAGA endògena ja demostra estar reprimint el promotor en certa mesura (Kosoy et al. 2002), només unes poques molècules de GAGA estarien sobre el promotor interaccionant amb unes poques molècules de repressor X, donant una feble repressió. Al sobreexpressar GAGA, aquesta ocuparia més llocs a la seqüència del promotor provocant que, o bé es potenciés d'alguna manera l'acció amb el repressor X o bé pogués ser portat més repressor al promotor i/o més corepressor Y, donant així un augment en la repressió que exerceixen GAGA, X i potser Y. Cap dels experiments que hem realitzat no aporten dades en contra d'aquesta hipòtesi, si bé, els intents que hem realitzat per trobar aquest factor repressor X han estat infructuosos, descartant que TTK69 o Psq puguin ser-ho. Pel que fa a l'hipotètic factor corepressor Y, si bé els nostres experiments descarten que pugui ser Groucho, dCtBP sí sembla tenir algun efecte sobre *Trl*, tot i que la seva implicació en aquest mecanisme concret de repressió de *Trl* per GAGA no està del tot resolta. Trobem que la sobreexpressió de dCtBP porta a la repressió de *Trl*, i que GAGA i dCtBP poden interaccionar físicament, el que aporta dades a favor de la implicació de dCtBP en aquest mecanisme repressor. Per contra, l'eliminació de dCtBP per RNAi no sembla afectar l'activitat basal del promotor de *Trl* ni la capacitat de GAGA per reprimir-lo. A més, la sobreexpressió de dCtBP no transforma a GAGA en un repressor de *Trl* a HeLa, cosa que seria d'esperar si fos dCtBP l'element no conservat entre HeLa i *Drosophila* i que determina que GAGA actuï com a repressor de *Trl*. Tot i amb això, també seria possible que a HeLa faltessin més factors essencials per a l'acció repressora de GAGA sobre *Trl* i que la simple sobreexpressió de dCtBP no compensés les diferències entre els dos sistemes.

El segon model aposta per l'existència d'un activador (Z a la figura E.5), d'unió específica al promotor de *Trl*, molt més potent que GAGA. En estat basal, aquest activador Z ocuparia els

seus llocs d'unió a la seqüència del promotor de *Trl* i només una petita quantitat de GAGA, provinent del component endogen, ocuparia uns pocs llocs d'alta afinitat a la seqüència del promotor. Al sobreexpressar-la, GAGA començaria a ocupar també llocs de baixa afinitat, desplaçant així a aquest hipotètic activador més potent. Com que el grau d'activació que donaria GAGA seria molt menor, resultaria en una disminució de l'activitat de *Trl*. A favor d'això està el fet que la forma de GAGA sense el domini de transactivació Q (GAGA Δ Q), a S2 sembla reprimir *Trl* de forma més eficient que GAGA, el que podria ser degut a que aquesta forma truncada de GAGA conserva la capacitat de desplaçar a l'activador de *Trl* però no la d'aportar una mínima activació al promotor. Aquest model és especialment atractiu perquè, a més de considerar GAGA com el que normalment és, un activador, explicaria de manera simple els resultats obtinguts a HeLa. Així, a HeLa mancaria aquest factor activador Z (el que explicaria la baixa activitat d'aquest promotor a aquestes cèl·lules) i per tant, la sobreexpressió de GAGA tindria el mateix resultat que a S2, una feble activació, que a HeLa seria molt gran en termes relatius perquè es passaria d'un estat de pràcticament no expressió a un d'expressió. Cap dels nostres experiments està en contra d'aquest segon model, tot i que s'ha de tenir en compte que, segons els nostres resultats amb les formes delecionades del promotor de *Trl*, aquest factor activador Z, al igual que GAGA, hauria de disposar de multitud de llocs d'unió al llarg de la seqüència del promotor de *Trl*, ja que les múltiples delecions del promotor assajades semblen no afectar a la repressió per GAGA.

Al tercer model és la proteïna GAGA, sense la intervenció directa de cap altre factor, la responsable de la repressió de *Trl*. Segons això, GAGA, per sobre de la seva funció activadora, al promotor de *Trl* actuaria més aviat com a un impediment per a l'acció de la maquinària basal de transcripció (Pol II a la figura E.5). Així, en una situació basal, amb només el component endogen de GAGA, l'escassa quantitat que s'uniria al promotor només destorbaria mínimament a la transcripció. Al sobreexpressar GAGA, aquesta ocuparia massivament el promotor representant un fort impediment per a la maquinària basal de transcripció, el que reduiria dràsticament l'activitat del promotor. Aquest model sorgeix sobretot de la observació, a la seqüència del promotor, d'una gran abundància de llocs GAGA a la regió 5'UTR (veure Resultats, apartat D.2.2.1, figura D.5.A, (Kosoy et al. 2002)). D'aquesta manera GAGA, en una situació de sobreexpressió, podria ocupar abundantment aquests llocs, entorpint l'avanç de la maquinària transcripcional. Tanmateix, les nostres dades experimentals aporten fortes

DISCUSSIÓ

evidències en contra d'aquesta hipòtesi. Primer, si això fos així, no es podria explicar fàcilment que GAGA a HeLa activi *Trl*. La maquinària basal de la transcripció en principi està força conservada al llarg de l'evolució (Young 1991) així que no hi ha raons per a pensar que aquesta pugui ser entorpidida per GAGA a *Drosophila*, però no a HeLa. Segon, si al sobreexpressar GAGA aquesta ocupés de forma massiva el promotor de *Trl*, això hauria d'haver-se observat als nostres *footprintings in vivo* amb DMS a embrions, i no es va observar pas. Tot i que això s'ha de matissar dient que, pel disseny de l'experiment, no totes les cèl·lules de l'embrió sobreexpressaven GAGA, ni les que ho feien ho feien en la mateixa mesura, així que aquest fet pot haver contribuït a diluir les diferències entre embrions control i els que sobreexpressaven GAGA. A més, dades recents obtingudes per *ChIP on chip* indiquen que la presència de GAGA sobre *Trl* ja és molt alta en estat normal a cèl·lules S2, de les més intenses de tot el genoma (Lee et al. 2008).

Finalment, una altra possibilitat seria que GAGA regulés l'expressió de *Trl* a nivell post-transcripcional, dificultant d'alguna manera el procés de traducció. Per exemple, podria ser que GAGA s'unís a l'mRNA de *Trl* per mitjà del seu DBD amb un dit de Zinc, retenint-lo al nucli, ja que GAGA és una proteïna estrictament nuclear (Soeller et al. 1993). Aquesta possibilitat l'hem descartat, atenent a l'absència d'evidències a favor i als resultats dels nostres experiments de RT-PCR i *Real Time* RT-PCR, tot i que encara podria ser que d'alguna manera GAGA estimulés la degradació dels transcrits de *Trl*.

Agafant totes les dades en conjunt, el primer i segon model semblen els més probables, a falta de trobar l'hipotètic factor repressor X o l'hipotètic potent factor activador Z, que acabarien d'esclarir definitivament aquest sorprenent mecanisme d'autoregulació de GAGA sobre el seu propi gen. Precisament en relació a aquest elusiu factor repressor X, dades molt recentment publicades d'experiments de *ChIP on chip* a embrions (Schuettengruber et al. 2009) podrien aportar certa llum a aquesta qüestió. En aquests experiments és pot observar com sobre el promotor i el principi de la regió gènica de *Trl*, col·localitzant de forma extensiva amb GAGA, es troba DSP1 (*Dorsal switch protein 1*). Aquest factor va ser identificat inicialment per la seva capacitat de transformar Dorsal, un conegut activador transcripcional, en repressor (Lehming et al. 1994). Posteriorment, s'ha vist que, al igual que GAGA, pot funcionar com a proteïna TrxG o PcG segons el *locus* considerat (Decoville et al. 2001) podent estar implicat en el silenciament gènic per reclutament de complexos PcG (Muller et al. 2006), tot i que també hi

ha cassos on la seva unió és de vital importància per a l'activació transcripcional (Rappailles et al. 2005). Així, a falta de la demostració experimental adient, la unió de DSP1 al promotor de *Trl* podria estar actuant, juntament amb GAGA, per a donar la repressió activa del promotor. De totes maneres, degut a la seva implicació en l'activació de certs gens, no podem descartar que sigui DSP1 el factor activador que anomenem Z al segon model, sent per tant l'equilibri entre GAGA i DSP1 el que determini l'activitat de *Trl*.

E.2. Estudi funcional de GAGA a S2 per mitjà de *microarrays*

La proteïna GAGA sembla estar implicada en multitud de processos crucials per a la mosca, possiblement regulant la transcripció d'una gran quantitat de gens i actuant a múltiples nivells, resultant en l'activació o en la repressió de la transcripció (Adkins et al. 2006). Per tractar de resoldre qüestions com si GAGA actua essencialment com a factor activador de la transcripció, quins altres gens a part de *Trl* pot reprimir i determinar quines funcions pot estar duent a terme, vàrem realitzar una sèrie d'experiments de *microarray* a la línia cel·lular S2, on analitzàrem com s'afecta l'estat transcripcional global per la sobreexpressió de GAGA i per la seva disminució per mitjà d'RNAi. Aquests experiments però, només són el primer pas de tot un estudi *in vivo* de les funcions de GAGA. Per la facilitat de manipulació i obtenció de les mostres que ofereixen les cèl·lules en cultiu, vàrem decidir començar amb aquesta sèrie d'experiments, que són els que corresponen al treball d'aquesta Tesi Doctoral. Posteriorment, experiments similars analitzaran els efectes en diversos teixits de *Drosophila melanogaster* com discos d'ala de larva, sistema nerviós, ovaris, etc. Així, els experiments aquí presentats s'han d'entendre com la part inicial d'un projecte molt més ampli, amb l'objectiu d'identificar els processos biològics en que aquest factor tant polifacètic està realment implicat, destacant així allò més biològicament rellevant, de tot el conjunt d'informació que al llarg dels anys s'ha anat generant sobre GAGA.

Dels experiments de transcriptòmica realitzats en aquesta Tesi Doctoral, el primer que és pot destacar és la gran fiabilitat dels resultats obtinguts, sobretot de l'experiment de sobreexpressió. Així, quan sobreexpressem GAGA₅₁₉ a la línia cel·lular establement transfectada d'S2, només s'alteren significativament 127 gens (descomptant *Trl* que també apareix com a

DISCUSSIÓ

activat degut a que l'estem sobreexpressant). Això ens indica que la variabilitat entre les mostres o la que pogués provocar la inducció amb CuSO_4 , no afecta significativament el nivell transcripcional de les cèl·lules, i que els pocs positius són, amb molta probabilitat, deguts a l'acció de GAGA. A més, gràcies al disseny de l'experiment de sobreexpressió, fent servir períodes molt curts per a la inducció amb CuSO_4 , aquests gens positius estan molt probablement controlats de forma directa per GAGA. També s'ha de destacar que som capaços de detectar als *microarrays* l'augment de GAGA que provoquem nosaltres amb la sobreexpressió, per mitjà de les sondes d'Affymetrix que cobreixen la regió final del domini Q de GAGA₅₁₉, la isoforma que sobreexpressem (veure figura E.6, sonda 2). Pel que fa a l'experiment d'RNAi, el número de gens significativament alterats per la reducció de la quantitat de GAGA és bastant més gran, 971. Tot i així, aquest no sembla un número de positius exagerat, sobre tot tenint en compte que estem induint les cèl·lules durant 96h amb CuSO_4 i és més probable veure efectes indirectes de la reducció de GAGA. Un fet destacable de l'experiment d'RNAi és que no es detecta la baixada dels nivells de *Trl*, que provoquem amb l'RNAi, i que sí podem veure pels controls de *western blot*. Per aquesta raó, vàrem analitzar els conjunts de sondes d'Affymetrix emprades als *microarrays* per detectar *Trl*. Al *microarray*, s'utilitzen com a sondes situades a la regió gènica de *Trl*, quatre jocs d'oligonucleòtids diferents (sondes 1, 2, 3 i 4 a la figura E.6). A la figura E.6 es representa esquemàticament la regió gènica de *Trl*, així com els possibles transcrits (font: NCBI), indicant a quina isoforma de GAGA donen lloc. S'ha de fer notar que, a les bases de dades (NCBI), degut als algorismes de predicció de llocs d'*splicing*, apareixen dos tipus de transcrits que donen lloc a hipotètiques isoformes de GAGA, que no han estat mai descrites experimentalment (indicades com a ~581 i ~519 per la similitud que tenen amb aquestes isoformes). A la figura també s'indiquen les sondes utilitzades als *microarrays* i quina regió reconeixen. Així, podem veure com la sonda 2, que detecta la sobreexpressió de GAGA al *microarray* de sobreexpressió, correspon al domini Q de GAGA₅₁₉. Pel que fa a l'experiment d'RNAi, no només no detectem la reducció de la quantitat d'mRNA de *Trl* per cap de les sondes, sinó que amb la sonda 1, situada a la regió final del 3'UTR dels dos primers possibles transcrits, detectem un augment de transcrit, a les cèl·lules que sobreexpressem l'RNAi contra GAGA. Parant atenció veiem que aquest augment hauria de ser degut al transcrit que codifica per a la hipotètica forma ~581, ja que la sonda 2 no detecta aquest augment. Tenint en compte que aquesta és una isoforma de GAGA que no s'ha descrit

mai, que l'RNAi contra GAGA hauria d'afectar tant a aquest transcrit com a la resta (ja que tots comparteixen el domini POZ, a partir del qual hem dissenyat l'RNAi), i que el primer transcrit de la figura E.6, que també té la regió amb la que hibrida la sonda 1, no és detectat per la sonda 2, pensem que aquest augment detectat per la sonda 1 a l'experiment d'RNAi és artefactual, potser degut a un RNA del gen *omega* no descrit fins ara, o algun altre tipus d'RNA.

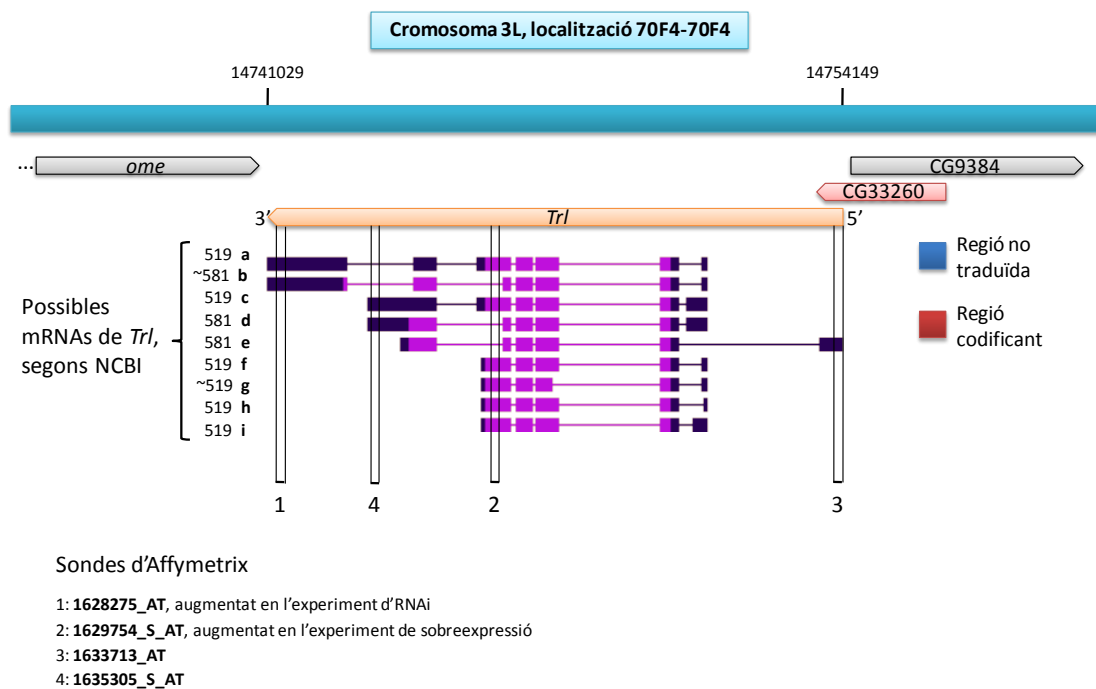


Figura E.6. Representació esquemàtica de la regió gènica de *Trl*, així com dels possibles transcrits (font: NCBI). A l'esquerra de cada transcrit s'indica a quin isoforma de GAGA donaria lloc. S'indiquen les regions on hibrida cada un dels quatre jocs de sondes emprats al *microarray* per a detectar els transcrits de GAGA. A la part inferior de la figura s'indica l'identificador d'Affymetrix de cada sonda i si ha aparegut com a positiu en algun dels experiments. A 3' de *Trl*, a l'altra cadena (la cadena +) es situa el gen *omega* (*ome*), que no solapa amb *Trl*. A 5' de *Trl*, a la mateixa cadena (cadena -), es troba CG33260, la regió 3' codificant del qual sí que solapa amb el 5'UTR del transcrit tipus e de *Trl*.

Fixant-nos en l'experiment de sobreexpressió, crida molt l'atenció el fet que dels 127 gens positius, només 10 corresponguin a gens que es reprimeixen per la sobreexpressió de GAGA. Això suggereix molt clarament que GAGA, a les cèl·lules S2, funciona gairebé de forma exclusiva com a activador transcripcional. Aquests resultats deixen en un segon pla, a cèl·lules S2, les funcions repressores que s'han atribuït a GAGA, ja que, a banda de reprimir el seu propi promotor, també s'ha vist que GAGA pot actuar com a repressor de certs gens, suposadament per la seva implicació en el mecanisme de repressió mediat per les proteïnes PcG (Hagstrom et al. 1997; Horard et al. 2000; Hodgson et al. 2001). En concordança amb els nostres resultats a cèl·lules, altres estudis fets a embrions apunten a que la funció primordial de GAGA no seria actuar amb les proteïnes del PcG, ja que per *ChIP on chip* s'ha vist que la colocalització de GAGA

DISCUSSIÓ

amb *Polycomb* (PC) i *Polyhomeotic* (PH), dues proteïnes del PcG, és molt limitada, suggerint que globalment GAGA no és un component necessari de la cromatina als PREs (Negre et al. 2006). D'altra banda, a l'experiment d'RNAi, al reduir els nivells de GAGA, cabria esperar una major proporció de gens que redueixen el seu nivell d'expressió, ja que GAGA és un activador. Els nostres resultats van en aquesta direcció donat que, dels 917 gens positius a l'experiment d'RNAi, 525 (un 57%) es troben reprimits. Tot i no ser una proporció tan alta com a l'experiment de sobreexpressió, s'ha de tenir en compte que el llarg temps de tractament al que han estat sotmeses aquestes cèl·lules augmenta el número d'efectes indirectes, que poden donar tant activació com repressió gènica i fins i tot emascarar els resultats, per la inducció de mecanismes compensatoris. A més, s'ha de recordar que amb l'RNAi només eliminem parcialment la quantitat de GAGA present a les cèl·lules (veure Resultats, apartat D.8, figura D.24), el que pot no ser suficient per alterar el nivell d'expressió d'alguns dels gens regulats per GAGA.

Pel que respecta a la llista de gens positius obtinguda a cada tipus de *microarray*, una de les primeres coses que vàrem trobar a faltar van ser alguns dels gens que clàssicament s'havien vist que poden ser activats per GAGA. Així, per exemple, certs gens de formació de patró com ara *Ultrabithorax* (*Ubx*), *engrailed* (*en*), *fushi tarazu* (*ftz*), *even-skipped* (*eve*) i *Krüppel* (*kr*) semblen estar sota el control de GAGA (Granok et al. 1995). D'aquests, només *engrailed* apareix com a positiu al *microarray* d'RNAi, reduint-se la seva expressió 1,23 vegades, amb una probabilitat del 88%. Certament, tot i que s'ha tingut en compte com a autèntic positiu, el nivell de canvi és molt escàs, a més de no sortir com a positiu a l'experiment de sobreexpressió. Part de l'explicació de que aquests gens no apareguin com a positius als nostres *microarrays* probablement estigui en el fet que a les cèl·lules S2, cap d'aquest gens s'expressa a uns nivells detectables al *microarray* (segons les dades obtingudes del nivell d'expressió d'aquests gens, respecte els controls interns del *microarray*), i per tant, veure canvis en els seus nivells d'expressió (especialment a la baixa) només alterant els nivells de GAGA és difícil, ja que, tot i que s'ha descrit que GAGA podria regular aquests gens, probablement no sigui l'únic factor implicat en aquesta regulació. En aquest sentit, cal remarcar que *Ubx*, *en* i *eve*, presenten senyals clars de GAGA, en experiments de *ChIP on chip* de cèl·lules S2 (Lee et al. 2008). A més, molt recentment s'han publicat les dades d'experiments de *ChIP on chip* realitzats a embrions de *Drosophila melanogaster* de 4 a 12 h (Schuettengruber et al. 2009), on tots els gens

anteriors, a excepció de *ftz*, presenten senyals clars de GAGA, reforçant la idea de que GAGA, a embrions, sí pot estar controlant l'expressió d'aquests gens.

Les dades, obtingudes per altres grups a partir d'experiments de *ChIP on chip* per a GAGA, donen informació precisa dels llocs al llarg del genoma on aquesta s'uneix. Com acabem de comentar, recentment s'han publicat experiments de *ChIP on chip*, a cèl·lules S2, tant per a NELF com per a GAGA, que permeten concloure que aquests dos factors estan freqüentment associats al *promoter-proximal pausing* (PrPP) de l'RNA Pol II (Lee et al. 2008) (veure Introducció, apartat A.2.5.2.). Donat que els nostres *microarrays* han estat realitzats també a cèl·lules S2, vàrem trobar interessant creuar les nostres dades amb les generades per Lee i col·laboradors. Això ens ha permès determinar que al *microarray* de sobreexpressió, dels 127 gens positius obtinguts en total, 26 gens tenen GAGA unida a la seva seqüència codificant o prou propera com per estar a una possible regió promotora (veure Annexos, apartat H.2.1). Al *microarray* d'RNAi trobem GAGA, en aquestes condicions, en 259 gens dels 917 gens positius del *microarray* (veure Annexos, apartat H.2.2). Així, al 20,47% dels gens del *microarray* de sobreexpressió, i al 28,24% dels gens del *microarray* d'RNAi, es confirma la presència de GAGA prou propera com per existir una alta probabilitat de que reguli directament aquests gens. El primer que sobta d'aquests percentatges és que la proporció de positius que ho són a la vegada al *ChIP on chip* i al nostre *microarray* és major pel *microarray* d'RNAi que pel de sobreexpressió, quan en principi esperàvem el contrari, ja que a l'experiment de sobreexpressió teòricament s'obté major proporció de gens regulats directament per la sobreexpressió de GAGA. Una possible explicació per aquest fet és que les dades de *ChIP on chip* han estat obtingudes de cèl·lules S2 normals, mentre que als nostres experiments de *microarray* els gens positius s'obtenen de la sobreexpressió de GAGA, i de la reducció d'aquesta, a les corresponents línies cel·lulars d'S2 establiment transfectades. Així, podria ser que els gens positius al *microarray* de sobreexpressió, tot i estar la majoria sota el control directe de GAGA, només una petita part d'aquests tinguin ja un nivell de GAGA detectable per *ChIP on chip* a les cèl·lules S2 en condicions normals de cultiu (veure figura E.7.A). En canvi, al reduir la quantitat de GAGA per mitjà de l'RNAi, no esperem observar ocupació *de novo* per part de GAGA (veure figura E.7.B). Així, seria més fàcil correlacionar les dades del *microarray* d'RNAi amb el *ChIP on chip*, ja que al *ChIP on chip* és veu a quins gens està GAGA unida a S2 normals, que són els que canvien directament el seu nivell transcripcional per la reducció de GAGA amb RNAi.

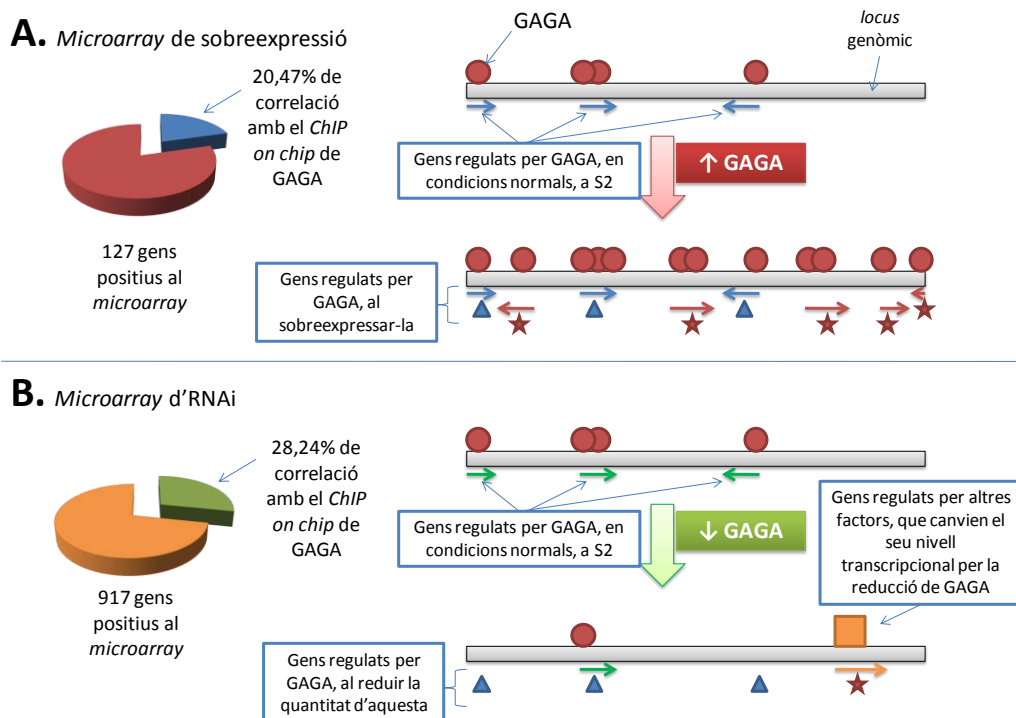


Figura E.7. Possible explicació per a la sensiblement major correlació entre el *microarray* d'RNAi i les dades de *CHIP on chip*. Els triangles blaus indiquen positius tant al *microarray* com al *CHIP on chip*. Les estrelles vermelles indiquen positius al *microarray* però no al *CHIP on chip*. **(A)** Comparant el *microarray* de sobreexpressió amb les dades de *CHIP on chip*, trobem que només el 20,47% dels 127 gens positius al *microarray* es poden correspondre a positius per a GAGA al *CHIP on chip*. Això podria ser degut a que al sobreexpressar GAGA, aquesta ocupi llocs on no està normalment a S2 i per tant, no positius al *CHIP on chip*. **(B)** En canvi, al *microarray* d'RNAi, al reduir la quantitat de GAGA, és més probable que canviïn el seu nivell transcripcional gens ja anteriorment regulats per GAGA, que no pas que la reducció de GAGA faci que aquesta ocupi llocs nous. Possiblement aquesta sigui la raó de la sensiblement major correlació entre el *microarray* d'RNAi i les dades de *CHIP on chip* (28,24%). No s'ha d'oblidar però, que la reducció de GAGA també pot fer variar el nivell transcripcional d'un gran nombre de gens, de manera indirecta (representat pel quadrat taronja a la figura), fet que també es pot donar al sobreexpressar GAGA, però teòricament de manera menys freqüent.

A banda d'això, també trobem que, dels gens positius al *microarray* de sobreexpressió i al *CHIP on chip*, un 96% (25 gens) augmenten l'expressió i un 4% (1 gen) la redueix (veure figura E.8.A). Si mirem això mateix a l'experiment d'RNAi, trobem que un 32% (83 gens) augmenten la seva expressió i un 68% (176 gens) la redueixen (veure figura E.8.B). Aquests percentatges aporten més dades a favor de GAGA com a proteïna activadora de la transcripció.

Per últim, en referència a la comparació de les dades de *CHIP on chip* i els nostres *microarrays*, s'ha de fer notar que, dels 25 gens que sortien diferencialment expressats tant al *microarray* de sobreexpressió com al d'RNAi, 8 són positius al *CHIP on chip* per a GAGA, confirmant aquests gens com a forts candidats a estar regulats directament per l'acció de GAGA (veure Annexos, apartat H.2.3).

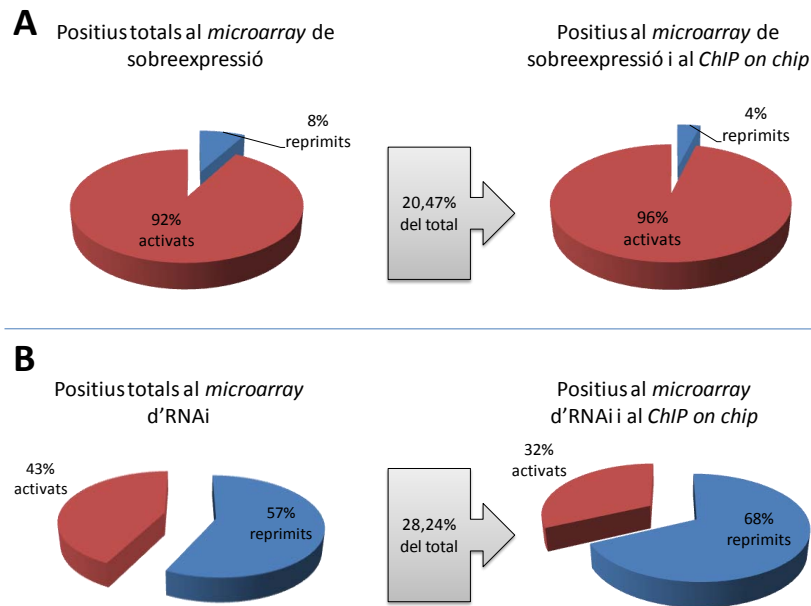


Figura E.8. Major esbiaix de GAGA cap a una funció activadora al analitzar, dels gens que es correlacionen entre cada *microarray* i el *ChIP on chip*, els percentatges de gens reprimits i activats. **(A)** Dels positius totals del *microarray* de sobreexpressió (part esquerra), si només tenim en compte el 20,47% que també són positius pel *ChIP on chip* (part dreta), es veu com la proporció de gens que s'activen per la sobreexpressió de GAGA és encara major. **(B)** Dels positius totals del *microarray* d'RNAi (part esquerra), si només tenim en compte el 28,24% que també són positius pel *ChIP on chip* (part dreta), es veu com la proporció de gens que es reprimeixen per la reducció de la quantitat de GAGA és encara major.

Si ara comparem la llista de funcions biològiques que realitzen els gens que s'activen al sobreexpressar GAGA, amb la llista de funcions biològiques que realitzen els gens que es reprimeixen al reduir la quantitat de GAGA, trobem que moltes d'aquestes funcions estan representades en ambdues llistes (veure figura E.9). Així, l'agrupació de gens més significativa (segons els nostres criteris d'anàlisi) és la mateixa, tant pels gens que s'activen per la sobreexpressió de GAGA, com pels gens que es reprimeixen per la reducció de GAGA: adhesió cel·lular. Això dona suport a la idea abans comentada de que, mirats de forma global, els *microarrays* de sobreexpressió i d'RNAi són experiments complementaris. Centrant l'atenció en el grup de gens d'adhesió cel·lular, podem veure que els dos tipus d'experiments de *microarray* comparteixen només un gen (*Neuroglian*) en aquest grup (veure figura E.10). Aprofitant les dades de l'experiment de *ChIP on chip* podem veure com la correlació entre *microarray* i *ChIP on chip*, dels gens d'aquest grup, és més alta que la correlació global. Pel *microarray* d'RNAi (veure figura E.10.B), es passa del 28,24% de correlació global al 56% si només tenim en compte aquest primer grup de gens. Pel *microarray* de sobreexpressió (veure figura E.10.A) es passa del 20,47% al 37,5%. Això sustenta la idea de que les nostres anàlisis tenen significat biològic, ja que GAGA, per *ChIP on chip*, apareix unida amb més probabilitat als gens dels grups més significatius obtinguts.

DISCUSSIÓ

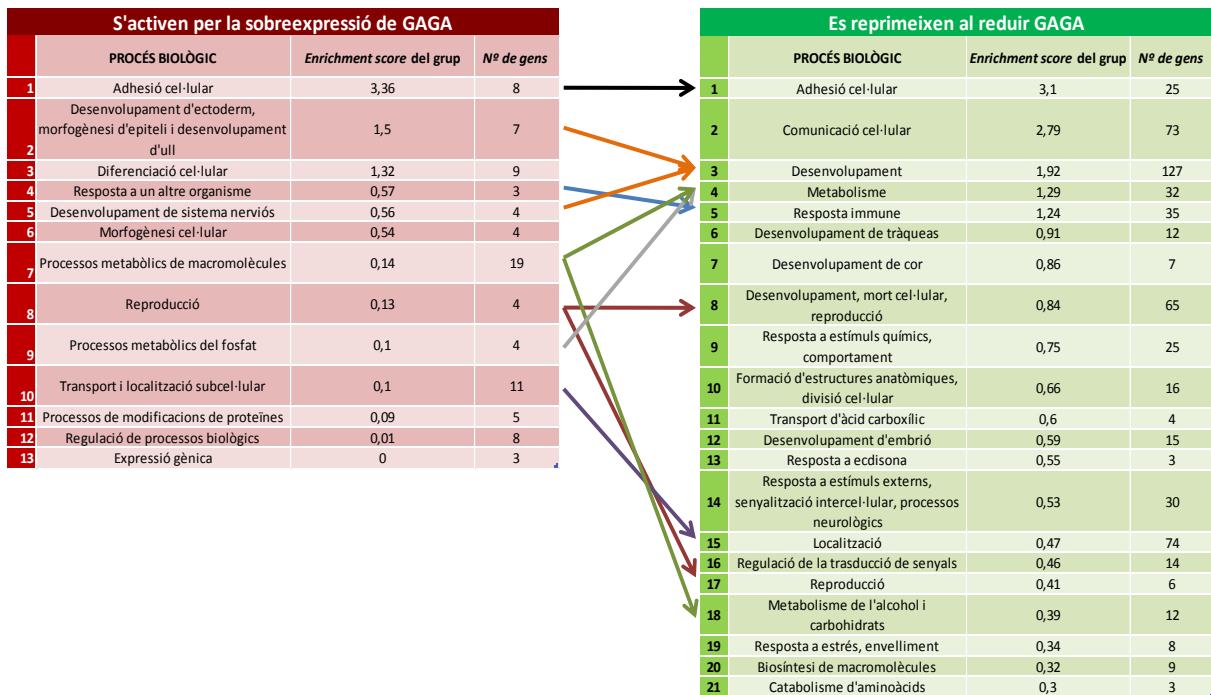


Figura E.9. Comparació de la relació de grups de gens, segons funció biològica, que s'activen per la sobreexpressió de GAGA i es reprimeixen al reduir-la. Les fletxes indiquen les funcions biològiques iguals, o estretament relacionades, que comparteixen ambdues llistes.

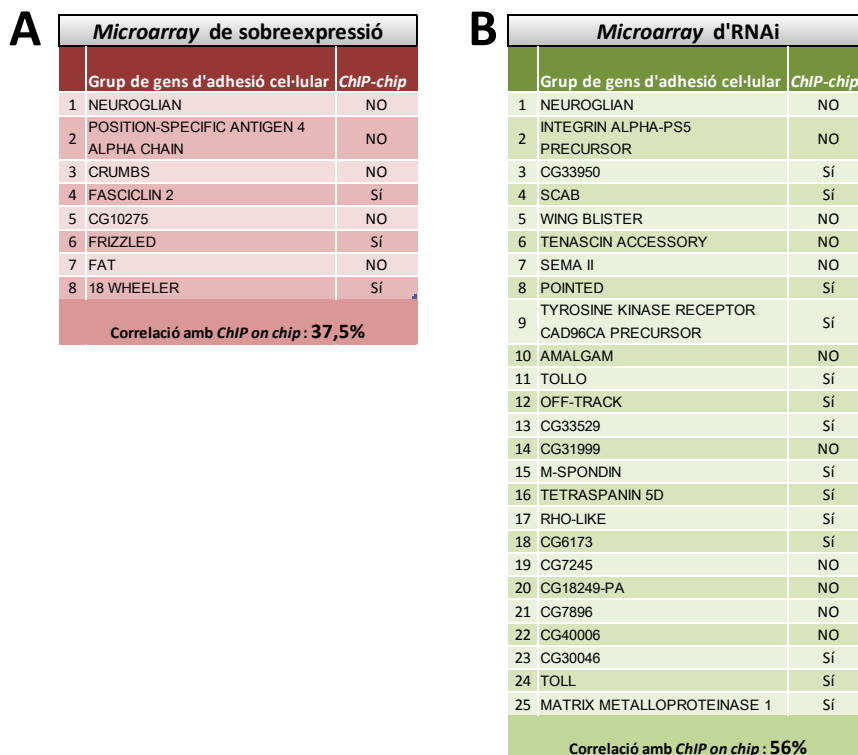


Figura E.10. Llistat de gens que pertanyen al grup d'adhesió cel·lular, que apareixen activats al *microarray* de sobreexpressió (A) i reprimits al *microarray* d'RNAi (B). S'indica el nom del gen i si és positiu o no segons les dades de *ChIP on chip*. A sota de cada llista és dona el percentatge de correlació entre aquests gens i les dades de *ChIP on chip*, és a dir, quin tant per cent de gens de la llista són positius també al *ChIP on chip*. Es pot veure com l'únic gen que comparteixen, *Neuroglian*, no té GAGA prou a prop segons les dades de *ChIP on chip*.

Pel fet que, com ja hem comentat, aquests experiments de *microarray* aquí presentats formen part d'un projecte més ampli, que pretén l'anàlisi funcional *in vivo* de GAGA a múltiples sistemes amb més rellevància biològica que no pas la línia cel·lular S2, no vàrem entrar en una comprovació sistemàtica dels possibles gens controlats per GAGA obtinguts als nostres *microarrays*. No obstant això, vàrem trobar molt interessant que *sickle*, un gen implicat en apoptosi (Srinivasula et al. 2002; Wing et al. 2002), aparegués com activat al sobreexpressar GAGA (veure Resultats, apartat D.8, figura D.25.A). Observacions anteriors com la baixa viabilitat de les mosques transgèniques que sobreexpressen GAGA, junt amb alguns dels fenotips obtinguts amb la sobreexpressió de GAGA conduïda per *dppGal4*, on la malformació o pèrdua de les potes suggereixen un mecanisme de mort cel·lular (Bernues et al. 2007), ens van portar a pensar que GAGA, en determinats llinatges cel·lulars, podria estar controlant alguna de les vies que porten a l'apoptosi. Així, per aquesta raó, vàrem voler corroborar aquest positiu del *microarray* per altres tècniques. Per immunolocalització amb anti-*sickle* tant a discos d'ala de larves de 3r estadi que sobreexpressen GAGA₅₁₉ sota el control del *driver* *dppGal4* (realitzades pel Dr. Jordi Bernués, veure Resultats, apartat D.8.1, figura D.30) com a cèl·lules S2 establiment transfectades que sobreexpressen GAGA (veure Resultats, apartat D.8.1, figura D.29), hem confirmat que la sobreexpressió de GAGA augmenta la quantitat de *sickle*. Així, el fet que *sickle* aparegui activat al *microarray* de sobreexpressió i que detectem l'augment de la proteïna *sickle* per sobreexpressió de GAGA, a més dels fenotips obtinguts a mosques transgèniques que sobreexpressen GAGA, recolzen la hipòtesi de que GAGA podria estar controlant d'alguna manera l'apoptosi, a certs llinatges cel·lulars. S'ha de tenir en compte però, que aquests són resultats molt preliminars, que haurien de confirmar-se en els posteriors *microarrays* de teixits larvaris, ja que a la nostra línia cel·lular no observem una mortalitat molt acusada deguda a la sobreexpressió de GAGA, almenys a curt termini. A més, *sickle* no apareix com a positiu al *ChIP on chip* de GAGA a cèl·lules S2 (veure Annexos, apartat H.2), indicant que el control de GAGA sobre *sickle* podria ser de forma indirecta, tot i que a experiments molt recents de *ChIP on chip* a embrions de 4 a 12 h (Schuettengruber et al. 2009), GAGA sí apareix unida al gen *sickle*.

Finalment, s'ha de fer notar l'enorme proporció de gens positius, tant al *microarray* de sobreexpressió com al d'RNAi, amb funció biològica desconeguda. Al *microarray* de sobreexpressió, un 55% dels gens que s'activen (veure Resultats, apartat D.8, figura D.26.A) i 6 dels 10 gens que es reprimeixen (veure Resultats, apartat D.8, figura D.26.B) no tenen funció

DISCUSSIÓ

biològica assignada i per tant, no han pogut entrar en les nostres anàlisis. Al *microarray* d'RNAi el 51% dels gens que s'activen (veure Resultats, apartat D.8, figura D.27.A) i el 46% dels gens que es reprimeixen (veure Resultats, apartat D.8, figura D.27.B) tampoc tenen funció biològica assignada. Això obre la possibilitat de que GAGA estigui implicada en multitud de processos més, a banda de la gran quantitat que ja han estat descrits, i obre el camí a investigar noves funcions d'aquest factor, vital per a *Drosophila melanogaster*.

F. CONCLUSIONS

- 1- Per transfeccions transitòries a diferents línies cel·lulars, hem comprovat que el mecanisme de repressió de *Trl* per GAGA està estrictament conservat entre *D. melanogaster* i *D. virilis*. Al mateix temps, hem vist que GAGA, a la línia de cèl·lules HeLa, no només no reprimeix sinó que activa el promotor de *Trl*.
- 2- Utilitzant assajos de *footprinting in vivo* amb DMS, hem analitzat les regions del promotor de *Trl* que *a priori* semblaven més importants per a la repressió. D'aquesta manera, hem determinat les seqüències que es troben protegides a cèl·lules S2 en cultiu, així com aquelles seqüències que esdevenen protegides per la sobreexpressió de GAGA a embrions transgènics de *D. melanogaster*. No obstant això, la mutació d'aquestes seqüències no afecta a la repressió de *Trl* per GAGA, tot i que l'activitat inicial del promotor si es veu reduïda en algun cas. A més, hem pogut comprovar que el promotor de *Trl* no és troba massivament protegit en una situació de sobreexpressió de GAGA, suggerint que aquesta no ocupa totes les seves seqüències d'unió potencials.
- 3- Per mitjà de delecions successives del promotor de *Trl*, hem definit entre -204 i +71 la regió mínima per a que es pugui observar repressió per GAGA. Tot i amb això, no podem assegurar que el determinant o determinants de la repressió per GAGA estiguin en aquesta regió ja que qualsevol de les delecions provades que afecten a aquesta regió afecten molt negativament a l'activitat inicial del promotor.
- 4- Hem comprovat que els aminoàcids del 222 al 303 de la regió X són prescindibles per a que es doni la repressió de *Trl*. A més, també hem comprovat que el DBD no és suficient per a que GAGA es localitzi al nucli i per tant pugui reprimir *Trl*. L'NLS tampoc sembla imprescindible per a que GAGA es localitzi al nucli, tot i que les construccions que no el porten tenen un patró de distribució nuclear anormal. Així, hem definit que els aminoàcids entre 122 i 222, que inclouen l'NLS, tot i ser prescindibles per a que GAGA es localitzi al nucli, sí són essencials per a la correcta disposició de GAGA dins del nucli i per tant, també per a la seva activitat.
- 5- Hem comprovat que els factors de transcripció TTK69 i Psq no semblen tenir cap efecte sobre el promotor de *Trl*. Al mateix temps, també hem vist que el corepressor Groucho sembla no estar implicat en la repressió de *Trl*. Per contra, el corepressor dCtBP sí sembla estar implicat en la repressió de *Trl*. Tot i que hem comprovat que GAGA i dCtBP poden interaccionar *in vitro*, els nostres experiments suggereixen que l'acció repressora de dCtBP és independent de GAGA.

CONCLUSIONS

6- La repressió de *Trl* per GAGA es dona també sobre la còpia endògena del gen *Trl*, tant a cèl·lules S2, com a discos d'ala de larves de 3r estadi. A més, experiments a embrions i a discos imaginals de mosques transgèniques indiquen que, al igual que en el cas de les cèl·lules S2, a la mosca les dues isoformes de GAGA són capaces de reprimir *Trl*, aquesta repressió és independent del domini Q i necessita que GAGA conservi la seva capacitat d'unir-se al DNA. També, de manera anàloga al que s'observa a S2, a la mosca el promotor de *Trl* sembla estar sotmès a un cert control per part del component endogen de GAGA, mantenint aquest un determinat nivell de repressió sobre *Trl*.

7- Els fenotips obtinguts de mosques transgèniques suggereixen que les dues isoformes de GAGA tenen funcions solapants i que GAGA₅₁₉ seria la més activa, ja que els fenotips derivats de la seva sobreexpressió són més severos. L'absència de fenotips derivats de la sobreexpressió de GAGA₅₁₉DBDKO indica que GAGA, per a dur a terme les seves funcions a la mosca, necessita unir-se al DNA.

8- Entre les diferents possibilitats per explicar la repressió de *Trl* per GAGA, els nostres resultats ens porten a decantar-nos per dos possibles models:

- a. A la seqüència del promotor de *Trl* existeixen llocs d'unió per a algun factor repressor. Al unir-se GAGA i aquest factor repressor al promotor actuarien de forma conjunta, potser amb l'ajuda d'algun factor corepressor, donant la repressió activa del promotor de *Trl*.
- b. Existeix un factor activador, d'unió específica al promotor de *Trl* i molt més potent que GAGA, que determina l'activitat d'aquest promotor. Al sobreexpressar GAGA, aquesta desplaçaria al potent activador de *Trl* i, donat que la possible acció activadora de GAGA seria d'un nivell molt inferior, això resultaria en una davallada de l'activitat de *Trl*.

9- Els *microarrays* realitzats a cèl·lules S2 establement transfectades suggereixen fortament que la funció predominant de GAGA a aquestes cèl·lules és l'activació transcripcional. A més, els nostres resultats obren la possibilitat de que GAGA estigui implicada en multitud de processos, a més dels ja coneguts, com poden ser l'adhesió cel·lular i la resposta immune. Hem comprovat que GAGA pot activar *sickle*, un gen pro-apoptòtic, que podria estar relacionat amb certs fenotips de letalitat derivats de la sobreexpressió de GAGA.

10- Aquests resultats, en conjunt, reforcen la noció de que GAGA és un factor de transcripció activador, que reprimeix la seva pròpia expressió i que aquest mecanisme, altament conservat a *Drosophila*, és essencial per a mantenir uns nivells correctes de proteïna i evitar així greus defectes derivats d'una possible expressió en excés.

G. BIBLIOGRAFIA

- Adkins, N. L., T. A. Hagerman, et al. (2006). "GAGA protein: a multi-faceted transcription factor." Biochem Cell Biol **84**(4): 559-67.
- Agianian, B., K. Leonard, et al. (1999). "The glutamine-rich domain of the Drosophila GAGA factor is necessary for amyloid fibre formation in vitro, but not for chromatin remodelling." J Mol Biol **285**(2): 527-44.
- Ahn, S. H., M. Kim, et al. (2004). "Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing." Mol Cell **13**(1): 67-76.
- Akashi, S., S. Nagakura, et al. (2008). "Structural characterization of human general transcription factor TFIIF in solution." Protein Sci **17**(3): 389-400.
- Albright, S. R. and R. Tjian (2000). "TAFs revisited: more data reveal new twists and confirm old ideas." Gene **242**(1-2): 1-13.
- Babu, M. M., N. M. Luscombe, et al. (2004). "Structure and evolution of transcriptional regulatory networks." Curr Opin Struct Biol **14**(3): 283-91.
- Bajic, V. B., S. L. Tan, et al. (2006). "Mice and men: their promoter properties." PLoS Genet **2**(4): e54.
- Bardwell, V. J. and R. Treisman (1994). "The POZ domain: a conserved protein-protein interaction motif." Genes Dev **8**(14): 1664-77.
- Bartfai, R., C. Balduf, et al. (2004). "TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish." Curr Biol **14**(7): 593-8.
- Bejarano, F. and A. Busturia (2004). "Function of the Trithorax-like gene during Drosophila development." Dev Biol **268**(2): 327-41.
- Belotserkovskaya, R., S. Oh, et al. (2003). "FACT facilitates transcription-dependent nucleosome alteration." Science **301**(5636): 1090-3.
- Belozero, V. E., P. Majumder, et al. (2003). "A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of Drosophila." EMBO J **22**(12): 3113-21.
- Bell, S. D., A. B. Brinkman, et al. (2001). "The archaeal TFIIE α homologue facilitates transcription initiation by enhancing TATA-box recognition." EMBO Rep **2**(2): 133-8.
- Bentley, D. (2002). "The mRNA assembly line: transcription and processing machines in the same factory." Curr Opin Cell Biol **14**(3): 336-42.
- Benyajati, C., L. Mueller, et al. (1997). "Multiple isoforms of GAGA factor, a critical component of chromatin structure." Nucleic Acids Res **25**(16): 3345-53.
- Bernues, J., D. Pineyro, et al. (2007). "General, negative feedback mechanism for regulation of Trithorax-like gene expression in vivo: new roles for GAGA factor in flies." Nucleic Acids Res **35**(21): 7150-9.
- Beverly, S. M. and A. C. Wilson (1984). "Molecular evolution in Drosophila and the higher Diptera II. A time scale for fly evolution." J Mol Evol **21**(1): 1-13.
- Bevilacqua, A., M. T. Fiorenza, et al. (2000). "A developmentally regulated GAGA box-binding factor and Sp1 are required for transcription of the hsp70.1 gene at the onset of mouse zygotic genome activation." Development **127**(7): 1541-51.
- Bhat, K. M., G. Farkas, et al. (1996). "The GAGA factor is required in the early Drosophila embryo not only for transcriptional regulation but also for nuclear division." Development **122**(4): 1113-24.
- Biggin, M. D. and R. Tjian (1988). "Transcription factors that activate the Ultrabithorax promoter in developmentally staged extracts." Cell **53**(5): 699-711.
- Blau, J., H. Xiao, et al. (1996). "Three functional classes of transcriptional activation domain." Mol Cell Biol **16**(5): 2044-55.
- Bleichenbacher, M., S. Tan, et al. (2003). "Novel interactions between the components of human and yeast TFIIA/TBP/DNA complexes." J Mol Biol **332**(4): 783-93.
- Bohr, V. A., C. A. Smith, et al. (1985). "DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall." Cell **40**(2): 359-69.
- Bonet, C., I. Fernandez, et al. (2005). "The GAGA protein of Drosophila is phosphorylated by CK2." J Mol Biol **351**(3): 562-72.

BIBLIOGRAFIA

- Brand, A. H. and N. Perrimon (1993). "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes." *Development* **118**(2): 401-15.
- Brent, R. and M. Ptashne (1985). "A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor." *Cell* **43**(3 Pt 2): 729-36.
- Brown, C. E., T. Lechner, et al. (2000). "The many HATs of transcription coactivators." *Trends Biochem Sci* **25**(1): 15-9.
- Brown, J. L. and C. Wu (1993). "Repression of Drosophila pair-rule segmentation genes by ectopic expression of tramtrack." *Development* **117**(1): 45-58.
- Buratowski, S. (2005). "Connections between mRNA 3' end processing and transcription termination." *Curr Opin Cell Biol* **17**(3): 257-61.
- Buratowski, S., S. Hahn, et al. (1989). "Five intermediate complexes in transcription initiation by RNA polymerase II." *Cell* **56**(4): 549-61.
- Burke, T. W. and J. T. Kadonaga (1997). "The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila." *Genes Dev* **11**(22): 3020-31.
- Bushnell, D. A., K. D. Westover, et al. (2004). "Structural basis of transcription: an RNA polymerase II-TFIIB cocystal at 4.5 Angstroms." *Science* **303**(5660): 983-8.
- Butler, J. E. and J. T. Kadonaga (2002). "The RNA polymerase II core promoter: a key component in the regulation of gene expression." *Genes Dev* **16**(20): 2583-92.
- Carcamo, J., L. Buckbinder, et al. (1991). "The initiator directs the assembly of a transcription factor IID-dependent transcription complex." *Proc Natl Acad Sci U S A* **88**(18): 8052-6.
- Carey, M., Y. S. Lin, et al. (1990). "A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives." *Nature* **345**(6273): 361-4.
- Carninci, P., A. Sandelin, et al. (2006). "Genome-wide analysis of mammalian promoter architecture and evolution." *Nat Genet* **38**(6): 626-35.
- Coin, F., L. Proietti De Santis, et al. (2006). "p8/TTD-A as a repair-specific TFIIF subunit." *Mol Cell* **21**(2): 215-26.
- Conaway, R. C. and J. W. Conaway (1989). "An RNA polymerase II transcription factor has an associated DNA-dependent ATPase (dATPase) activity strongly stimulated by the TATA region of promoters." *Proc Natl Acad Sci U S A* **86**(19): 7356-60.
- Costello, J. F. and C. Plass (2001). "Methylation matters." *J Med Genet* **38**(5): 285-303.
- Courey, A. J. and S. Jia (2001). "Transcriptional repression: the long and the short of it." *Genes Dev* **15**(21): 2786-96.
- Cramer, P. (2004). "RNA polymerase II structure: from core to functional complexes." *Curr Opin Genet Dev* **14**(2): 218-26.
- Cramer, P., K. J. Armache, et al. (2008). "Structure of eukaryotic RNA polymerases." *Annu Rev Biophys* **37**: 337-52.
- Creti, R., P. Londei, et al. (1993). "Complete nucleotide sequence of an archaeal (Pyrococcus woesei) gene encoding a homolog of eukaryotic transcription factor IIB (TFIIB)." *Nucleic Acids Res* **21**(12): 2942.
- Croston, G. E., L. A. Kerrigan, et al. (1991). "Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription." *Science* **251**(4994): 643-9.
- Croston, G. E., P. J. Laybourn, et al. (1992). "Mechanism of transcriptional antirepression by GAL4-VP16." *Genes Dev* **6**(12A): 2270-81.
- Crowley, T. E., T. Hoey, et al. (1993). "A new factor related to TATA-binding protein has highly restricted expression patterns in Drosophila." *Nature* **361**(6412): 557-61.
- Cuesta, I., K. S. Zaret, et al. (2007). "The forkhead factor FoxE1 binds to the thyroperoxidase promoter during thyroid cell differentiation and modifies compacted chromatin structure." *Mol Cell Biol* **27**(20): 7302-14.
- Chalkley, G. E. and C. P. Verrijzer (1999). "DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator." *EMBO J* **18**(17): 4835-45.

- Cho, E. J., M. S. Kobor, et al. (2001). "Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain." *Genes Dev* **15**(24): 3319-29.
- Cho, E. J., T. Takagi, et al. (1997). "mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain." *Genes Dev* **11**(24): 3319-26.
- Cho, H., G. Orphanides, et al. (1998). "A human RNA polymerase II complex containing factors that modify chromatin structure." *Mol Cell Biol* **18**(9): 5355-63.
- Chopra, V. S., A. Srinivasan, et al. (2008). "Transcriptional activation by GAGA factor is through its direct interaction with dmTAF3." *Dev Biol* **317**(2): 660-70.
- Dahmus, M. E. (1996). "Reversible phosphorylation of the C-terminal domain of RNA polymerase II." *J Biol Chem* **271**(32): 19009-12.
- Dantoni, J. C., J. M. Wurtz, et al. (1999). "The TBP-like factor: an alternative transcription factor in metazoa?" *Trends Biochem Sci* **24**(9): 335-9.
- Deato, M. D. and R. Tjian (2007). "Switching of the core transcription machinery during myogenesis." *Genes Dev* **21**(17): 2137-49.
- Decoville, M., E. Giacomello, et al. (2001). "DSP1, an HMG-like protein, is involved in the regulation of homeotic genes." *Genetics* **157**(1): 237-44.
- DeJong, J. and R. G. Roeder (1993). "A single cDNA, hTFIIA/alpha, encodes both the p35 and p19 subunits of human TFIIA." *Genes Dev* **7**(11): 2220-34.
- Deng, W. and S. G. Roberts (2005). "A core promoter element downstream of the TATA box that is recognized by TFIIB." *Genes Dev* **19**(20): 2418-23.
- Deng, W. and S. G. Roberts (2007). "TFIIB and the regulation of transcription by RNA polymerase II." *Chromosoma* **116**(5): 417-29.
- Dion, M. F., S. J. Altschuler, et al. (2005). "Genomic characterization reveals a simple histone H4 acetylation code." *Proc Natl Acad Sci U S A* **102**(15): 5501-6.
- Drapkin, R., J. T. Reardon, et al. (1994). "Dual role of TFIIF in DNA excision repair and in transcription by RNA polymerase II." *Nature* **368**(6473): 769-72.
- Dubnicoff, T., S. A. Valentine, et al. (1997). "Conversion of dorsal from an activator to a repressor by the global corepressor Groucho." *Genes Dev* **11**(22): 2952-7.
- Dvir, A., R. C. Conaway, et al. (1997). "A role for TFIIF in controlling the activity of early RNA polymerase II elongation complexes." *Proc Natl Acad Sci U S A* **94**(17): 9006-10.
- Dvir, A., K. P. Garrett, et al. (1996). "A role for ATP and TFIIF in activation of the RNA polymerase II preinitiation complex prior to transcription initiation." *J Biol Chem* **271**(13): 7245-8.
- Ebright, R. H. (2000). "RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II." *J Mol Biol* **304**(5): 687-98.
- Egloff, S. and S. Murphy (2008). "Role of the C-terminal domain of RNA polymerase II in expression of small nuclear RNA genes." *Biochem Soc Trans* **36**(Pt 3): 537-9.
- Eisenberg, J. C. (2001). "Decisive factors: a transcription activator can overcome heterochromatin silencing." *Bioessays* **23**(9): 767-71.
- Espinosa, M. L., S. Canudas, et al. (2000). "The GAGA factor of Drosophila interacts with SAP18, a Sin3-associated polypeptide." *EMBO Rep* **1**(3): 253-9.
- Espinosa, M. L., E. Jimenez-Garcia, et al. (1999). "The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity." *J Biol Chem* **274**(23): 16461-9.
- Farkas, G., J. Gausz, et al. (1994). "The Trithorax-like gene encodes the Drosophila GAGA factor." *Nature* **371**(6500): 806-8.
- Faro-Trindade, I. and P. R. Cook (2006). "Transcription factories: structures conserved during differentiation and evolution." *Biochem Soc Trans* **34**(Pt 6): 1133-7.
- Faucheux, M., J. Y. Roignant, et al. (2003). "batman Interacts with polycomb and trithorax group genes and encodes a BTB/POZ protein that is included in a complex containing GAGA factor." *Mol Cell Biol* **23**(4): 1181-95.

BIBLIOGRAFIA

- Feaver, W. J., O. Gileadi, et al. (1991). "CTD kinase associated with yeast RNA polymerase II initiation factor b." *Cell* **67**(6): 1223-30.
- FitzGerald, P. C., A. Shlyakhtenko, et al. (2004). "Clustering of DNA sequences in human promoters." *Genome Res* **14**(8): 1562-74.
- FitzGerald, P. C., D. Sturgill, et al. (2006). "Comparative genomics of Drosophila and human core promoters." *Genome Biol* **7**(7): R53.
- Flores, O., I. Ha, et al. (1990). "Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF." *J Biol Chem* **265**(10): 5629-34.
- Flores, O., H. Lu, et al. (1991). "The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex." *Proc Natl Acad Sci U S A* **88**(22): 9999-10003.
- Francis, N. J. and R. E. Kingston (2001). "Mechanisms of transcriptional memory." *Nat Rev Mol Cell Biol* **2**(6): 409-21.
- Freiman, R. N. and R. Tjian (2003). "Regulating the regulators: lysine modifications make their mark." *Cell* **112**(1): 11-7.
- Freire-Picos, M. A., S. Krishnamurthy, et al. (2005). "Evidence that the Tfg1/Tfg2 dimer interface of TFIIF lies near the active center of the RNA polymerase II initiation complex." *Nucleic Acids Res* **33**(16): 5045-52.
- Frustaci, J. M., I. Sangwan, et al. (1995). "gsa1 is a universal tetrapyrrole synthesis gene in soybean and is regulated by a GAGA element." *J Biol Chem* **270**(13): 7387-93.
- Gautier, L., L. Cope, et al. (2004). "affy--analysis of Affymetrix GeneChip data at the probe level." *Bioinformatics* **20**(3): 307-15.
- Gentleman, R. C., V. J. Carey, et al. (2004). "Bioconductor: open software development for computational biology and bioinformatics." *Genome Biol* **5**(10): R80.
- Gerasimova, T. I. and V. G. Corces (1998). "Polycomb and trithorax group proteins mediate the function of a chromatin insulator." *Cell* **92**(4): 511-21.
- Gershenson, N. I. and I. P. Ioshikhes (2005). "Synergy of human Pol II core promoter elements revealed by statistical sequence analysis." *Bioinformatics* **21**(8): 1295-300.
- Gershenson, N. I., E. N. Trifonov, et al. (2006). "The features of Drosophila core promoters revealed by statistical analysis." *BMC Genomics* **7**: 161.
- Ghazy, M. A., S. A. Brodie, et al. (2004). "Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II." *Mol Cell Biol* **24**(24): 10975-85.
- Giglia-Mari, G., F. Coin, et al. (2004). "A new, tenth subunit of TFIIF is responsible for the DNA repair syndrome trichothiodystrophy group A." *Nat Genet* **36**(7): 714-9.
- Gilmour, D. S., G. H. Thomas, et al. (1989). "Drosophila nuclear proteins bind to regions of alternating C and T residues in gene promoters." *Science* **245**(4925): 1487-90.
- Godt, D., J. L. Couderc, et al. (1993). "Pattern formation in the limbs of Drosophila: bric a brac is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus." *Development* **119**(3): 799-812.
- Gonzalez, G. A. and M. R. Montminy (1989). "Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133." *Cell* **59**(4): 675-80.
- Gowher, H., O. Leisemann, et al. (2000). "DNA of Drosophila melanogaster contains 5-methylcytosine." *EMBO J* **19**(24): 6918-23.
- Granok, H., B. A. Leibovitch, et al. (2001). "A heat-shock-activated cDNA encoding GAGA factor rescues some lethal mutations in the Drosophila melanogaster Trithorax-like gene." *Genet Res* **78**(1): 13-21.
- Granok, H., B. A. Leibovitch, et al. (1995). "Chromatin. Ga-ga over GAGA factor." *Curr Biol* **5**(3): 238-41.
- Greenberg, A. J. and P. Schedl (2001). "GAGA factor isoforms have distinct but overlapping functions in vivo." *Mol Cell Biol* **21**(24): 8565-74.

- Gu, W. and R. G. Roeder (1997). "Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain." *Cell* **90**(4): 595-606.
- Guenther, M. G., S. S. Levine, et al. (2007). "A chromatin landmark and transcription initiation at most promoters in human cells." *Cell* **130**(1): 77-88.
- Hagemann, T. L., D. Mares, et al. (2000). "Gene regulation of Wiskott-Aldrich syndrome protein and the human homolog of the Drosophila Su(var)3-9: WASP and SUV39H1, two adjacent genes at Xp11.23." *Biochim Biophys Acta* **1493**(3): 368-72.
- Hagstrom, K., M. Muller, et al. (1997). "A Polycomb and GAGA dependent silencer adjoins the Fab-7 boundary in the Drosophila bithorax complex." *Genetics* **146**(4): 1365-80.
- Hampsey, M. (1998). "Molecular genetics of the RNA polymerase II general transcriptional machinery." *Microbiol Mol Biol Rev* **62**(2): 465-503.
- Hampsey, M. and D. Reinberg (2003). "Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation." *Cell* **113**(4): 429-32.
- Hansen, S. K., S. Takada, et al. (1997). "Transcription properties of a cell type-specific TATA-binding protein, TRF." *Cell* **91**(1): 71-83.
- Hawkes, N. A. and S. G. Roberts (1999). "The role of human TFIIB in transcription start site selection in vitro and in vivo." *J Biol Chem* **274**(20): 14337-43.
- Herr, A. J., M. B. Jensen, et al. (2005). "RNA polymerase IV directs silencing of endogenous DNA." *Science* **308**(5718): 118-20.
- Hirose, F., M. Yamaguchi, et al. (1993). "Novel 8-base pair sequence (Drosophila DNA replication-related element) and specific binding factor involved in the expression of Drosophila genes for DNA polymerase alpha and proliferating cell nuclear antigen." *J Biol Chem* **268**(3): 2092-9.
- Hisatake, K., S. Malik, et al. (1991). "Conserved structural motifs between Xenopus and human TFIIB." *Nucleic Acids Res* **19**(23): 6639.
- Hochheimer, A., S. Zhou, et al. (2002). "TRF2 associates with DREF and directs promoter-selective gene expression in Drosophila." *Nature* **420**(6914): 439-45.
- Hodgson, J. W., B. Argiropoulos, et al. (2001). "Site-specific recognition of a 70-base-pair element containing d(GA)(n) repeats mediates bithoraxoid polycomb group response element-dependent silencing." *Mol Cell Biol* **21**(14): 4528-43.
- Holstege, F. C., P. C. van der Vliet, et al. (1996). "Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH." *EMBO J* **15**(7): 1666-77.
- Hoogstraten, D., A. L. Nigg, et al. (2002). "Rapid switching of TFIIF between RNA polymerase I and II transcription and DNA repair in vivo." *Mol Cell* **10**(5): 1163-74.
- Hooper, J. E. and M. P. Scott (1989). "The Drosophila patched gene encodes a putative membrane protein required for segmental patterning." *Cell* **59**(4): 751-65.
- Horard, B., C. Tatout, et al. (2000). "Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor." *Mol Cell Biol* **20**(9): 3187-97.
- Hozak, P., P. R. Cook, et al. (1994). "Site of transcription of ribosomal RNA and intranucleolar structure in HeLa cells." *J Cell Sci* **107** (Pt 2): 639-48.
- Huang, D. H., Y. L. Chang, et al. (2002). "pipsqueak encodes a factor essential for sequence-specific targeting of a polycomb group protein complex." *Mol Cell Biol* **22**(17): 6261-71.
- Huynh, K. D. and V. J. Bardwell (1998). "The BCL-6 POZ domain and other POZ domains interact with the co-repressors N-CoR and SMRT." *Oncogene* **17**(19): 2473-84.
- Iben, S., H. Tschochner, et al. (2002). "TFIIF plays an essential role in RNA polymerase I transcription." *Cell* **109**(3): 297-306.
- Imbalzano, A. N., K. S. Zaret, et al. (1994). "Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA." *J Biol Chem* **269**(11): 8280-6.
- Irizarry, R. A., B. Hobbs, et al. (2003). "Exploration, normalization, and summaries of high density oligonucleotide array probe level data." *Biostatistics* **4**(2): 249-64.

BIBLIOGRAFIA

- Isogai, Y., S. Keles, et al. (2007a). "Transcription of histone gene cluster by differential core-promoter factors." *Genes Dev* **21**(22): 2936-49.
- Isogai, Y., S. Takada, et al. (2007b). "Novel TRF1/BRF target genes revealed by genome-wide analysis of *Drosophila* Pol III transcription." *EMBO J* **26**(1): 79-89.
- Itoh, Y., S. Unzai, et al. (2005). "Investigation of molecular size of transcription factor TFIIE in solution." *Proteins* **61**(3): 633-41.
- Jackson, D. A., F. J. Iborra, et al. (1998). "Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei." *Mol Biol Cell* **9**(6): 1523-36.
- Jackson, S. P. and R. Tjian (1988). "O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation." *Cell* **55**(1): 125-33.
- Johnson, W. E., W. Li, et al. (2006). "Model-based analysis of tiling-arrays for ChIP-chip." *Proc Natl Acad Sci U S A* **103**(33): 12457-62.
- Juven-Gershon, T., J. Y. Hsu, et al. (2008). "The RNA polymerase II core promoter - the gateway to transcription." *Curr Opin Cell Biol* **20**(3): 253-9.
- Kadonaga, J. T. (2002). "The DPE, a core promoter element for transcription by RNA polymerase II." *Exp Mol Med* **34**(4): 259-64.
- Kadonaga, J. T. (2004). "Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors." *Cell* **116**(2): 247-57.
- Kaltenbach, L., M. A. Horner, et al. (2000). "The TBP-like factor CeTLF is required to activate RNA polymerase II transcription during *C. elegans* embryogenesis." *Mol Cell* **6**(3): 705-13.
- Kangaspeska, S., B. Stride, et al. (2008). "Transient cyclical methylation of promoter DNA." *Nature* **452**(7183): 112-5.
- Kaplan, J. and K. Calame (1997). "The ZiN/POZ domain of ZF5 is required for both transcriptional activation and repression." *Nucleic Acids Res* **25**(6): 1108-16.
- Katsani, K. R., M. A. Hajibagheri, et al. (1999). "Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology." *EMBO J* **18**(3): 698-708.
- Kerrigan, L. A., G. E. Croston, et al. (1991). "Sequence-specific transcriptional antirepression of the *Drosophila* Kruppel gene by the GAGA factor." *J Biol Chem* **266**(1): 574-82.
- Kershner, E., S. Y. Wu, et al. (1998). "Immunoaffinity purification and functional characterization of human transcription factor IIH and RNA polymerase II from clonal cell lines that conditionally express epitope-tagged subunits of the multiprotein complexes." *J Biol Chem* **273**(51): 34444-53.
- Kim, M. H., J. S. Shin, et al. (2002). "Retinoic acid response element in HOXA-7 regulatory region affects the rate, not the formation of anterior boundary expression." *Int J Dev Biol* **46**(3): 325-8.
- Kim, T. K., R. H. Ebright, et al. (2000). "Mechanism of ATP-dependent promoter melting by transcription factor IIH." *Science* **288**(5470): 1418-22.
- Kim, Y. J., S. Bjorklund, et al. (1994). "A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II." *Cell* **77**(4): 599-608.
- Kimura, H., Y. Tao, et al. (1999). "Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure." *Mol Cell Biol* **19**(8): 5383-92.
- Kimura, K., A. Wakamatsu, et al. (2006). "Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes." *Genome Res* **16**(1): 55-65.
- Kingston, R. E. and G. J. Narlikar (1999). "ATP-dependent remodeling and acetylation as regulators of chromatin fluidity." *Genes Dev* **13**(18): 2339-52.
- Kireeva, M. L., B. Hancock, et al. (2005). "Nature of the nucleosomal barrier to RNA polymerase II." *Mol Cell* **18**(1): 97-108.
- Kireeva, M. L., W. Walter, et al. (2002). "Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription." *Mol Cell* **9**(3): 541-52.
- Klug, A. and J. W. Schwabe (1995). "Protein motifs 5. Zinc fingers." *FASEB J* **9**(8): 597-604.

- Knoepfler, P. S. and R. N. Eisenman (1999). "Sin meets NuRD and other tails of repression." *Cell* **99**(5): 447-50.
- Koch, F., F. Jourquin, et al. (2008). "Genome-wide RNA polymerase II: not genes only!" *Trends Biochem Sci* **33**(6): 265-73.
- Koleske, A. J. and R. A. Young (1994). "An RNA polymerase II holoenzyme responsive to activators." *Nature* **368**(6470): 466-9.
- Komarnitsky, P., E. J. Cho, et al. (2000). "Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription." *Genes Dev* **14**(19): 2452-60.
- Kornberg, R. D. (2007). "The molecular basis of eukaryotic transcription." *Proc Natl Acad Sci U S A* **104**(32): 12955-61.
- Kosoy, A., S. Pagans, et al. (2002). "GAGA factor down-regulates its own promoter." *J Biol Chem* **277**(44): 42280-8.
- Krishnamurthy, S., X. He, et al. (2004). "Ssu72 Is an RNA polymerase II CTD phosphatase." *Mol Cell* **14**(3): 387-94.
- Kuhn, E. J. and P. K. Geyer (2003). "Genomic insulators: connecting properties to mechanism." *Curr Opin Cell Biol* **15**(3): 259-65.
- Kumar, K. P., S. Akoulitchev, et al. (1998). "Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event." *Proc Natl Acad Sci U S A* **95**(17): 9767-72.
- Kunert, N., J. Marhold, et al. (2003). "A Dnmt2-like protein mediates DNA methylation in Drosophila." *Development* **130**(21): 5083-90.
- Kurdistani, S. K., S. Tavazoie, et al. (2004). "Mapping global histone acetylation patterns to gene expression." *Cell* **117**(6): 721-33.
- Kutach, A. K. and J. T. Kadonaga (2000). "The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters." *Mol Cell Biol* **20**(13): 4754-64.
- Lagrange, T., A. N. Kapanidis, et al. (1998). "New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB." *Genes Dev* **12**(1): 34-44.
- Lagrange, T., T. K. Kim, et al. (1996). "High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex." *Proc Natl Acad Sci U S A* **93**(20): 10620-5.
- Lee, C., X. Li, et al. (2008). "NELF and GAGA factor are linked to promoter-proximal pausing at many genes in Drosophila." *Mol Cell Biol* **28**(10): 3290-300.
- Lee, D. H., N. Gershenzon, et al. (2005). "Functional characterization of core promoter elements: the downstream core element is recognized by TAF1." *Mol Cell Biol* **25**(21): 9674-86.
- Lee, D. K., J. DeJong, et al. (1992a). "TFIIA induces conformational changes in TFIID via interactions with the basic repeat." *Mol Cell Biol* **12**(11): 5189-96.
- Lee, H., K. W. Kraus, et al. (1992b). "DNA sequence requirements for generating paused polymerase at the start of hsp70." *Genes Dev* **6**(2): 284-95.
- Lee, Y. S. and R. W. Carthew (2003). "Making a better RNAi vector for Drosophila: use of intron spacers." *Methods* **30**(4): 322-9.
- Lehmann, M. (2004). "Anything else but GAGA: a nonhistone protein complex reshapes chromatin structure." *Trends Genet* **20**(1): 15-22.
- Lehmann, M., T. Siegmund, et al. (1998). "The pipsqueak protein of Drosophila melanogaster binds to GAGA sequences through a novel DNA-binding domain." *J Biol Chem* **273**(43): 28504-9.
- Lehming, N., D. Thanos, et al. (1994). "An HMG-like protein that can switch a transcriptional activator to a repressor." *Nature* **371**(6493): 175-9.
- Lemon, B. and R. Tjian (2000). "Orchestrated response: a symphony of transcription factors for gene control." *Genes Dev* **14**(20): 2551-69.

BIBLIOGRAFIA

- Lewis, B. A., T. K. Kim, et al. (2000). "A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts." *Proc Natl Acad Sci U S A* **97**(13): 7172-7.
- Lewis, E. B. (1978). "A gene complex controlling segmentation in *Drosophila*." *Nature* **276**(5688): 565-70.
- Li, B., M. Carey, et al. (2007). "The role of chromatin during transcription." *Cell* **128**(4): 707-19.
- Lim, C. Y., B. Santoso, et al. (2004). "The MTE, a new core promoter element for transcription by RNA polymerase II." *Genes Dev* **18**(13): 1606-17.
- Lin, Y. C., W. S. Choi, et al. (2005). "TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape." *Nat Struct Mol Biol* **12**(7): 603-7.
- Lintermann, K. G., G. E. Roth, et al. (1998). "Comparison of the GAGA factor genes of *Drosophila melanogaster* and *Drosophila virilis* reveals high conservation of GAGA factor structure beyond the BTB/POZ and DNA-binding domains." *Dev Genes Evol* **208**(8): 447-56.
- Lu, H., R. P. Fisher, et al. (1997). "The CDK7-cycH-p36 complex of transcription factor IIH phosphorylates p53, enhancing its sequence-specific DNA binding activity in vitro." *Mol Cell Biol* **17**(10): 5923-34.
- Lu, H., L. Zawel, et al. (1992). "Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II." *Nature* **358**(6388): 641-5.
- Lu, Q., L. L. Wallrath, et al. (1993). "(CT)_n (GA)_n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the *Drosophila* hsp26 gene." *Mol Cell Biol* **13**(5): 2802-14.
- Luger, K., A. W. Mader, et al. (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution." *Nature* **389**(6648): 251-60.
- Lyko, F., B. H. Ramsahoye, et al. (2000). "DNA methylation in *Drosophila melanogaster*." *Nature* **408**(6812): 538-40.
- Mahmoudi, T., K. R. Katsani, et al. (2002). "GAGA can mediate enhancer function in trans by linking two separate DNA molecules." *EMBO J* **21**(7): 1775-81.
- Maldonado, E., R. Shiekhattar, et al. (1996). "A human RNA polymerase II complex associated with SRB and DNA-repair proteins." *Nature* **381**(6577): 86-9.
- Manley, J. L., A. Fire, et al. (1980). "DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract." *Proc Natl Acad Sci U S A* **77**(7): 3855-9.
- Martinez-Campa, C., P. Politis, et al. (2004). "Precise nucleosome positioning and the TATA box dictate requirements for the histone H4 tail and the bromodomain factor Bdf1." *Mol Cell* **15**(1): 69-81.
- Martinez, E., H. Ge, et al. (1998). "Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex." *Mol Cell Biol* **18**(11): 6571-83.
- Mason, P. B., Jr. and J. T. Lis (1997). "Cooperative and competitive protein interactions at the hsp70 promoter." *J Biol Chem* **272**(52): 33227-33.
- Matsui, T., J. Segall, et al. (1980). "Multiple factors required for accurate initiation of transcription by purified RNA polymerase II." *J Biol Chem* **255**(24): 11992-6.
- Matsukage, A., F. Hirose, et al. (2008). "The DRE/DREF transcriptional regulatory system: a master key for cell proliferation." *Biochim Biophys Acta* **1779**(2): 81-9.
- Maxon, M. E., J. A. Goodrich, et al. (1994). "Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance." *Genes Dev* **8**(5): 515-24.
- McCracken, S. and J. Greenblatt (1991). "Related RNA polymerase-binding regions in human RAP30/74 and *Escherichia coli* sigma 70." *Science* **253**(5022): 900-2.
- Meisterernst, M. and R. G. Roeder (1991). "Family of proteins that interact with TFIID and regulate promoter activity." *Cell* **67**(3): 557-67.
- Melfi, R., F. Palla, et al. (2000). "Functional characterization of the enhancer blocking element of the sea urchin early histone gene cluster reveals insulator properties and three essential cis-acting sequences." *J Mol Biol* **304**(5): 753-63.
- Melnikova, L., F. Juge, et al. (2004). "Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*." *Proc Natl Acad Sci U S A* **101**(41): 14806-11.

- Mellon, I. and P. C. Hanawalt (1989). "Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand." *Nature* **342**(6245): 95-8.
- Mihaly, J., I. Hogga, et al. (1998). "Chromatin domain boundaries in the Bithorax complex." *Cell Mol Life Sci* **54**(1): 60-70.
- Mishra, K., V. S. Chopra, et al. (2003). "Trl-GAGA directly interacts with lola like and both are part of the repressive complex of Polycomb group of genes." *Mech Dev* **120**(6): 681-9.
- Mitsiou, D. J. and H. G. Stunnenberg (2000). "TAC, a TBP-sans-TAFs complex containing the unprocessed TFIIAalpha precursor and the TFIIBgamma subunit." *Mol Cell* **6**(3): 527-37.
- Mizuguchi, G., X. Shen, et al. (2004). "ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex." *Science* **303**(5656): 343-8.
- Moore, P. A., J. Ozer, et al. (1999). "A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators." *Mol Cell Biol* **19**(11): 7610-20.
- Müller, P., G. Parmigiani, et al. (2004). "Optimal sample size for multiple testing: the case of gene expression microarrays." *Journal of the American Statistical Association* **99**: 990-1001.
- Muller, F., M. A. Demeny, et al. (2007). "New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors." *J Biol Chem* **282**(20): 14685-9.
- Muller, J. and J. A. Kassis (2006). "Polycomb response elements and targeting of Polycomb group proteins in Drosophila." *Curr Opin Genet Dev* **16**(5): 476-84.
- Muse, G. W., D. A. Gilchrist, et al. (2007). "RNA polymerase is poised for activation across the genome." *Nat Genet* **39**(12): 1507-11.
- Naar, A. M., D. J. Taatjes, et al. (2002). "Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation." *Genes Dev* **16**(11): 1339-44.
- Nakayama, T., K. Nishioka, et al. (2007). "Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading." *Genes Dev* **21**(5): 552-61.
- Negre, N., J. Hennetin, et al. (2006). "Chromosomal distribution of PcG proteins during Drosophila development." *PLoS Biol* **4**(6): e170.
- Ni, Z., B. E. Schwartz, et al. (2004). "Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes." *Mol Cell* **13**(1): 55-65.
- Nibu, Y., H. Zhang, et al. (2001). "Local action of long-range repressors in the Drosophila embryo." *EMBO J* **20**(9): 2246-53.
- Norwood, L. E., S. K. Grade, et al. (2004). "Conserved properties of HP1(Hsalpha)." *Gene* **336**(1): 37-46.
- Novina, C. D. and A. L. Roy (1996). "Core promoters and transcriptional control." *Trends Genet* **12**(9): 351-5.
- O'Brien, T., R. C. Wilkins, et al. (1995). "Distribution of GAGA protein on Drosophila genes in vivo." *Genes Dev* **9**(9): 1098-110.
- O'Donnell, K. H., C. T. Chen, et al. (1994). "Insulating DNA directs ubiquitous transcription of the Drosophila melanogaster alpha 1-tubulin gene." *Mol Cell Biol* **14**(9): 6398-408.
- O'Malley, B. W. (2007). "Coregulators: from whence came these "master genes"." *Mol Endocrinol* **21**(5): 1009-13.
- Oelgeschlager, T. (2002). "Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control." *J Cell Physiol* **190**(2): 160-9.
- Ohbayashi, T., Y. Makino, et al. (1999). "Identification of a mouse TBP-like protein (TLP) distantly related to the drosophila TBP-related factor." *Nucleic Acids Res* **27**(3): 750-5.
- Ohkuma, Y., S. Hashimoto, et al. (1995). "Analysis of the role of TFIIE in basal transcription and TFIIE-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha." *Mol Cell Biol* **15**(9): 4856-66.
- Ohkuma, Y., M. Horikoshi, et al. (1990a). "Engrailed, a homeodomain protein, can repress in vitro transcription by competition with the TATA box-binding protein transcription factor IID." *Proc Natl Acad Sci U S A* **87**(6): 2289-93.

BIBLIOGRAFIA

- Ohkuma, Y., H. Sumimoto, et al. (1990b). "Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIE." Proc Natl Acad Sci U S A **87**(23): 9163-7.
- Ohler, U. (2006). "Identification of core promoter modules in Drosophila and their application in accurate transcription start site prediction." Nucleic Acids Res **34**(20): 5943-50.
- Ohler, U., G. C. Liao, et al. (2002). "Computational analysis of core promoters in the Drosophila genome." Genome Biol **3**(12): RESEARCH0087.
- Ohshima, N., M. Takahashi, et al. (2003). "Identification of a human homologue of the DREF transcription factor with a potential role in regulation of the histone H1 gene." J Biol Chem **278**(25): 22928-38.
- Ohtsuki, S. and M. Levine (1998). "GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo." Genes Dev **12**(21): 3325-30.
- Okada, M. and S. Hirose (1998). "Chromatin remodeling mediated by Drosophila GAGA factor and ISWI activates fushi tarazu gene transcription in vitro." Mol Cell Biol **18**(5): 2455-61.
- Okamoto, T., S. Yamamoto, et al. (1998). "Analysis of the role of TFIIE in transcriptional regulation through structure-function studies of the TFIIEbeta subunit." J Biol Chem **273**(31): 19866-76.
- Omichinski, J. G., P. V. Pedone, et al. (1997). "The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode." Nat Struct Biol **4**(2): 122-32.
- Onodera, Y., J. R. Haag, et al. (2005). "Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation." Cell **120**(5): 613-22.
- Orphanides, G., T. Lagrange, et al. (1996). "The general transcription factors of RNA polymerase II." Genes Dev **10**(21): 2657-83.
- Orphanides, G. and D. Reinberg (2002). "A unified theory of gene expression." Cell **108**(4): 439-51.
- Osborne, C. S., L. Chakalova, et al. (2004). "Active genes dynamically colocalize to shared sites of ongoing transcription." Nat Genet **36**(10): 1065-71.
- Ossipow, V., J. P. Tassan, et al. (1995). "A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation." Cell **83**(1): 137-46.
- Otero, G., J. Fellows, et al. (1999). "Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation." Mol Cell **3**(1): 109-18.
- Ozer, J., P. A. Moore, et al. (1994). "Molecular cloning of the small (gamma) subunit of human TFIIA reveals functions critical for activated transcription." Genes Dev **8**(19): 2324-35.
- Pagans, S., M. Ortiz-Lombardia, et al. (2002). "The Drosophila transcription factor tramtrack (TTK) interacts with Trithorax-like (GAGA) and represses GAGA-mediated activation." Nucleic Acids Res **30**(20): 4406-13.
- Pagans, S., D. Pineyro, et al. (2004). "Repression by TTK69 of GAGA-mediated activation occurs in the absence of TTK69 binding to DNA and solely requires the contribution of the POZ/BTB domain of TTK69." J Biol Chem **279**(11): 9725-32.
- Palenchar, J. B., W. Liu, et al. (2006). "A divergent transcription factor TFIIB in trypanosomes is required for RNA polymerase II-dependent spliced leader RNA transcription and cell viability." Eukaryot Cell **5**(2): 293-300.
- Park, Y. J., J. V. Chodaparambil, et al. (2005). "Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding." J Biol Chem **280**(3): 1817-25.
- Pedone, P. V., R. Ghirlando, et al. (1996). "The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding." Proc Natl Acad Sci U S A **93**(7): 2822-6.
- Pei, Y., B. Schwer, et al. (2003). "Interactions between fission yeast Cdk9, its cyclin partner Pch1, and mRNA capping enzyme Pct1 suggest an elongation checkpoint for mRNA quality control." J Biol Chem **278**(9): 7180-8.
- Persengiev, S. P., X. Zhu, et al. (2003). "TRF3, a TATA-box-binding protein-related factor, is vertebrate-specific and widely expressed." Proc Natl Acad Sci U S A **100**(25): 14887-91.

- Peterlin, B. M. and D. H. Price (2006). "Controlling the elongation phase of transcription with P-TEFb." Mol Cell **23**(3): 297-305.
- Phatnani, H. P. and A. L. Greenleaf (2006). "Phosphorylation and functions of the RNA polymerase II CTD." Genes Dev **20**(21): 2922-36.
- Pile, L. A. and I. L. Cartwright (2000). "GAGA factor-dependent transcription and establishment of DNase hypersensitivity are independent and unrelated events in vivo." J Biol Chem **275**(2): 1398-404.
- Pinto, I., D. E. Ware, et al. (1992). "The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo." Cell **68**(5): 977-88.
- Platero, J. S., A. K. Csink, et al. (1998). "Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in Drosophila." J Cell Biol **140**(6): 1297-306.
- Pointud, J. C., J. Larsson, et al. (2001). "The BTB/POZ domain of the regulatory proteins Bric a brac 1 (BAB1) and Bric a brac 2 (BAB2) interacts with the novel Drosophila TAF(II) factor BIP2/dTAF(II)155." Dev Biol **237**(2): 368-80.
- Pokholok, D. K., N. M. Hannett, et al. (2002). "Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo." Mol Cell **9**(4): 799-809.
- Pombo, A., D. A. Jackson, et al. (1999). "Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III." EMBO J **18**(8): 2241-53.
- Poortinga, G., M. Watanabe, et al. (1998). "Drosophila CtBP: a Hairy-interacting protein required for embryonic segmentation and hairy-mediated transcriptional repression." EMBO J **17**(7): 2067-78.
- Poux, S., B. Horard, et al. (2002). "The Drosophila trithorax protein is a coactivator required to prevent re-establishment of polycomb silencing." Development **129**(10): 2483-93.
- Price, D. H., A. E. Sluder, et al. (1989). "Dynamic interaction between a Drosophila transcription factor and RNA polymerase II." Mol Cell Biol **9**(4): 1465-75.
- Proudfoot, N. J., A. Furger, et al. (2002). "Integrating mRNA processing with transcription." Cell **108**(4): 501-12.
- Radonjic, M., J. C. Andrau, et al. (2005). "Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon *S. cerevisiae* stationary phase exit." Mol Cell **18**(2): 171-83.
- Ranish, J. A., S. Hahn, et al. (2004). "Identification of TFB5, a new component of general transcription and DNA repair factor IIH." Nat Genet **36**(7): 707-13.
- Rappailles, A., M. Decoville, et al. (2005). "DSP1, a Drosophila HMG protein, is involved in spatiotemporal expression of the homoeotic gene *Sex combs reduced*." Biol Cell **97**(10): 779-85.
- Reinberg, D., M. Horikoshi, et al. (1987a). "Factors involved in specific transcription in mammalian RNA polymerase II. Functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site." J Biol Chem **262**(7): 3322-30.
- Reinberg, D. and R. G. Roeder (1987b). "Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE." J Biol Chem **262**(7): 3310-21.
- Reinke, H. and W. Horz (2003). "Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter." Mol Cell **11**(6): 1599-607.
- Reyes, J. C., C. Muchardt, et al. (1997). "Components of the human SWI/SNF complex are enriched in active chromatin and are associated with the nuclear matrix." J Cell Biol **137**(2): 263-74.
- Richards, E. J. and S. C. Elgin (2002). "Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects." Cell **108**(4): 489-500.
- Robert, F., M. Douziech, et al. (1998). "Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF." Mol Cell **2**(3): 341-51.
- Rodriguez, C. R., E. J. Cho, et al. (2000). "Kin28, the TFIIF-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II." Mol Cell Biol **20**(1): 104-12.

BIBLIOGRAFIA

- Roeder, R. G. (1998). "Role of general and gene-specific cofactors in the regulation of eukaryotic transcription." *Cold Spring Harb Symp Quant Biol* **63**: 201-18.
- Rossell, D., R. Guerra, et al. (2008). "Semi-parametric differential expression analysis via partial mixture estimation." *Stat Appl Genet Mol Biol* **7**(1): Article15.
- Rougvie, A. E. and J. T. Lis (1988). "The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged." *Cell* **54**(6): 795-804.
- Roy, A. L., S. Malik, et al. (1993). "An alternative pathway for transcription initiation involving TFII-I." *Nature* **365**(6444): 355-9.
- Roy, A. L., M. Meisterernst, et al. (1991). "Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF." *Nature* **354**(6350): 245-8.
- Roy, R., L. Schaeffer, et al. (1994). "The DNA-dependent ATPase activity associated with the class II basic transcription factor BTF2/TFIIH." *J Biol Chem* **269**(13): 9826-32.
- Ruan, J. P., G. K. Arhin, et al. (2004). "Functional characterization of a *Trypanosoma brucei* TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes." *Mol Cell Biol* **24**(21): 9610-8.
- Sadowski, I., J. Ma, et al. (1988). "GAL4-VP16 is an unusually potent transcriptional activator." *Nature* **335**(6190): 563-4.
- Saha, A., J. Wittmeyer, et al. (2006). "Chromatin remodelling: the industrial revolution of DNA around histones." *Nat Rev Mol Cell Biol* **7**(6): 437-47.
- Salvaing, J., A. Lopez, et al. (2003). "The *Drosophila* Corto protein interacts with Polycomb-group proteins and the GAGA factor." *Nucleic Acids Res* **31**(11): 2873-82.
- Sambrook, J. and D. Russell, Eds. (1989). *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- Sandelin, A., P. Carninci, et al. (2007). "Mammalian RNA polymerase II core promoters: insights from genome-wide studies." *Nat Rev Genet* **8**(6): 424-36.
- Sangwan, I. and M. R. O'Brian (2002). "Identification of a soybean protein that interacts with GAGA element dinucleotide repeat DNA." *Plant Physiol* **129**(4): 1788-94.
- Santi, L., Y. Wang, et al. (2003). "The GA octodinucleotide repeat binding factor BBR participates in the transcriptional regulation of the homeobox gene *Bkn3*." *Plant J* **34**(6): 813-26.
- Saunders, A., L. J. Core, et al. (2006). "Breaking barriers to transcription elongation." *Nat Rev Mol Cell Biol* **7**(8): 557-67.
- Saxonov, S., P. Berg, et al. (2006). "A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters." *Proc Natl Acad Sci U S A* **103**(5): 1412-7.
- Scacheri, P. C., G. E. Crawford, et al. (2006). "Statistics for ChIP-chip and DNase hypersensitivity experiments on NimbleGen arrays." *Methods Enzymol* **411**: 270-82.
- Schaeffer, L., V. Moncollin, et al. (1994). "The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor." *EMBO J* **13**(10): 2388-92.
- Schaeffer, L., R. Roy, et al. (1993). "DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor." *Science* **260**(5104): 58-63.
- Schaeper, U., J. M. Boyd, et al. (1995). "Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation." *Proc Natl Acad Sci U S A* **92**(23): 10467-71.
- Scherer, W. F., J. T. Syverton, et al. (1953). "Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix." *J Exp Med* **97**(5): 695-710.
- Schneider, I. (1972). "Cell lines derived from late embryonic stages of *Drosophila melanogaster*." *J Embryol Exp Morphol* **27**(2): 353-65.
- Schneider, R. and R. Grosschedl (2007). "Dynamics and interplay of nuclear architecture, genome organization, and gene expression." *Genes Dev* **21**(23): 3027-43.
- Schroeder, S. C., B. Schwer, et al. (2000). "Dynamic association of capping enzymes with transcribing RNA polymerase II." *Genes Dev* **14**(19): 2435-40.

- Schuettengruber, B., M. Ganapathi, et al. (2009). "Functional anatomy of polycomb and trithorax chromatin landscapes in *Drosophila* embryos." *PLoS Biol* **7**(1): e13.
- Schug, J., W. P. Schuller, et al. (2005). "Promoter features related to tissue specificity as measured by Shannon entropy." *Genome Biol* **6**(4): R33.
- Schulze, S. R. and L. L. Wallrath (2007). "Gene regulation by chromatin structure: paradigms established in *Drosophila melanogaster*." *Annu Rev Entomol* **52**: 171-92.
- Schwartz, B. E. and K. Ahmad (2005). "Transcriptional activation triggers deposition and removal of the histone variant H3.3." *Genes Dev* **19**(7): 804-14.
- Schweinsberg, S., K. Hagstrom, et al. (2004). "The enhancer-blocking activity of the Fab-7 boundary from the *Drosophila* bithorax complex requires GAGA-factor-binding sites." *Genetics* **168**(3): 1371-84.
- Schwendemann, A. and M. Lehmann (2002). "Pipsqueak and GAGA factor act in concert as partners at homeotic and many other loci." *Proc Natl Acad Sci U S A* **99**(20): 12883-8.
- Serizawa, H., R. C. Conaway, et al. (1992). "A carboxyl-terminal-domain kinase associated with RNA polymerase II transcription factor delta from rat liver." *Proc Natl Acad Sci U S A* **89**(16): 7476-80.
- Seto, E., Y. Shi, et al. (1991). "YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro." *Nature* **354**(6350): 241-5.
- Shao, H., M. Revach, et al. (2005). "Core promoter binding by histone-like TAF complexes." *Mol Cell Biol* **25**(1): 206-19.
- Shilatifard, A., R. C. Conaway, et al. (2003). "The RNA polymerase II elongation complex." *Annu Rev Biochem* **72**: 693-715.
- Shimojima, T., M. Okada, et al. (2003). "*Drosophila* FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor." *Genes Dev* **17**(13): 1605-16.
- Shogren-Knaak, M., H. Ishii, et al. (2006). "Histone H4-K16 acetylation controls chromatin structure and protein interactions." *Science* **311**(5762): 844-7.
- Shopland, L. S., K. Hirayoshi, et al. (1995). "HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites." *Genes Dev* **9**(22): 2756-69.
- Siegfried, Z., S. Eden, et al. (1999). "DNA methylation represses transcription in vivo." *Nat Genet* **22**(2): 203-6.
- Simon, J. A. and J. W. Tamkun (2002). "Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes." *Curr Opin Genet Dev* **12**(2): 210-8.
- Simonis, M., P. Klous, et al. (2006). "Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C)." *Nat Genet* **38**(11): 1348-54.
- Sinclair, D. A., S. Schulze, et al. (2000). "Essential genes in autosomal heterochromatin of *Drosophila melanogaster*." *Genetica* **109**(1-2): 9-18.
- Singer, V. L., C. R. Wobbe, et al. (1990). "A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation." *Genes Dev* **4**(4): 636-45.
- Smale, S. T. and J. T. Kadonaga (2003). "The RNA polymerase II core promoter." *Annu Rev Biochem* **72**: 449-79.
- Smyth, G. K. (2004). "Linear models and empirical bayes methods for assessing differential expression in microarray experiments." *Stat Appl Genet Mol Biol* **3**: Article3.
- Soeller, W. C., C. E. Oh, et al. (1993). "Isolation of cDNAs encoding the *Drosophila* GAGA transcription factor." *Mol Cell Biol* **13**(12): 7961-70.
- Soeller, W. C., S. J. Poole, et al. (1988). "In vitro transcription of the *Drosophila* engrailed gene." *Genes Dev* **2**(1): 68-81.
- Srinivasula, S. M., P. Datta, et al. (2002). "sickle, a novel *Drosophila* death gene in the reaper/hid/grim region, encodes an IAP-inhibitory protein." *Curr Biol* **12**(2): 125-30.
- Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." *Nature* **403**(6765): 41-5.

BIBLIOGRAFIA

- Strutt, H., G. Cavalli, et al. (1997). "Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression." *EMBO J* **16**(12): 3621-32.
- Sullivan, W., M. Ashburner, et al., Eds. (2000). *Drosophila Protocols*, Cold Spring Harbor Laboratory Press.
- Sun, L. V., L. Chen, et al. (2003). "Protein-DNA interaction mapping using genomic tiling path microarrays in *Drosophila*." *Proc Natl Acad Sci U S A* **100**(16): 9428-33.
- Svejstrup, J. Q. (2007). "Contending with transcriptional arrest during RNAPII transcript elongation." *Trends Biochem Sci* **32**(4): 165-71.
- Svejstrup, J. Q., Y. Li, et al. (1997). "Evidence for a mediator cycle at the initiation of transcription." *Proc Natl Acad Sci U S A* **94**(12): 6075-8.
- Svejstrup, J. Q., P. Vichi, et al. (1996). "The multiple roles of transcription/repair factor TFIIH." *Trends Biochem Sci* **21**(9): 346-50.
- Takada, S., J. T. Lis, et al. (2000). "A TRF1:BRF complex directs *Drosophila* RNA polymerase III transcription." *Cell* **101**(5): 459-69.
- Thomas, M. C. and C. M. Chiang (2006). "The general transcription machinery and general cofactors." *Crit Rev Biochem Mol Biol* **41**(3): 105-78.
- Tokusumi, Y., Y. Ma, et al. (2007). "The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters." *Mol Cell Biol* **27**(5): 1844-58.
- Tora, L. (2002). "A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription." *Genes Dev* **16**(6): 673-5.
- Trigon, S., H. Serizawa, et al. (1998). "Characterization of the residues phosphorylated in vitro by different C-terminal domain kinases." *J Biol Chem* **273**(12): 6769-75.
- Tsukiyama, T., P. B. Becker, et al. (1994). "ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor." *Nature* **367**(6463): 525-32.
- Tsukiyama, T., C. Daniel, et al. (1995a). "ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor." *Cell* **83**(6): 1021-6.
- Tsukiyama, T. and C. Wu (1995b). "Purification and properties of an ATP-dependent nucleosome remodeling factor." *Cell* **83**(6): 1011-20.
- Usheva, A., E. Maldonado, et al. (1992). "Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein." *Cell* **69**(5): 871-81.
- Valentine, S. A., G. Chen, et al. (1998). "Dorsal-mediated repression requires the formation of a multiprotein repression complex at the ventral silencer." *Mol Cell Biol* **18**(11): 6584-94.
- Van Dyke, M. W., M. Sawadogo, et al. (1989). "Stability of transcription complexes on class II genes." *Mol Cell Biol* **9**(1): 342-4.
- van Steensel, B., J. Delrow, et al. (2003). "Genomewide analysis of *Drosophila* GAGA factor target genes reveals context-dependent DNA binding." *Proc Natl Acad Sci U S A* **100**(5): 2580-5.
- Vaquero, A., M. Blanch, et al. (2008). "Activation properties of GAGA transcription factor." *Biochim Biophys Acta* **1779**(5): 312-7.
- Vaquero, A., M. L. Espinas, et al. (2000). "Functional mapping of the GAGA factor assigns its transcriptional activity to the C-terminal glutamine-rich domain." *J Biol Chem* **275**(26): 19461-8.
- Vaughn, M. W. and R. A. Martienssen (2005). "Finding the right template: RNA Pol IV, a plant-specific RNA polymerase." *Mol Cell* **17**(6): 754-6.
- Veenstra, G. J., D. L. Weeks, et al. (2000). "Distinct roles for TBP and TBP-like factor in early embryonic gene transcription in *Xenopus*." *Science* **290**(5500): 2312-5.
- Veraksa, A., A. Bauer, et al. (2005). "Analyzing protein complexes in *Drosophila* with tandem affinity purification-mass spectrometry." *Dev Dyn* **232**(3): 827-34.
- Volpe, T. A., C. Kidner, et al. (2002). "Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi." *Science* **297**(5588): 1833-7.

- Volpi, S., C. Rabadan-Diehl, et al. (2002). "Transcriptional regulation of the pituitary vasopressin V1b receptor involves a GAGA-binding protein." *J Biol Chem* **277**(31): 27829-38.
- Walter, J., C. A. Dever, et al. (1994). "Two homeo domain proteins bind with similar specificity to a wide range of DNA sites in *Drosophila* embryos." *Genes Dev* **8**(14): 1678-92.
- Wall, G., P. D. Varga-Weisz, et al. (1995). "Chromatin remodeling by GAGA factor and heat shock factor at the hypersensitive *Drosophila* hsp26 promoter in vitro." *EMBO J* **14**(8): 1727-36.
- Wampler, S. L. and J. T. Kadonaga (1992). "Functional analysis of *Drosophila* transcription factor IIB." *Genes Dev* **6**(8): 1542-52.
- Wang, Y., M. Jorda, et al. (2006). "Functional CpG methylation system in a social insect." *Science* **314**(5799): 645-7.
- Wang, Y. V., H. Tang, et al. (2005). "Identification in vivo of different rate-limiting steps associated with transcriptional activators in the presence and absence of a GAGA element." *Mol Cell Biol* **25**(9): 3543-52.
- Wang, Z., S. Buratowski, et al. (1995). "The yeast TFB1 and SSL1 genes, which encode subunits of transcription factor IIH, are required for nucleotide excision repair and RNA polymerase II transcription." *Mol Cell Biol* **15**(4): 2288-93.
- Watanabe, T., K. Hayashi, et al. (2003). "The carboxy terminus of the small subunit of TFIIE regulates the transition from transcription initiation to elongation by RNA polymerase II." *Mol Cell Biol* **23**(8): 2914-26.
- Weber, J. A., D. J. Taxman, et al. (1997). "Molecular architecture of the hsp70 promoter after deletion of the TATA box or the upstream regulation region." *Mol Cell Biol* **17**(7): 3799-808.
- Wen, Y., D. Nguyen, et al. (2000). "The N-terminal BTB/POZ domain and C-terminal sequences are essential for Tramtrack69 to specify cell fate in the developing *Drosophila* eye." *Genetics* **156**(1): 195-203.
- Wilkins, R. C. and J. T. Lis (1997). "Dynamics of potentiation and activation: GAGA factor and its role in heat shock gene regulation." *Nucleic Acids Res* **25**(20): 3963-8.
- Wilkins, R. C. and J. T. Lis (1998). "GAGA factor binding to DNA via a single trinucleotide sequence element." *Nucleic Acids Res* **26**(11): 2672-8.
- Wilkins, R. C. and J. T. Lis (1999). "DNA distortion and multimerization: novel functions of the glutamine-rich domain of GAGA factor." *J Mol Biol* **285**(2): 515-25.
- Wing, J. P., J. S. Karres, et al. (2002). "*Drosophila* sickle is a novel grim-reaper cell death activator." *Curr Biol* **12**(2): 131-5.
- Workman, J. L. and R. E. Kingston (1998). "Alteration of nucleosome structure as a mechanism of transcriptional regulation." *Annu Rev Biochem* **67**: 545-79.
- Wu, S. Y. and C. M. Chiang (1998). "Properties of PC4 and an RNA polymerase II complex in directing activated and basal transcription in vitro." *J Biol Chem* **273**(20): 12492-8.
- Xiao, H., R. Sandaltzopoulos, et al. (2001). "Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions." *Mol Cell* **8**(3): 531-43.
- Yamashita, S., K. Wada, et al. (1992). "Isolation and characterization of a cDNA encoding *Drosophila* transcription factor TFIIB." *Proc Natl Acad Sci U S A* **89**(7): 2839-43.
- Yan, Q., R. J. Moreland, et al. (1999). "Dual roles for transcription factor IIF in promoter escape by RNA polymerase II." *J Biol Chem* **274**(50): 35668-75.
- Yang, X., K. Su, et al. (2001). "O-linkage of N-acetylglucosamine to Sp1 activation domain inhibits its transcriptional capability." *Proc Natl Acad Sci U S A* **98**(12): 6611-6.
- Yang, X., F. Zhang, et al. (2002). "Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression." *Cell* **110**(1): 69-80.
- Yokomori, K., A. Admon, et al. (1993). "*Drosophila* TFIIA-L is processed into two subunits that are associated with the TBP/TAF complex." *Genes Dev* **7**(11): 2235-45.
- Young, R. A. (1991). "RNA polymerase II." *Annu Rev Biochem* **60**: 689-715.
- Yuan, G. C., Y. J. Liu, et al. (2005). "Genome-scale identification of nucleosome positions in *S. cerevisiae*." *Science* **309**(5734): 626-30.

BIBLIOGRAFIA

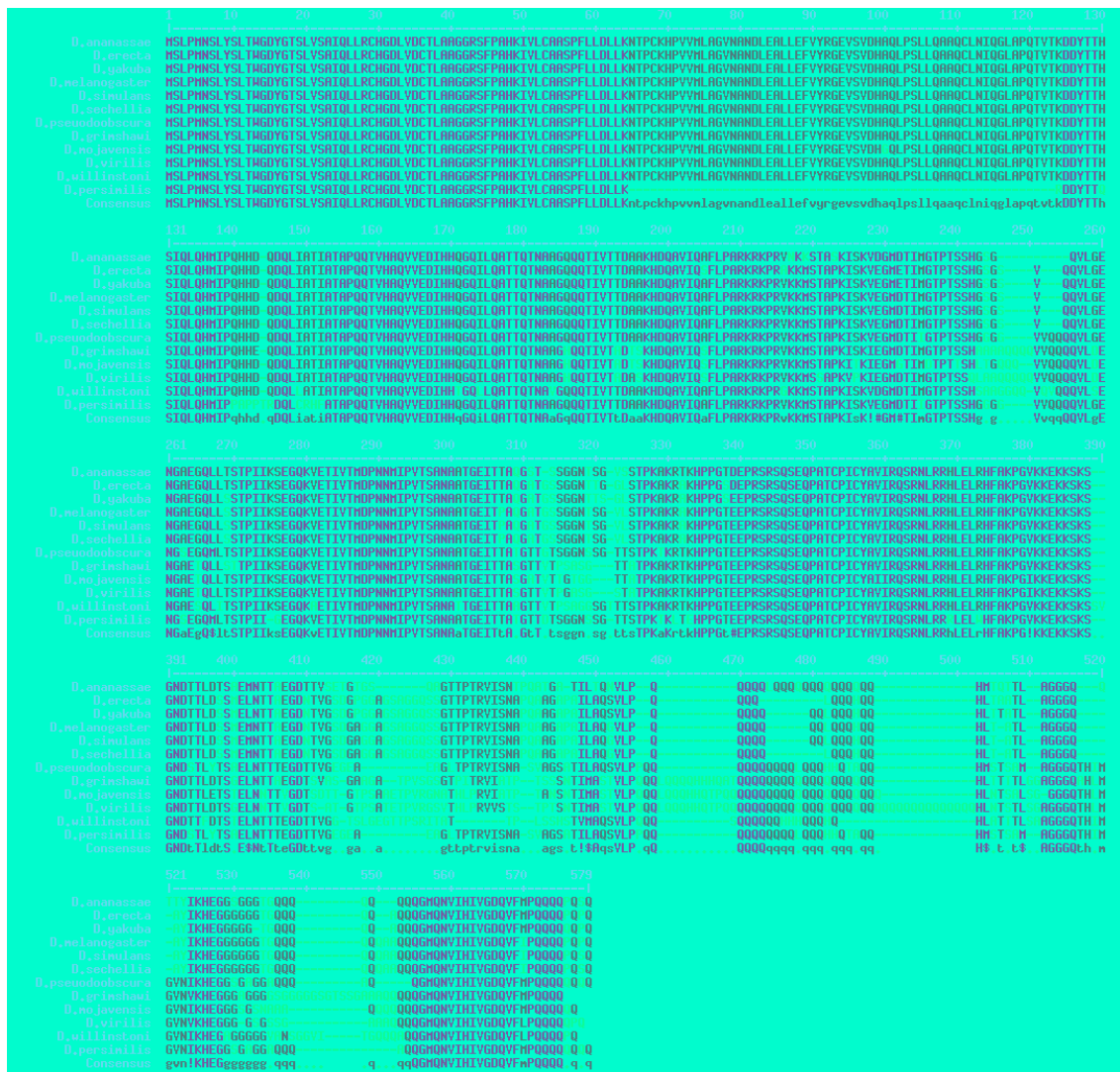
- Yudkovsky, N., J. A. Ranish, et al. (2000). "A transcription reinitiation intermediate that is stabilized by activator." *Nature* **408**(6809): 225-9.
- Yue, Z., E. Maldonado, et al. (1997). "Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II." *Proc Natl Acad Sci U S A* **94**(24): 12898-903.
- Yuryev, A., M. Patturajan, et al. (1996). "The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins." *Proc Natl Acad Sci U S A* **93**(14): 6975-80.
- Zawel, L., K. P. Kumar, et al. (1995). "Recycling of the general transcription factors during RNA polymerase II transcription." *Genes Dev* **9**(12): 1479-90.
- Zeitlinger, J., A. Stark, et al. (2007). "RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo." *Nat Genet* **39**(12): 1512-6.
- Zollman, S., D. Godt, et al. (1994). "The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*." *Proc Natl Acad Sci U S A* **91**(22): 10717-21.

H. ANNEXOS

H.1. Alineaments

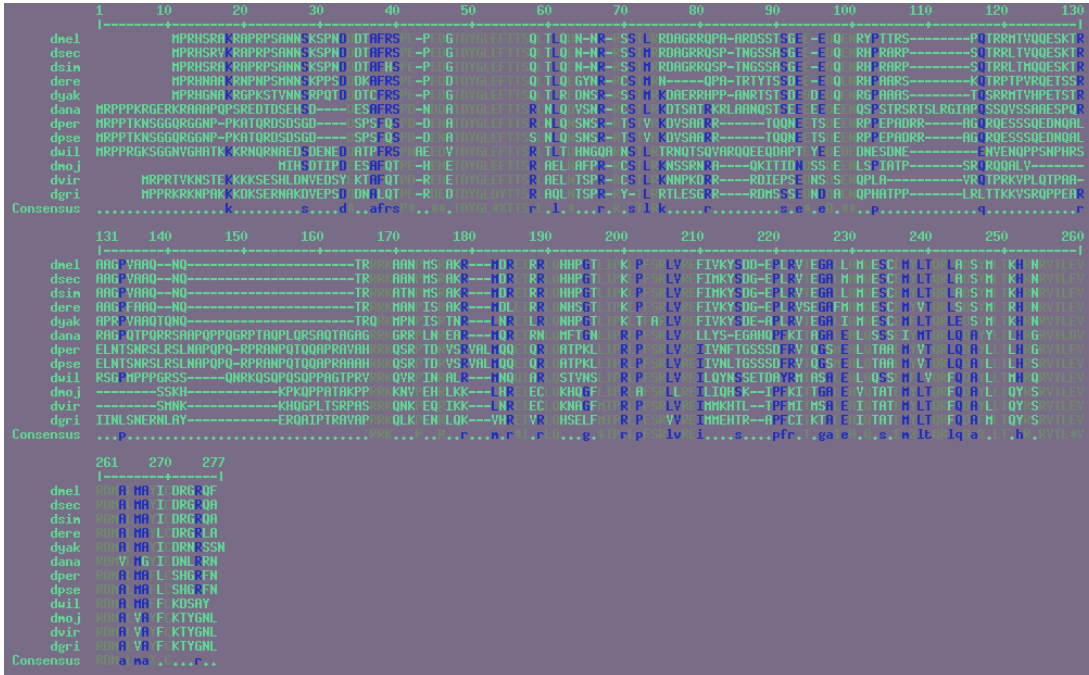
H.1.1. Alineament de GAGA a 12 espècies de *Drosophila*

A la següent imatge es mostra l'alineament de la proteïna GAGA₅₁₉ de *D. melanogaster* amb el seus homòlegs de les 11 espècies restants de *Drosophila* de les que es disposa de la seqüència. L'alineament ha estat realitzat amb el programa MultAlin (disponible *on line* a: <http://bioinfo.genotoul.fr/multalin/multalin.html>), utilitzant els paràmetres per defecte. Cal destacar l'elevat grau de conservació de GAGA a les 12 espècies, extensiu a tota la proteïna, tot i que és menor a la regió que correspon al domini Q.



H.1.2. Alineament de Cid a 12 espècies de *Drosophila*

A la següent imatge es mostra l’alineament de la proteïna Cid de 12 espècies de *Drosophila*. L’alineament ha estat realitzat amb el programa MultAlin, utilitzant els paràmetres per defecte. Com es pot comprovar, la conservació de seqüència és molt inferior que per a GAGA.



H.1.3. Alineament dels promotors de *Trl* de *D. melanogaster* i *D. virilis*

A la següent imatge es mostra l’alineament dels 2979 nucleòtids *upstream* el codó ATG del gen *Trl* de *D. melanogaster* (seqüència superior) i de *D.virilis* (seqüència del mig). A la seqüència inferior de l’alineament apareixen en vermell els nucleòtids que coincideixen, i en blau els que difereixen, entre ambdues espècies. Amb punts apareixen els espais buits (*gaps*) que realitza el programa per maximitzar l’alineament. L’alineament ha estat realitzat amb el programa MultAlin, utilitzant els paràmetres per defecte. Cal destacar l’escassa homologia de seqüència entre els promotors d’ambdues espècies, tret de la zona més propera al codó ATG, que d’altra banda s’ha mostrat prescindible per a la repressió donada per GAGA (veure Discussió, apartat E.1.2). A la imatge també s’han marcat els tres inicis de transcripció més ben caracteritzats al promotor de *D. melanogaster* (indicant el que anomenem +1) i el codó ATG, com a referències.

ANNEXOS

H.1.4. Estructura gènica de *Trl* a 12 espècies de *Drosophila*

Comparació de l'estructura d'exons i introns de la isoforma GAGA₅₁₉ de *D. melanogaster* amb les equivalents d'11 espècies més de *Drosophila*. En números s'indica la mida de cada intró. En verd el més petit i en vermell el més gran.

<i>D. ananassae</i>	exó1-(2172)-exó2-(065)-exó3-(087)-exó4
<i>D. erecta</i>	exó1-(2479)-exó2-(126)-exó3-(158)-exó4
<i>D. grimshawi</i>	exó1-(2851)-exó2-(082)-exó3-(084)-exó4
<i>D. melanogaster</i>	exó1-(2354)-exó2-(122)-exó3-(160)-exó4
<i>D. mojavensis</i>	exó1-(2720)-exó2-(072)-exó3-(074)-exó4
<i>D. pseudoobscura</i>	exó1-(2086)-exó2-(069)-exó3-(074)-exó4
<i>D. persimilis</i>	exó1-(2277)-exó2-(069)-exó3-(074)-exó4
<i>D. sechellia</i>	exó1-(2380)-exó2-(109)-exó3-(160)-exó4
<i>D. simulans</i>	exó1-(2361)-exó2-(214)-exó3-(161)-exó4
<i>D. virilis</i>	exó1-(2593)-exó2-(072)-exó3-(079)-exó4
<i>D. willinstoni</i>	exó1-(2999)-exó2-(076)-exó3-(142)-exó4
<i>D. yakuba</i>	exó1-(2585)-exó2-(112)-exó3-(156)-exó4

H.2. Comparacions de les dades dels experiments de *microarray* i de *ChIP on chip* de GAGA

H.2.1. Comparació del *microarray* de sobreexpressió i el *ChIP on chip* de GAGA

Gens activats al <i>microarray</i> i positius al <i>ChIP on chip</i> de GAGA		
Sondes d'Affymetrix positives al <i>microarray</i>	"DAVID ID" del gen	Símbol del gen
1626509_AT	18 WHEELER	18W
1626839_S_AT	BLOATED TUBULES	BLOT
1641120_AT	CG13506-PA	CG13506
1629085_AT	CG14227-PA	CG14227
1633795_A_AT	CG17124-PB, ISOFORM B	CG17124
1625421_AT	CG30377-PA	CG30377
1639203_AT	CG32556-PA, ISOFORM A	CG32556
1624885_AT	CG32645-PB	CG32645
1624774_A_AT; 1638956_AT	FASCICLIN 2	FAS2
1631231_AT	FRIZZLED	FZ
1632317_AT	GH04850P	CG3036
1623197_A_AT	GH09541P	ZORMIN
1631474_S_AT	GH10595P	L(2)K05713
1627073_A_AT	GH22994P	CG10126
1623957_S_AT	GLUTATHIONE S TRANSFERASE S1	GSTS1
1641530_S_AT	LD30894P	CG2264
1628226_AT	LP11145P	CG8539
1636934_AT	CG32307-PA	CG32307
1624074_AT	ODORANT-BINDING PROTEIN 56A	OBP56A
1630007_S_AT	PHOSPHOLIPASE C AT 21C	PLC21C
1637662_AT	PROTEIN MSTA, ISOFORM A	MSTA
1637957_S_AT	RNA-BINDING PROTEIN 9	RBP9
1636998_AT	SCABROUS	SCA
1629754_S_AT	TRITHORAX-LIKE	CG33260
1640235_AT	WRINKLED	W

Gens reprimits al <i>microarray</i> i positius al <i>ChIP on chip</i> de GAGA		
Sondes d'Affymetrix positives al <i>microarray</i>	"DAVID ID" del gen	Símbol del gen
1636931_at	BROAD	BR

H.2.2. Comparació del *microarray* d'RNAi i el *ChIP on chip* de GAGA

Gens activats al <i>microarray</i> i positius al <i>ChIP on chip</i> de GAGA		
Sondes d'Affymetrix positives al <i>microarray</i>	"DAVID ID" del gen	Símbol del gen
1623466_AT	BRIC A BRAC 1	BAB1
1623718_AT	UNPAIRED 2	UPD2
1624300_S_AT	NEMO	NMO
1624412_AT; 1639498_A_AT	APTEROUS	AP
1624686_A_AT	DUNCE	DNC
1624726_S_AT	GM01289P	CG1745
1624833_A_AT	MUSCARINIC ACETYLCHOLINE RECEPTOR 60C	MACR-60C
1625031_AT	CG7841-PA	CG7841
1625035_AT	INSCUTEABLE	INSC
1625960_AT	NO OCCELLI	NOC
1626749_A_AT	CG18135-PD, ISOFORM D	CG18135
1626821_S_AT	HEAT SHOCK PROTEIN 70BBB	HSP70BBB
1626821_S_AT; 1632841_X_AT	HEAT-SHOCK-PROTEIN-70BA	HSP70BC
1626884_A_AT	CG32062-PB, ISOFORM B	CG32062
1627582_A_AT	CG30035-PA, ISOFORM A	CG30035
1627647_AT	CG3940-PA	CG3940
1627653_AT	GH2758P	LI1 G0469
1627761_AT	PUTATIVE FAT-LIKE CADHERIN-RELATED TUMOR SUPPRESSOR HOMOLOG PRECURSOR	FAT2
1628231_AT	ADENYLYL CYCLASE 76E	AC76E
1628275_AT	TRITHORAX-LIKE	CG33260
1628479_AT	AT09986P	CG15625
1628886_AT	CG4629-PB, ISOFORM B	CG4629
1629207_AT	GH3305P	CG6287
1629701_A_AT	ABIRUPT	AB
1629806_A_AT	CG32103-PC, ISOFORM C	CG32103
1630001_AT; 1635794_A_AT	CG12535-PB, ISOFORM B	CG12535
1630044_S_AT	PUTATIVE GLYCOCEN [STARCH] SYNTHASE	CG6904
1630212_AT	CG2065-PA	CG2065
1630258_AT	GLUTATHIONE S TRANSFERASE D2	GSTD2
1630361_AT	NAKED CUTICLE	NKD
1631095_AT	POSTERIOR SEX COMBS	PSC
1631154_AT	CG13321-PA	CG13321
1631261_AT	LP03706P	CG1806
1631330_AT	REI100P	CG3408
1631474_S_AT	GH10595P	LI2 K05713
1631485_A_AT	ECHINUS	EC
1631620_AT	GLYCOGEN PHOSPHORYLASE	GLYP
1631628_S_AT	CG12896-PA	CG12896
1631628_S_AT	PEROXIREDOXIN 2540	PRX2540-1
1631628_S_AT	PEROXIREDOXIN 2540	PRX2540-2
1631704_AT	CG8505-PA	CG8505
1631993_S_AT; 1633473_S_AT	ALDOLASE	ALD
1632074_AT; 1634948_AT	LUNA	CG13235
1632530_S_AT	CG8776-PD, ISOFORM D	CG8776
1632676_S_AT	CG11887-PA, ISOFORM A	CG11897
1632908_S_AT	LYSOSOMAL ENZYME RECEPTOR PROTEIN	LERP
1633032_S_AT	CG8177-PC, ISOFORM C	CG8177
1633142_AT	CG17681-PA	CG17681
1633866_AT	SUPPRESSOR OF DZITE 2	SUZI2
1634019_AT	SD07613P	CG2064
1634074_A_AT	YIN	YIN
1634129_AT	GH24286P	CG6663
1634325_A_AT	MISEXPRESSION SUPPRESSOR OF KSR 2	MESK2
1634440_S_AT	ECDYSONE-INDUCED PROTEIN 74EF	EIP74EF
1634658_A_AT	NO EXTENDED MEMORY	NEYM
1634995_AT	CG3788-PA	CG3788
1635044_AT	HEAT SHOCK PROTEIN 26	HSP26
1635227_AT	ECDYSONE-INDUCIBLE GENE L3	IMPL3
1635462_AT	RHOMBROID	RHO
1636202_S_AT	ECDYSONE-INDUCIBLE GENE L2	IMPL2
1637129_AT	RE40129P	GSTE3
1637286_A_AT	RAW	RAW
1637355_AT	LD0326P	MTH781
1637605_S_AT	CG1146-PA, ISOFORM A	CG1146
1637788_AT	CG31176-PA	CG31176
1638005_S_AT	SMALIG	SMG
1638277_AT	CG11550-PA	CG11550
1638657_AT	NA	NA
1638693_S_AT	CG1882-PB, ISOFORM B	CG1882
1638844_S_AT	GH04243P	CG3714
1639178_S_AT	MAELSTROM	MAEL
1639223_A_AT	STRIFE	SR
1639571_S_AT	HEAT-SHOCK-PROTEIN-70AB	HSP70AB
1639785_S_AT	DAPLED	DPLD
1639808_AT	CG6892-PA, ISOFORM A	ETS96B
1640039_AT	CG13681-PA	CG13681
1640168_AT	CG9454-PA	CG9454
1640341_S_AT	CG5729-PB, ISOFORM B	DGP-1
1640363_A_AT	GH09623P	CG6330
1640423_AT	RE11961P	CG11811
1640822_S_AT	QUAKING RELATED 54B	QKR54B
1640884_AT	CG15784-PA	CG15784
1640979_AT	CG1681-PA	CG1681

Gens reprimits al <i>microarray</i> i positius al <i>ChIP on chip</i> de GAGA		
Sondes d'Affymetrix positives al <i>microarray</i>	"DAVID ID" del gen	Símbol del gen
1622908_A_AT; 1628935_AT	RE29090P	CG10962
1622932_S_AT	SINGED	SN
1623026_A_AT	CG10249-PB, ISOFORM B	CG10249
1623069_S_AT	CG17544-PB, ISOFORM B	CG17544
1623080_AT	CG18410-PA	CG18410
1623084_AT	SHAGGY	SGG
1623189_AT	POLO KINASE 1	SLIK
1623200_AT	FINGE	FMG
1623252_A_AT	GM05016P	KDN
1623411_AT; 1633089_A_AT	LONGITUDINALS LACKING	LOLA
1623418_AT	TYROSINE KINASE RECEPTOR CAD96CA PRECURSOR	CAD96CA
1623441_AT	POLYHOMEOTIC PROXIMAL	PH-P
1623497_AT	E(SPL) REGION TRANSCRIPT M3	HLHM3
1623514_A_AT; 1626174_A_AT	J DOMAIN-CONTAINING PROTEIN	JDP
1623605_A_AT	CG4427-PA, ISOFORM A	CBT
1623699_A_AT	SMOOTH	SM
1623769_AT	RH5757P	CG7322
1623957_S_AT	GLUTATHIONE S TRANSFERASE S1	GST51
1623963_AT	CG32115-PA	CG32115
1623996_AT	CG5001-PA	CG5001
1624059_AT	SD04822P	CG32223
1624122_A_AT	CG33529-PB, ISOFORM B	RAPGAP1
1624237_AT	CG15099-PA	CG15099
1624273_AT	CG8034-PA	CG8034
1624377_S_AT	NA	NA
1624410_AT	GH29222P	CG15658
1624522_AT	CG5302-PA	CG5302
1624574_AT	CG24656-PA	CG14656
1624660_AT	PROTEIN TYROSINE PHOSPHATASE 10D	PTP10D
1624744_A_AT	CG33991-PD, ISOFORM D	NUF
1624841_S_AT	CG13585-PB, ISOFORM B	CG13585
1624982_S_AT	CG5080-PB, ISOFORM B	CG5080
1625023_A_AT	NICOTINIC ACETYLCHOLINE RECEPTOR BETA 21C	NACRBETA-21C
1625048_AT	NEUROFIBROMIN 1	NF1
1625116_AT; 1640223_A_AT	CG33950-PB, ISOFORM B	TROL
1625183_AT	SMELL IMPAIRED 35A	SMI35A
1625227_AT	CG30456-PA	CG30456
1625369_AT	CG10877-PA	CG10877
1625373_AT	AT05246P	CG6966
1625471_S_AT	GH20759P	CG4928
1625475_S_AT	DOWNSTREAM OF RECEPTOR KINASE	DRK
1625644_AT	GH15037P	CG2641
1625761_A_AT; 1632204_AT	MATRIX METALLOPROTEINASE 1	MMP1
1625844_S_AT	APONTIC	APT
1625885_AT	-TUBULIN AT 97EF	BETATUB97EF
1626016_S_AT	SANPOOD PROTEIN	TMOD
1626079_A_AT	LD18613P	CG2201
1626314_AT	VIBRATOR	VIB
1626641_S_AT	GLOBIN 1	GLOB1
1626666_S_AT	CG7122-PA, ISOFORM A	RHOGAP16F
1626727_AT	MONOCARBOXYLATE TRANSPORTER 1	MCT1
1626839_S_AT	BLOATED TUBULES	BLOT
1626846_S_AT	CG5973-PA, ISOFORM A	CG5973
1626869_A_AT; 1630226_AT; 1642004_AT	SD16815P	CG31619
1626917_AT	CG31337-PA	CG31337
1626965_AT	CG31743-PA	CG31743
1627041_S_AT	INERBIATED	INE
1627073_A_AT	GH29949P	CG10126
1627191_A_AT	ENABLED	ENA
1627228_AT	YIPPEE-LIKE PROTEIN CG15309	CG15309
1627400_A_AT	AT22312P	CG5613
1627445_S_AT	ENGRAILED	EN
1627599_S_AT	HOPE11	HOE1
1627511_AT	CHARYBDE	CHR
1627589_S_AT	CG9520-PA, ISOFORM A	CG9520
1627617_AT	BESTROPHIN 1	BEST1
1627649_AT; 1635984_AT	SD09317P	NA
1627746_AT	CG6124-PA	CG6124
1628052_AT	CYPA17	CYPA17
1628081_S_AT	CG7530-PB, ISOFORM B	CG7530
1628224_A_AT	EARLY GENE AT 23	E23
1628232_AT	CG30046-PB	CG30046
1628398_AT	CG5853-PA	CG5853
1628493_AT	GH04511P	KAL-1
1628845_AT	CG6018-PA	CG6018
1628778_AT	ION TRANSPORT PEPTIDE	ITP
1628947_S_AT	MINIDISCS	MND
1629015_AT	RH61354P	DIAM
1629043_AT	GUSTATORY RECEPTOR 93D	GR93D
1629062_AT	CG13252-PA	CG13252
1629430_S_AT; 1629889_S_AT	REGUCALCIN HOMOLOGUE	REGUCALCIN
1629469_S_AT	CG10960-PC, ISOFORM C	CG10960
1629803_A_AT	SCAB	SCB
1629963_AT	SUGAR TRANSPORTER 1	SUT1
1629981_AT	LAMIN C	LAMC
1630007_S_AT	PHOSPHOLIPASE C AT 21C	PLC21C
1630010_A_AT	POINTED	PNT
1630026_S_AT	DECAPENTALEGIC	DPP
1630113_AT	CG10660-PA	CG10660
1630248_AT	CG5346-PA	CG5346
1630265_AT	CG31370-PA	CG31370
1630531_AT	CG3421-PA	RHOGAP93B
1630696_AT	CG2972-PA	CG12972
1630968_AT	CG13907-PA	CG13907
1631129_AT	IP01335P	CG31612
1631183_AT	RHO-LIKE	RHOL
1631285_AT	TETRASPANNIN 5D	TP5D
1631327_AT	M-SPONDIN	MSPD
1631426_AT	CG32412-PA	CG32412
1631610_AT	GM03781P	CG31886
1631783_AT	RH13166P	CG4267
1632011_AT; 1636736_S_AT	MISEXPRESSION SUPPRESSOR OF KSR 4	MESK4
1632213_S_AT	CG31292-PA	CG31292
1632213_S_AT	LD39912P	CG3303
1632231_A_AT	MYOSIN HEAVY CHAIN-LIKE	MHCL
1632313_AT	CG1998-PA	CG1998
1632457_S_AT	MASTERMIND	MAM
1632675_S_AT	CG5455-PA, ISOFORM A	CG5455
1632974_S_AT	GH22734P	CG30015
1632975_AT	CG4398-PA	CG4398
1633439_AT	G PROTEIN 73B	GAU73B
1633467_AT	LP01642P	CG2444
1633611_S_AT	CG32043-PA, ISOFORM A	CG32043
1633641_A_AT	RE34668P	CG15611
1633682_A_AT	CG32697-PE, ISOFORM E	LI1 G0232
1633753_AT	OFF-TRACK	OTK
1633975_S_AT	GH13040P	CG12214
1634099_AT; 1639574_A_AT	RETINAL DEGENERATION B	ROGB
1634146_AT	CG15926-PA	CG15926
1634296_S_AT	SD03967P	CG17646
1634431_AT	CG15071-PA	CG15071
1634546_AT	TIGGRIN	TIG
1634766_AT	SERRATE	SER
1634893_AT; 1636311_AT	GLYCEROL 3 PHOSPHATE DEHYDROGENASE	GPDH
1634922_S_AT	HYPOTHETICAL PROTEIN CG32795 IN CHROMOSOME 1	CG32795
1634952_S_AT	CG15094-PA, ISOFORM A	CG15094
1634983_S_AT	CG32521-PA, ISOFORM A	CG32521
1635000_AT	GLYCOCENIN	GLYCOCENIN
1635336_AT	DRK	CGM02
1635347_A_AT	CG8557-PA, ISOFORM A	CG8557
1635469_AT	CG18228-PA	CG18228
1635583_A_AT	FRIZZLED 2	FZZ
1635870_AT	LD16544P	CG1628
1636145_AT	CG7219-PA	CG7219
1636149_AT	CG31705-PB, ISOFORM B	CG31705
1636347_AT	HEPARAN SULFATE 6-O-SULFOTRANSFERASE	HS6ST
1636493_AT	DELIHAH	DEI
1636730_AT	LD26007P	MESO18E
1636815_A_AT	NUMB	NUMB
1636835_AT	LP07767P	CG16700
1636864_AT	CORTO	CORTO
1636961_A_AT	CG9027-PB, ISOFORM B	CG9027
1637275_A_AT	GH09427P	CG13335
1637312_A_AT	BROAD	BR
1637478_S_AT	HEPHAESTUS	HEPH
1637481_AT	TOLLO	TOLLO
1637491_S_AT	CG3835-PB, ISOFORM B	CG3835
1637625_S_AT	CG32813-PC, ISOFORM C	CG32813
1637658_AT	METHUSELAH-LIKE 3	MTHL3
1638050_S_AT	SOX BOX PROTEIN 14	SOX14
1638345_AT	FOLDED GASTRULATION	F06
1638368_AT	CG31674-PA	CG31674
1638462_AT	CG18088-PA	CG18088
1638498_S_AT	CG1600-PC, ISOFORM C	CG1600
1639221_S_AT	TOLL	TL
1639439_AT	CG2471-PA	CG2471
1639469_A_AT	PUNCH	PU
1639542_AT	SEMAPHORIN 5C	SEMA-5C
1639928_A_AT	TUNGUS	CG30084
1639956_AT	CG17760-PA	CG17760
1640002_AT	LD40103P	CG4586
1640090_A_AT	SILUGOSH A	SLGA
1640200_AT	DYSTROPHIN	DYS
1640214_AT	CG31216-PA	CG31216
1640235_AT	WRINKLED	W
1640300_AT	GH08706P	CG33144
1640590_AT	CG32195-PA	CG32195
1640770_A_AT	ECDYSONE-INDUCED PROTEIN 78C	EIP78C
1640798_A_AT	GAP JUNCTION PROTEIN PRP33	INX2
1641031_S_AT	LD33329P	JGR1
1641063_S_AT	ORGANIC ANION TRANSPORTING POLYPEPTIDE 30B	OATP30B
1641068_A_AT	SPATZLE	SPZ
1641152_AT	CG1718-PA	CG1718
1641428_AT	CYTOCHROME P450-9C1	CYP9C1
1641496_A_AT	CG5896-PA, ISOFORM A	GRASS
1641530_S_AT	LD30894P	CG2264

H.2.3. Comparació dels gens que solapen entre els dos experiments de *microarray* i les dades de *ChIP on chip*

A continuació es mostra una relació dels gens que hem trobat diferencialment expressats tant al *microarray* de sobreexpressió com al d'RNAi i que, a més, són positius a l'experiment de *ChIP on chip* de GAGA.

Gens que són positius als dos tipus de <i>microarray</i> i a la vegada al <i>ChIP on chip</i> de GAGA			
"DAVID ID" del gen	Símbol del gen	<i>Microarray</i> de sobreexpressió	<i>Microarray</i> d'RNAi
BLOATED TUBULES	BLOT	+	-
BROAD	BR	-	-
GH10595P	L(2)K05713	+	+
GH22994P	CG10126	+	-
GLUTATHIONE S TRANSFERASE S1	GSTS1	+	-
LD30894P	CG2264	+	-
PHOSPHOLIPASE C AT 21C	PLC21C	+	-
WRINKLED	W	+	-

+ Activat

- Reprimit

Nota: Per a Broad, la sonda d'Affymetrix que detecta la repressió al *microarray* de sobreexpressió és diferent del que la detecta al *microarray* d'RNAi