

**UNIVERSITAT AUTÒNOMA DE BARCELONA**  
Departament de Ciència Animal i dels Aliments  
Facultat de Veterinària

**CENTRE DE RECERCA EN AGRIGENÒMICA**  
Unitat de Genètica Animal

# **Caracterització de microRNAs d'interès en l'espècie porcina**

Oriol Timoneda i Heredia

Tesi Doctoral  
juliol de 2013

Director:  
Armand Sánchez Bonastre



La ignorància et pot fer feliç;  
el coneixement, lliure.

I must be strong and carry on,  
'cause I know I don't belong here in heaven.

*(Tears in heaven, Eric Clapton)*



El Dr. **Armand Sánchez Bonastre**, catedràtic del Departament de Ciència Animal i dels Aliments de la Universitat Autònoma de Barcelona

CERTIFICA:

Que l'**Oriol Timoneda i Heredia** ha realitzat sota la seva direcció el treball de recerca "Caracterització de microRNAs d'interès en l'espècie porcina" per a obtenir el grau de doctor per la Universitat Autònoma de Barcelona.

Que aquest treball s'ha dut a terme al Departament de Ciència Animal i dels Aliments de la Facultat de Veterinària de la Universitat Autònoma de Barcelona i a la Unitat de Genètica Animal del Centre de Recerca en Agrigenòmica.

Bellaterra, 27 de maig de 2013

Dr. Armand Sánchez Bonastre



## RESUM

El descobriment dels microRNAs (miRNAs) com a nous reguladors de l'expressió gènica ha obert un nou camp en l'estudi de quin és el comportament d'aquests RNAs de mida petita i descriure en quins processos actuen i de quina manera regulen l'expressió gènica. L'aparició de les tècniques de seqüenciació massiva ha permès la descripció i estudi dels perfils d'expressió de miRNAs en diferents situacions, a fi d'observar la seva implicació en els processos biològics, tant fisiològics com patològics. Aquesta tesi exemplifica, mitjançant dues aproximacions diferents, l'ús d'aquestes tècniques per a la descripció, descobriment i estudi de perfils de miRNAs en l'espècie porcina.

La primera part del treball es va dissenyar amb l'objectiu d'ampliar el nombre de miRNAs descrits en el porc. Les aproximacions utilitzades per a la determinació de nous miRNAs varen ser, primer de tot, la descripció del perfil d'expressió de miRNAs en el ronyó del porc, incloent els miRNAs ortòlegs i, segon, l'ús d'un protocol de descobriment i validació de nous miRNAs porcins. Una altra motivació del treball va ser el fet d'estudiar possibles canvis en l'expressió dels miRNAs en races de porc de diferents orígens, des de races europees fins a asiàtiques, incloent races europees amb influència asiàtica, on hem descrit miRNAs diferencialment expressats i s'ha estudiat la seva possible funcionalitat.

En la segona part del treball es va dur a terme una infecció experimental amb el virus de la malaltia d'Aujeszky (ADV), l'herpesvirus porcí tipus 1 (SHV-1), amb una soca virulenta (NIA-3) i una soca vacunal (Begonia). Es varen realitzar dues aproximacions diferents: un aproximació *in vitro* utilitzant les línies cel·lulars PK-15 derivades del ronyó de porc, i una aproximació d'infecció experimental *in vivo* utilitzant el bulb olfactori i el gangli trigemin com a teixits diana per l'estudi. Amb l'objectiu d'estudiar la implicació dels miRNAs, tant virals com de l'hoste, en les interaccions hoste – patògen durant una infecció vírica, s'han descrit els perfils de miRNAs i s'han avaluat les diferències en la seva expressió, no només entre el grup infectat i el grup control, sinó també entre les soques utilitzades i entre les dues aproximacions. També s'han descrit nous miRNAs virals que s'expressen durant la infecció. Finalment, s'ha elaborat, amb estudis funcionals *in silico*, una xarxa d'interaccions entre els miRNAs virals, els miRNAs de l'hoste diferencialment expressats i els gens que codifica l'agent infeccios SHV-1.

Donada la importància de la tècnica de l'RT-qPCR per a la validació de l'expressió de miRNAs, també s'ha realitzat un estudi addicional per avaluar l'estabilitat en l'expressió d'alguns miRNAs per a poder ser utilitzats com a gens de referència en estudis de quantificació relativa de dades d'RT-qPCR, ja que fins l'actualitat s'han fet servir poc per a aquest propòsit.

**Paraules clau:** microRNA, porc, seqüenciació massiva, quantificació relativa per RT-qPCR, perfil d'expressió de miRNAs, malaltia d'Aujeszky, interaccions hoste – patògen.

## SUMMARY

The discovery of microRNAs (miRNAs) as novel gene expression regulators has opened a new field in the study of the roles of these small RNAs, as well as on describing in what processes they act and how they regulate gene expression. The emergence of next generation sequencing methods has allowed the description and study of miRNA expression profiles in different situations, in order to observe its involvement in biological processes, both physiological and pathological. This thesis illustrates, through two different approaches, the use of these techniques for the description, discovery and study of miRNA profiles in the porcine species.

The first part of the study was designed with the aim of increasing the number of described miRNAs in pigs. The approaches used for the determination of novel miRNAs were, first of all, the expression profiling of miRNAs in the swine kidney, including the orthologous ones and, second, using a pipeline for the discovery and validation of new porcine miRNAs. Another motivation of this work was to study the possible changes in the kidney miRNAs expression patterns among pig breeds from different origins, from European to Asian breeds, including European breeds with Asian influences. In this sense, differentially expressed miRNAs have been described and their functional roles have been studied.

In the second part of the study, an experimental infection with Aujeszky's disease virus (ADV), also known as suid herpesvirus type 1 (SHV-1), was carried out, using a virulent strain (NIA-3) and a vaccine strain (Begonia). Two different approaches were conducted: an *in vitro* approach using PK-15 cell lines, derived from pig kidney, and an *in vivo* approach using the olfactory bulb and trigeminal ganglia as target tissues. With the aim of studying the role of both host and viral miRNAs in host – pathogen interactions during SHV-1 infection, miRNAs expression profiles have been described and their expression differences evaluated, not only between infected and mock-infected groups, but also between strains and between the two performed approaches. New viral miRNAs have been described and their expression during the infection has been confirmed. Finally, a network of interactions between viral miRNAs, host differentially expressed miRNAs and SHV-1 encoded genes was developed using *in silico* functional studies.

Given the importance of RT-qPCR method for validating the expression of miRNAs, we also performed an additional study to assess the expression stability of some miRNAs to be used as reference genes in relative quantification studies of RT-qPCR data, which they have not been widely used for this purpose.

**Key words:** microRNA, pig, high throughput sequencing, relative quantification, RT-qPCR, miRNAs expression profile, Aujeszky's disease virus, host – pathogen interactions.



# CONTINGUTS

---

<b>1. INTRODUCCIÓ</b>	<b>11</b>
<b>1.1. L'espècie porcina</b>	<b>11</b>
1.1.1. La domesticació de l'espècie porcina	11
1.1.2. Filogènia del porc domèstic	16
1.1.3. Programes de millora genètica en la producció porcina	18
1.1.4. Abast nacional i internacional de la producció porcina	20
1.1.5. El porc com a model biomèdic	20
1.1.6. El virus de la malaltia d'Aujeszky	22
<b>1.2. Expressió gènica</b>	<b>25</b>
1.2.1. Mecanismes de regulació de l'expressió gènica	26
1.2.1.1. Els microRNAs	28
1.2.1.1.1. Biogènesi i mecanismes d'actuació	29
1.2.1.1.2. Conservació i especificitat	30
1.2.1.1.3. Anotació i validació de la seva expressió	31
1.2.1.1.4. Predicció de gens diana i anàlisis funcionals de miRNAs	32
1.2.1.1.5. Els miRNAs virals	33
1.2.2. Mesura de l'expressió gènica	34
1.2.2.1. Tècniques de seqüenciació massiva	36
1.2.2.1.1. Primera generació de seqüenciació: el mètode de Sanger	36
1.2.2.1.2. Segona generació de seqüenciació: seqüenciació en paral·lel massiva amb pre-amplificació	36
1.2.2.1.3. Tercera generació de seqüenciació: seqüenciació en paral·lel massiva de molècula única	39
1.2.2.1.4. Aplicació de les tècniques de seqüenciació massiva a l'estudi de miRNAs	40
1.2.2.2. La tècnica de la reacció en cadena de la polimerasa a temps real mitjançant transcripció reversa (RT-qPCR)	42
1.2.2.2.1. Metodologia de l'RT-qPCR	42

1.2.2.2.2.	Estratègies de quantificació de l'expressió gènica mitjançant l'RT-qPCR.....	45
1.2.2.2.3.	Adaptació de l'RT-qPCR per a l'anàlisi de miRNAs.....	46
1.2.2.2.4.	L'ús de microRNAs com a gens de referència.....	48
1.2.2.2.5.	La PCR digital, la tercera generació de la PCR.....	49
<b>2.</b>	<b>OBJECTIUS</b> .....	<b>51</b>
<b>3.</b>	<b>ARTICLES I ESTUDIS</b> .....	<b>53</b>
3.1.	Estudi del perfil d'expressió dels microRNAs al ronyó de diferents races porcines.....	53
3.2.	Estudi de les funcions dels microRNAs virals i de l'hoste durant el procés d'infecció amb el virus de la malaltia d'Aujeszky.....	69
3.3.	Estudi dels microRNAs per a ser utilitzats com a gens de referència en estudis de quantificació relativa.....	125
<b>4.</b>	<b>DISCUSSIÓ GENERAL</b> .....	<b>141</b>
<b>5.</b>	<b>CONCLUSIONS</b> .....	<b>155</b>
<b>6.</b>	<b>BIBLIOGRAFIA</b> .....	<b>159</b>

# 1. INTRODUCCIÓ

---

## 1.1. L'espècie porcina

### 1.1.1. La domesticació de l'espècie porcina

Per determinar l'origen de l'espècie porcina, la seva domesticació, la interacció amb el seu ancestre, el porc senglar (senglar en endavant) i la procedència d'aquest han calgut anys d'estudis en diferents camps com la genètica i l'arqueologia. Per definir els primers punts de domesticació cal remetre's, no tan sols a informació de base genètica, sinó també a evidències arqueològiques i històriques. L'ús d'aproximacions interdisciplinàries introduint coneixements sociològics, econòmics, culturals, ambientals i ecològics ha permès desxifrar o, si més no, començar a esclarir quines corrents migratòries hi ha hagut al llarg de la història tant dels senglars com dels porcs domèstics (porc en endavant) i perquè el procés de domesticació ha estat llarg, complicat i divergent d'altres espècies domèstiques (White, 2011, Larson *et al.*, 2011).

La domesticació és una acció en la que l'ésser humà acostuma o adapta un animal salvatge a viure amb ell en un ambient captiu, desencadenant una sèrie de canvis morfològics, fisiològics i comportamentals a través de diverses generacions. És un procés llarg i de canvi continuat, gradual. Així doncs, cal que es plantegi aquest procés com un període en el que els seus dos extrems (salvatge – domèstic) són termes oposats antònims (graduals) i no complementaris (excloents) (Dobney i Larson, 2006), fet que implica que el porc ha conviscut amb el seu ancestre, el senglar. De fet, és una de les poques espècies domesticades que encara conviu temporalment amb el seu ancestre, a diferència d'altres espècies domesticades a Euràsia com la vaca o el cavall, i, a més, el senglar manté una àmplia estructura filogeogràfica, fet que no succeeix amb el llop, per exemple (Larson *et al.*, 2005). Aquest fet és important ja que en alguns moments durant la domesticació el porc s'ha arribat a creuar amb el senglar especialment a Europa, i això és determinant per aclarir els patrons d'haplotips que es troben en alguns treballs (Groenen *et al.*, 2012). Actualment, la diferenciació genètica entre els porcs i els senglars és sensiblement petita, degut a la interacció genètica que han

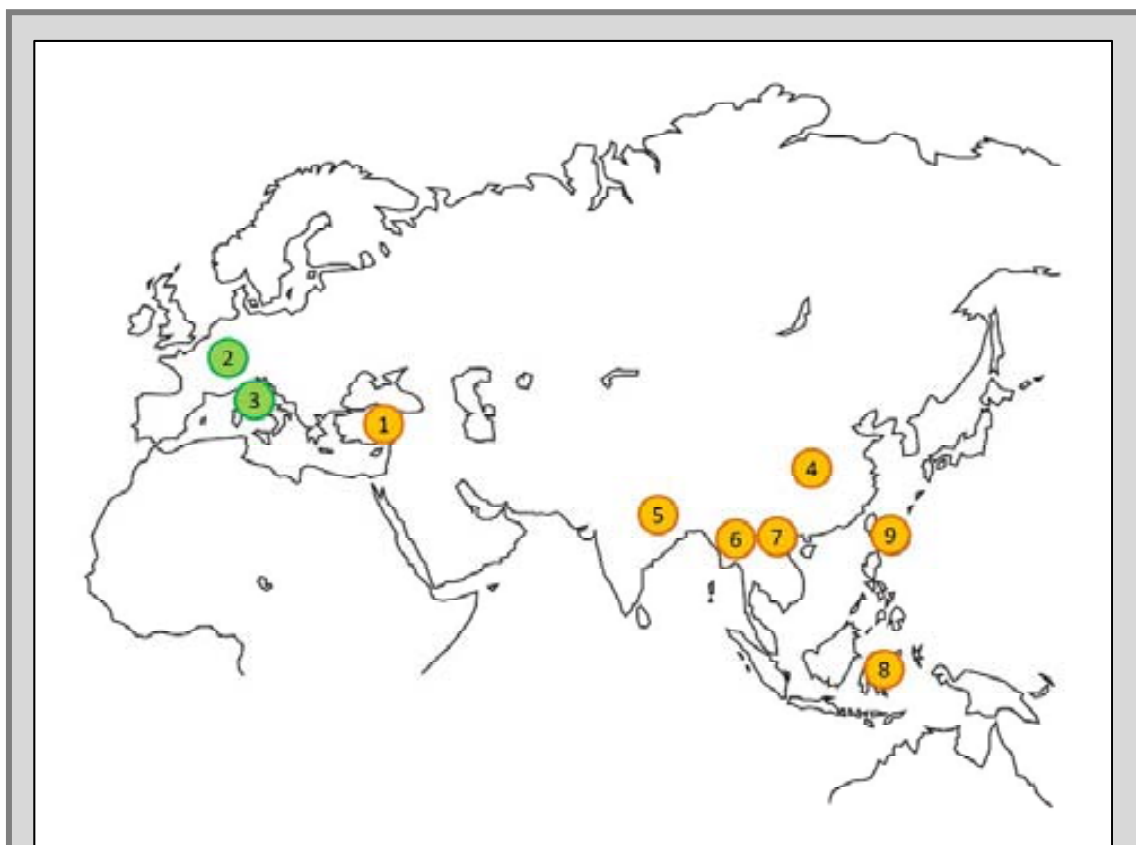
mantingut les dues poblacions (Ramirez *et al.*, 2009) i és de gran interès l'estudi de llinatges de senglars amb un cariotip  $2n = 36$  i altres llinatges de senglars i totes les races de porcs amb cariotip  $2n = 38$  (Fang *et al.*, 2006, Kijas i Andersson, 2001, Giuffra *et al.*, 2000).

Hi ha estudis que situen l'origen del senglar a l'oest de la zona marítima de l'Àsia sud-oriental (Larson *et al.*, 2005, Larson *et al.*, 2010). Groenen *et al.*, (2012) data la seva aparició a l'inici del Pliocè, fa uns 5,3-3,5 milions d'anys. Una primera migració cap al continent asiàtic i el subcontinent indi, seguida d'una progressiva expansió per Euràsia van fer que el senglar arribés fins a l'Europa occidental (Larson *et al.*, 2005). Paral·lelament, el senglar també es va escampar per tot el sud d'Àsia (Larson *et al.*, 2007b). A mitjan Pleistocè (fa 1,8-0,8 milions d'anys), durant l'última gran glaciació, les poblacions de senglars europeus i asiàtics van quedar aïllades. Diversos estudis daten la divergència entre els llinatges europeu i asiàtic de senglars entre 500.000 i 1 milió d'anys, ajustant-se, més o menys, a aquesta època (Groenen *et al.*, 2012, Kijas and Andersson, 2001, Giuffra *et al.*, 2000, Alves *et al.*, 2003).

A principis de l'Holocè (fa 11.000 anys) l'ésser humà irromp en l'Era Neolítica, primerament al Pròxim Orient, i un nou model de societat s'imposa. El desenvolupament de l'agricultura i la ramaderia va comportar el canvi de vida nòmada a sedentària de l'espècie humana, el canvi de ser caçadors i recol·lectors a productors i agricultors. Així, es va iniciar la domesticació d'algunes espècies a les quals se'ls va començar a trobar utilitats, com l'ovella, la cabra o el vedell. A diferència d'aquestes espècies, el porc no és un remugant i, per tant, al tenir una dieta omnívora, la traçabilitat de transformar la cel·lulosa (no comestible) en carn (comestible) deixava de veure's clara. A més, es considerava que el porc competia directament amb l'espècie humana per l'aliment. Aquests fets van propiciar que el procés de domesticació fos llarg i costós, condicionant-lo a cercar el nínxol ecològic adient en la relació de simbiosi amb l'espècie humana. Així doncs, la motivació inicial de la domesticació del porc no va ésser amb objectiu de conversió alimentària sinó més aviat per un suposat emmagatzematge d'aliment (White, 2011). Més endavant, amb l'evolució de l'espècie humana i ja havent definit el rol de l'espècie porcina, es justifica la seva domesticació

per a servir com a font de proteïnes per a l'espècie humana i també pel gust singular de la carn de porc (Buchanan i Stalder, 2011).

Estudis arqueològics daten la domesticació del porc fa uns 9.000 – 10.000 anys (Larson *et al.*, 2011). A través d'estudis de DNA mitocondrial i nuclear entre races domèstiques i els seus ancestres s'ha conclòs que la domesticació del porc es va fer de manera independent a diferents zones d'Euràsia i s'han definit diferents punts de domesticació mitjançant arbres filogenètics i també en funció de la concordança del patró d'haplotips del porc domèstic amb el senglar de cada zona (Figura 1.1., Larson *et al.*, 2005, Ramirez *et al.*, 2009, Kijas and Andersson, 2001, Giuffra *et al.*, 2000, Alves *et al.*, 2003, Luetkemeier *et al.*, 2010). Es descriuen dos punts de domesticació a Europa, un estès per tot centre Europa i l'altre situat al sud d'Itàlia (Larson *et al.*, 2005, Kijas i



**Figura 1.1.** Mapa amb els diferents punts de domesticació descrits (figura adaptada de Larson *et al.*, 2011). Els punts verds (2-3) provenen de llinatges de senglar europeus i els punts taronges (1, 4-9) de llinatges asiàtics. (1) Pròxim Orient. (2) Europa. (3) Itàlia. (4) Xina continental. (5) subcontinent indi. (6, 7) Sud-est asiàtic. (8) Illa de Sulawesi. (9) Illa de Taiwan.

Andersson, 2001, Giuffra *et al.*, 2000, Alves *et al.*, 2003) i varis punts de domesticació a l'Àsia, des del Pròxim Orient, l'Índia, la Xina, el sud-est peninsular asiàtic i les illes de Taiwan i Sulawesi (Larson *et al.*, 2011, Larson *et al.*, 2005, Larson *et al.*, 2010, Wu *et al.*, 2007, Watanobe *et al.*, 1999). La divergència entre els dos llinatges europeus es pot explicar per la barrera geogràfica que els Alps Europeus varen suposar durant les últimes grans glaciacions del període Quaternari (Taberlet *et al.*, 1998).

La domesticació, però, va ser costosa. El primer ús al sud-est asiàtic dels porcs va ésser com a escombriaires per a residus humans i de rebuig, primer com a animals recol·lectors i finalment com a veritables animals de granja (Yuan i Flad, 2002), una transició que es va produir fa 8.000 anys, conduint el porc domèstic asiàtic cap a un animal més dòcil i pàl·lid, de cames curtes, panxut i amb forces races regionals característiques de casa zona (Epstein, 1971). Més cap a l'oest, a l'anomenat Creixent Fèrtil, aquesta transició va ser més lenta degut a l'elevada densitat dels assentaments neolítics en aquesta zona, que va propiciar el rebuig d'aquests animals per ser un destorb pels cultius i, a més, competència directa amb recursos escassos com l'aigua (Rosenberg i Redding, 1998). Paral·lelament, diverses poblacions europees, de manera independent, van desenvolupar la tècnica de la pastura de porcs domèstics (de l'anglès *pannage*) alimentant-se bàsicament de fages i glans, de poc valor per la subsistència humana. Aquesta pràctica va suposar l'alliberament dels porcs de cria controlats i sotmetre'ls a les condicions ambientals i de competència que tenien els seus antecessors. Com estaven la major part de l'any vagant pels boscos i camps, es feia difícil portar un control de la cria i afavoria el creuament amb els senglars. Així, el porc europeu es caracteritzava per ser ferotge i àgil, magre i amb les potes llargues, crestat i amb ullals residuals (White, 2011). S'ha vist que una primera forma domèstica de porcs del Pròxim Orient es va introduir a Europa durant l'Era Neolítica (fa 5.000 – 6.000 anys). Aquest fet segurament va propiciar la domesticació del senglar europeu desplaçant, finalment, la forma asiàtica. Estudis d'haplotips amb DNA mitocondrial demostren que no queda rastre d'aquesta primera introgressió del porc asiàtic a Europa en les races de porc europees actuals (Larson *et al.*, 2007a). També ho demostren estudis amb DNA autosòmic, tot i que s'ha arribat a trobar certa petjada de senglar del pròxim orient en mostres de senglars russos i belgues (Manunza *et al.*,

2013). Fins i tot, hi ha indicis que l'expansió del porc europeu arribà fins al Pròxim Orient, on s'han trobat haplotips Europeus en mostres de porcs de fa 1.000-3.000 anys (Larson *et al.*, 2007a, Ottoni *et al.*, 2013).

Les dues civilitzacions anaven evolucionant de manera contraposada. A l'Àsia, les poblacions creixien en densitat i l'agricultura evolucionava de manera intensiva, afavorint l'engreix dels porcs i posicionant-lo en l'animal de granja de referència (Simoons, 1991). A Europa, per contra, la densitat poblacional es va mantenir baixa, es practicava una agricultura més extensiva i amb predominança del bestiar remugant (Malcolmson i Mastoris, 1998). El porc europeu, doncs, seguia amb una pastura extensiva que resultava ser menys eficient i més complicada pels pagesos. El canvi definitiu a una producció porcina intensiva i una millora de les races es va produir a cavall dels segles XVII i XVIII a Anglaterra. Va ésser en aquesta època quan es varen produir una sèrie d'acostaments ecològics per tota Euràsia, amb una població en auge i limitacions en terres cultivables i boscos de pastura. Aquesta convergència també va succeir en el camp tecnològic i en la qualitat de vida entre les civilitzacions europees i asiàtiques, creant contactes entre elles i essent les precursors del comerç mundial i la globalització (Gunn, 2003). En el decurs d'un creixement poblacional, s'afavorí l'ús dels porcs com a escombriaires urbans, i amb la introducció de tubercles i llegums com les patates, els pèsols i les faves, millorà el seu rendiment com a animal de granja i n'accentuà la seva comercialització. Va ésser just en aquesta transició de millora de l'espècie que es va introduir el porc asiàtic, jugant un rol clau en la transformació del porc europeu. Charles Darwin (1868) ja va diferenciar dos llinatges de porcs el s.XIX: un d'europeu (*Sus scrofa*) i un d'asiàtic (*Sus indicus*). Estudis amb DNA mitocondrial proposen que l'arribada dels porcs asiàtics es va produir durant els segles XVIII i XIX (Giuffra *et al.*, 2000, Jones, 1998). També s'ha demostrat, per validar aquesta introgressió de porcs asiàtics a Europa, que les noves races europees, com la Landrace, Large White, Piétrain o Berkshire, comparteixen haplotips asiàtics i als arbres filogenètics se situen més properes al porc asiàtic que a l'europeu (Kijas i Andersson, 2001, Watanobe *et al.*, 1999, Clop *et al.*, 2004, Kim *et al.*, 2002). Per altra banda, també s'ha vist que les races europees tradicionals ubicades a la zona mediterrània no

presenten haplotips asiàtics i, per tant, el creuament amb races asiàtiques sembla improbable (Kijas i Andersson, 2001, Alves *et al.*, 2003, Clop *et al.*, 2004).

Literàriament, l'entusiasme amb que es descriu l'arribada d'aquestes races asiàtiques indica l'admiració i la seva utilitat en la intensificació de l'agricultura al nord d'Europa (Lawrence, 1727). Nombroses enciclopèdies d'animals i de procediments en agricultura descriuen els avantatges de les noves races creuades amb els porcs xinesos, principalment en la rapidesa de l'engreix, la dieta menys exigent i l'augment de la ventrada (Beilby, 1800). Amb cinquanta anys van aconseguir una posició dominant en l'agricultura anglesa i van contribuir a la revolució agrícola de finals del segle XVIII a Anglaterra (Overton, 1996). Tanmateix, l'auge del comerç i l'ampliació de les distàncies comercials cap a principis del segle XIX afavorí l'aparició i el comerç de nombroses races locals típiques de cada regió (Malcolmson i Mastoris, 1998). Aquest comerç a gran escala també va posar de manifest les diferències que s'havien creat entre les zones rurals i les urbanes i també entre Europa i Àsia. En aquest sentit, l'economia rural, predominant a Àsia, no estava tan desenvolupada i es centrava en una estratègia de subsistència en petites explotacions a diferència d'una estratègia d'empresa comercial especialitzada predominant a Europa (Trow-Smith, 2006). El canvi de forma que el porc domèstic havia experimentat en aquests últims cent anys i el sistema de producció també van precipitar el canvi del porc en la cultura anglesa. Així doncs, el porc va passar de ser aquella criatura forta, ferotge i obscura de l'Edat Mitjana, associada a rituals centenaris i simbolismes religiosos, a adquirir un rol de criatura inofensiva, divertida i satírica (Malcolmson i Mastoris, 1998). Les típiques guardioles angleses en forma de porc es van inventar a finals del segle XVIII, per exemple.

### **1.1.2. Filogènia del porc domèstic**

Filogènicament, les races de porc que coneixem avui dia provenen de l'espècie *Sus scrofa*. L'expansió eurasiàtica explicada anteriorment la va protagonitzar aquesta espècie, on estudis recents proposen que l'espècie es va originar al sud-est asiàtic (Larson *et al.*, 2005, Larson *et al.*, 2010), juntament amb d'altres espècies del gènere *Sus* (Ruvinsky *et al.*, 2011), totes elles actualment confinades a zones concretes del sud-est asiàtic: *Sus verrucosus* (Illa de Java), *Sus barbatus* (Malàisia continental i illes de Borneo, Sumatra, Bangka, Java, Palawan i Palau), *Sus celebensis* (Illa de Sulawesi i



Timor Oriental), *Sus philippensis* (est arxipèlag filipí) i *Sus cebifrons* (centre-oest arxipèlag filipí).

L'expansió o colonització de la resta del món per part de la humanitat va comportar, també, l'expansió de l'espècie porcina. Estudis arqueològics han intentat desxifrar l'evolució del gènere *Sus* arreu del món (Groves, 1981). Així, al nord del continent africà s'hi troben restes de porc que daten de l'Era Neolítica, fa entre 6.000 i 3.000 anys. Ramírez *et al.*, (2009) descriu que la població occidental africana de porcs descendeix de llinatges autòctons i llinatges europeus, mentre que la població oriental africana és una barreja de llinatges europeus i asiàtics. Larson *et al.*, (2007b) descriu l'arribada de porcs de l'espècie *Sus scrofa* del continent asiàtic a l'arxipèlag del sud-est asiàtic, tot i haver-hi altres espècies. Precisament, la introducció de porcs a Oceania també s'atribueix a individus de l'espècie *Sus scrofa* durant l'Era Neolítica. Finalment, la colonització del continent americà durant els segles XV a XVII va significar també l'expansió del porc europeu a Amèrica, demostrant la presència d'haplotips europeus a les poblacions de porcs criolls a sud-Amèrica (Ramírez *et al.*, 2009).

El gènere *Sus scrofa* està format per 16 subespècies (FAO, <http://dad.fao.org>), una de les quals, *Sus scrofa domestica*, inclou totes les races de porc. Definir el nombre de races existents a l'actualitat no és fàcil, ja que tampoc ho és definir què significa el terme raça. *A priori*, una raça podria ser un conjunt d'animals de la mateixa espècie amb característiques físiques semblants. Tot i així, existeixen races (no només de porc) amb una àmplia variació en algun o alguns trets físics, i també existeixen animals de races diferents que són molt semblants. Hi ha un consens general, però, en determinar que el concepte de raça va lligat a una ascendència comuna. Tanmateix, Buchanan i Stalder (2011) descriuen fins a 65 races de porc segons el seu origen, distribució i les principals característiques fenotípiques. Si comptéssim les línies de cada raça i les seves variacions, però, podríem arribar a comptar uns quants centenars de varietats (FAO, <http://dad.fao.org>).

La introducció de les races asiàtiques a Europa durant els segles XVIII i XIX va esdevenir un punt d'inflexió en la creació de les races modernes de porc. White (2011) proposa

dues introduccions separades, cada una amb una varietat de porc asiàtic diferent: una primera relacionada amb la creació de les races britàniques Large White i Berkshire i una de més tardana amb les races Landrace, Duroc i Welsh. Aquest intercanvi genètic es va realitzar amb els objectius d'augmentar el greix de l'animal i disminuir l'edat d'arribada a la maduresa degut a les bones característiques reproductives que tenien les races asiàtiques. L'impacte d'aquests creuaments es pot mesurar en el fet que les races britàniques es varen fer servir posteriorment per a crear noves races a Europa, Àsia i Amèrica (Buchanan i Stalder, 2011, Amills *et al.*, 2010).

En l'actualitat, podríem classificar les principals poblacions comercials en funció de les seves característiques productives i reproductives. Hi ha les races amb una elevada taxa de creixement i un bon nivell reproductiu, com la Landrace, la Large White i la Yorkshire; les races amb una elevada taxa de creixement i poc prolífiques, com són la Piétrain, la Hampshire i la Duroc; les races hiperprolífiques són principalment les de tronc asiàtic, com la Meishan i la Erhualian; i finalment hi ha el grup on s'inclouen les races locals, ben adaptades al seu medi ambiental però amb taxes de creixement i nivells reproductius discrets, com la Ibèrica.

### **1.1.3. Programes de millora genètica en la producció porcina**

Després de la domesticació i amb el pas dels anys, Europa i Àsia van esdevenir els dos grans focus de cria de porcs, desenvolupant varietats locals adaptades a les necessitats i requeriments de cada zona mitjançant la selecció. Durant segles, la manera de seleccionar els porcs va ésser el creuament entre els individus que presentaven les millors característiques en productivitat, docilitat i resistència. Més endavant aquesta tasca es va sistematitzar amb l'aparició de les primeres genealogies (*pedigrees*) per part dels criadors anglesos a principis del segle XIX (Lush, 1943), on utilitzaven aquesta informació per seleccionar els millors animals i creuar-los amb línies ancestrals. Amb la integració de les lleis de Mendel i l'assumpció de la genètica com a disciplina científica, el mètode de selecció van esdevenir més acurats i eficients. Durant molt de temps, la millora genètica del porc s'ha dut a terme amb els índexs de selecció i la metodologia BLUP (*Best Linear Unbiased Prediction*) per calcular el valor millorant dels animals. Tot i així, aquestes aproximacions només fan servir informació fenotípica i genealògica.

L'aplicació de la genètica molecular als programes de millora genètica als anys 80 va suposar un avenç en l'eficiència dels programes de millora (Andersson, 2001). L'estudi de gens o regions cromosòmiques que afecten o estan relacionades amb caràcters quantitius d'interès productiu, anomenats QTL (*Quantitative Trait Loci*), suposa una estratègia que pretén complementar la selecció convencional, especialment en caràcters on els registres fenotípics són cars o difícils d'obtenir (Dekkers, 2007). Així, en alguns casos es poden arribar a seleccionar directament regions cromosòmiques que afecten el caràcter en qüestió a través de la selecció assistida de marcadors (MAS) o gens (GAS) (Dekkers, 2004). L'aparició de nous marcadors moleculars com els SNP (*Single Nucleotide Polymorphism*) i les noves tècniques de genotipatge massiu permeten treballar, actualment, amb tot el genoma d'un individu gràcies a l'avenç tecnològic. Aquesta nova metodologia, la selecció genòmica, permet obtenir valors predictius genètics més acurats d'aquells caràcters d'interès productiu i reproductiu (Goddard i Hayes, 2007, Meuwissen *et al.*, 2001).

Els caràcters seleccionats en els programes de millora genètica han anat evolucionat amb el pas del temps. A principis del 1950, els esforços de selecció genètica es centraven en caràcters productius, com reduir el greix dorsal i millorar la taxa de creixement, aconseguint veritables progressos en la producció de carn magra i disminuint els dies necessaris per arribar al pes de comercialització dels animals (Dekkers *et al.*, 2011). Més endavant es varen començar a tenir en compte factors com la qualitat de la carn i la reducció de costos en la producció. Ja en les últimes dècades del segle XX, la selecció i millora genètica porcina va començar a centrar el seu interès en els caràcters reproductius, ja que una bona fecundació es podia relacionar directament amb una bona vida productiva de les truges (Distl, 2007). Aprofitant les bones aptituds maternals i les bones característiques en els caràcters reproductius de les races asiàtiques, es va començar a treballar en la millora de la mida de la ventrada en les races productores de carn (Bidanel, 2011). Aquesta última dècada ha començat a suscitar interès, per pressió i demanda de la societat, la millora d'altres caràcters que no són purament econòmics i no estan directament relacionats amb l'eficiència productiva, com el benestar i la salut dels animals, les conseqüències ecològiques

d'una producció porcina altament intensiva i la qualitat sensorial i la salubritat de la carn de porc (Kanis *et al.*, 2005).

#### **1.1.4. Abast nacional i internacional de la producció porcina**

El porc domèstic es troba estès per quasi tot el món degut a la seva extraordinària importància en la producció de carn vermella, llard i productes curats (Amills *et al.*, 2010). La pell s'aprofita per confeccionar articles de roba, sabates i complements; els ossos per elaborar armes, gomes, esmalts, vidres i productes de neteja; del greix també se'n fan olis lubricants i productes químics, com pintures de maquillatge i vernissos; i amb pèls s'elaboren pinzells i raspalls. Així doncs, la dita popular que diu que del porc s'aprofita tot és ben bé certa.

La carn de porc és la més consumida a nivell mundial i, en conseqüència, la més produïda. Representa el 40% de la producció de carn el 2011, amb 110M de tones i amb tendència a l'alça. La producció de carn de porc es concentra principalment a Àsia (56%), Europa (25%) i Amèrica (17%). Per països, Espanya és el 4rt país que més produeix carn de porc (3,5M de tones, 3%), tan sols superada per la Xina (51,5M de tones, 47%), els Estats Units (10,3M de tones, 9%) i Alemanya (5,6M de tones, 5%. FAOSTAT, 2011).

A l'Estat espanyol, amb dades del 2011, el pes de la producció porcina és molt important, amb 42M de caps sacrificats i 3,5M de tones de carn produïdes. El cens porcí espanyol arriba als 25,6M de caps, i Catalunya concentra el 27,5% del cens (7M d'animals) i amb 1,4M de tones de carn de porc representa el 40% de la producció de carn de porc a l'Estat. La producció de bestiar de porquí a Catalunya va aportar, el 2011, el 55% de la producció final ramadera amb més de 17,5M de caps sacrificats (MAGRAMA i DAAM, 2011).

#### **1.1.5. El porc com a model biomèdic**

A banda de la seva importància productiva en el sector agrari, el porc és un valuós model biomèdic per a l'estudi de la salut i les malalties humanes degut a les seves semblances anatòmiques i fisiològiques amb l'espècie humana. Un model animal o

model biomèdic és un substitut per un ésser humà, o un sistema biològic humà, que es pot utilitzar per comprendre la funció normal i anormal del genotip al fenotip i per proporcionar una base per al tractament preventiu o la intervenció terapèutica en malalties humanes (Swanson *et al.*, 2004).

L'ús dels animals en l'estudi de la fisiologia i anatomia humana remunta al segle XVI (Kuzmuk i Schook, 2011). Primer s'utilitzaven com a observacions en condicions estipulades; més endavant ja es va començar la inducció de l'estat de malaltia de manera química i física. A principis del segle XX es van començar a infectar animals amb malalties infeccioses per testar medicaments antibacterians. L'evolució de l'ús de models animals al segle XX ha estat molt important amb l'ús de tècniques com la recombinació genètica, la clonació o l'ús de cèl·lules mare per crear organismes genèticament modificats i establir la condicions desitjades en l'estudi de malalties humanes.

Tradicionalment, els animals més utilitzats com a models han estat els rosegadors (ratolins i rates) i alguns mamífers com els gossos. Tot i així, malgrat les facilitats de maneig que presenten i les semblances fisiològiques en alguns òrgans i sistemes, sovint no són un bon model per a moltes malalties ja que no són prou semblants anatòmicament i no mimetitzen de manera fidel els símptomes ni desenvolupen els desordres com ho farien els humans. El porc és un mamífer que presenta una mida i una fisiologia semblant a l'espècie humana, té un desenvolupament dels òrgans i una progressió de les malalties que el fa ser un model ideal per ser estudiat amb les tecnologies que s'empren en humans (Lunney, 2007). A més a més, l'oportunitat d'aprofitar tots els garrins que surten en una ventrada juntament amb les tècniques de clonació i models transgènics facilita el mapatge genètic. També és una espècie que té nombroses línies cel·lulars desenvolupades en un ampli ventall de teixits que facilita i amplia les possibilitats dels estudis científics.

Un altre avantatge en l'ús del porc com a model biomèdic és que el seu genoma ja està totalment seqüenciat i anotat ([http://www.ensembl.org/Sus\\_scrofa/Info/Index](http://www.ensembl.org/Sus_scrofa/Info/Index)), amb 21.630 gens codificants (RNAs missatgers) i 3.124 gens no codificants (RNAs no

codificants) (Groenen *et al.*, 2012). Comparativament parlant, el genoma porcí, comprès de 18 cromosomes autosòmics i 2 de sexuals (X i Y), té una mida d'uns 2,7Gb i és un 7% més petit que el genoma humà (el genoma del gos i del ratolí ho són un 14%). A nivell nucleotídic, el genoma porcí és 3 vegades més similar al genoma humà que no pas el genoma del gos o del ratolí (Walters *et al.*, 2012).

Així, el porc s'ha utilitzat en estudis de patologies humanes en diferents camps com la fisiologia cardíaca, funció reproductiva, transplantaments, fisiologia de la pell, vessaments cerebrals i reacció de certes drogues al cervell, fisiologia i nutrició de l'intestí, models biomecànics, enginyeria tissular, funció respiratòria i malalties infeccioses (Kuzmuk i Schook, 2011). A més a més, el porc també s'utilitza en l'estudi de patologies i malalties pròpies de l'espècie.

#### **1.1.6. El virus de la malaltia d'Aujeszky**

La malaltia d'Aujeszky, també coneguda com a Pseudoràbia, és una malaltia contagiosa causada per un virus que afecta la majoria dels animals domèstics, fonamentalment l'espècie porcina. No afecta el primats majors, inclosa l'espècie humana. El porc és l'hoste primari del virus i el principal disseminador de la malaltia. En els porcs, el virus es pot mantenir en fase de latència, mostrant signes de recuperació, essent així l'única espècie capaç de sobreviure a una infecció productiva i posteriorment actuar com a reservori (Pejsak i Truszczynski, 2006). La infecció viral en altres espècies causa desordres neurològics i la mort posterior. Per tant, l'eliminació del virus en l'espècie porcina acabaria eradicant la malaltia. Els porcs senglars també poden contraure la malaltia i es consideren reservoris potencials d'infecció dels porcs domèstics.

És una malaltia de gran importància econòmica al sector de producció porcina, sobretot pels sistemes que es concentren en grans granges situades relativament a prop les unes de les altres. El comerç internacional intensiu ha permès el moviment de grans quantitats d'animals, particularment de porcs i els seus productes derivats, i també ha ajudat a la disseminació de la malaltia. Causa importants pèrdues econòmiques en les explotacions porcines, principalment per reducció de la mida de la

garrinada, avortaments i creixement lent dels animals, així com per les restriccions de moviments relacionats amb el control d'aquesta. Tot i així, el progrés de la biologia molecular els últims 15 anys ha permès millores notables en el diagnòstic, control i prevenció de la malaltia. Es tracta d'una malaltia de declaració obligatòria segons l'OIE (Oficina Internacional d'Epizooties). A Espanya està en fase avançada d'eradicació i dades del 2011 de la Generalitat de Catalunya apunten que Catalunya ja és un territori lliure de la malaltia.

L'agent etiològic de la malaltia és l'Herpesvirus porcí tipus 1 (SHV-1), un virus amb un genoma lineal de doble cadena de DNA d'unes 142 Kb. El seu genoma està seqüenciat (Klupp *et al.*, 2004) i codifica uns 70 gens. Una característica biològica dels herpesvirus és la capacitat de mantenir latent una infecció circularitzant el seu genoma i persistent com un episoma amb una expressió gènica viral molt limitada (Boss *et al.*, 2009). Així, podríem classificar els seus gens en dos principals grups, els lítics, que s'expressen durant la infecció, i els latents, que s'expressen durant la fase de latència. Durant l'estat de latència, l'expressió gènica viral està restringida a la transcripció d'una part concreta del genoma anomenada transcrit de latència gran (LLT), de la qual es generen els transcrits associats a la latència (LAT). Les bases moleculars de la latència en els herpesvirus són encara desconegudes. Els principals punts on el virus roman latent són el gangli trigemin, el bulb olfatori i la tonsil·la (Pomeranz *et al.*, 2005). En aquests llocs, el DNA viral es pot detectar en absència d'infecció vírica productiva. Hi ha diverses línies cel·lulars que s'usen a molts estudis científics que permeten el cultiu al laboratori de l'SHV-1, com la línia cel·lular PK-15 derivada del ronyó de porc (Pejsak i Truszczynski, 2006). En aquesta tesi s'ha fet servir aquesta línia cel·lular per a realitzar una infecció experimental *in vitro* amb l'SHV-1.

El virus és transmès principalment per contacte físic entre porcs a través de les secrecions. També es pot transmetre per via sexual, transplacentària, per fòmits i per via aerògena. Després de la infecció oronasal, el virus es replica primerament als teixits epitelials i pot entrar directament a les terminacions nervioses sensorials per la nasofaringe. El període normal d'incubació del virus és de 2 a 6 dies. El virus infecta el sistema nerviós central i altres òrgans, com el tracte respiratori. Segons l'edat dels

animals infectats i el seu estat immunològic, la ruta d'infecció i la virulència de la soca viral, els signes clínics poden ser més o menys severes. La mortalitat és del 100% en garrins de menys de dues setmanes, on el quadre que presenten és neurològic. A mesura que augmenta l'edat, la mortalitat baixa fins al 50% en garrins de 3-4 setmanes i fins al 10% en garrins de 5 a 9 setmanes. El porcs d'engreix o adults presenten un quadre respiratori sovint amb pneumònia, amb una morbiditat del 100% i una mortalitat de l'1-2%. Les femelles reproductores prenyades que s'infecten al primer terç de l'embaràs pateixen una reabsorció dels fetus; si la infecció esdevé al segon o tercer trimestre, presenten quadres d'avortament, momificacions i maceracions dels fetus. En els animals vacunats, la infecció cursa de forma subclínica o poc aparent (Pejsak i Truszczynski, 2006). En la present tesi s'ha realitzat una infecció experimental *in vivo* amb porcs de 4 setmanes i s'ha treballat amb el bulb olfactori i el gangli trigemin, punts de replicació inicial.

Actualment, el principal objectiu és l'eradicació de la malaltia. Segons l'estat de cada país o regió s'aplica un programa de prevenció o un altre. Al país/regió on el virus hi és present, s'usa principalment el programa de vacunació-eradicació, amb vacunes marcades. Primer es fa una campanya de vacunació, i per controlar l'aparició de brots es fan controls serològics dels anticossos específics de les soques silvestres, que es diferencien de les soques vacunals ja que aquestes últimes han estat genèticament modificades amb l'eliminació de l'expressió d'alguna glicoproteïna, com la glicoproteïna E (gE), que juga un paper important en la disseminació del virus pel sistema nerviós. Si apareixen brots es sacrifiquen tots els animals infectats i els que han estat en contacte amb l'animal i, quan el país/regió té la malaltia controlada amb la vacuna, s'elimina l'ús d'aquesta (Pejsak i Truszczynski, 2006). En el cas d'un país/regió amb una prevalença molt alta, primer es realitzen els tests i es sacrifiquen els animals i després s'inicia la vacunació general. La vacunació no evita completament la infecció, però permet evitar l'aparició de casos clínics, disminueix la durada de l'excreció viral, en dificulta la transmissió i ajuda a reduir la prevalença.



## 1.2. Expressió gènica

L'expressió gènica és el procés pel qual s'interpreta el material genètic d'un organisme, és a dir, amb la informació que un gen amaga a la seva cadena de DNA es sintetitza un producte que serà funcional per al desenvolupament de l'organisme. La transcripció del DNA a RNA, la maduració d'aquest (*splicing*) i, posteriorment, la traducció del transcrit a proteïna són els passos necessaris per a l'expressió del genoma, del genotip al fenotip. Bona part dels gens del genoma codifiquen un producte funcional que és una proteïna (gens codificants de proteïnes), tot i que hi ha gens que el seu producte funcional és un RNA que no codifica per cap proteïna, els RNAs no codificants de proteïnes: ribosomals, de transferència, nuclears, nucleolars, microRNAs i altres RNAs de mida petita, etc.

Així, els efectors reals del fenotip són el transcriptoma i el proteoma. El transcriptoma és el conjunt de transcrits que s'expressa en una cèl·lula, en un espai i un moment concret. En l'espècie humana, el genoma conté uns 23.500 gens que codifiquen per alguna proteïna, generant més de 140.000 transcrits diferents i, per tant, indicant que el transcriptoma és força més complex que el genoma. El genoma del porc està estimat que codifiqui per uns 20.000 – 25.000 gens, i tot i que el nombre de transcrits no està determinat s'estima que sigui semblant al de l'espècie humana (Groenen *et al.*, 2011).

De fet, la diversitat fenotípica observada no s'explica només per la variabilitat genòmica observada a nivell de seqüència, sinó que també hi ha tot un conjunt de mecanismes involucrats en la modificació transcripcional i post transcripcional (Ruan *et al.*, 2004). Així doncs, totes les passes en el procés de transformació del DNA en una proteïna estan sotmeses a una regulació que modula tots aquests productes.

Les cèl·lules eucariotes tenen mecanismes per a poder reconèixer les condicions ambientals en les que han d'activar o reprimir la transcripció dels gens corresponents i són capaces d'activar i desactivar la transcripció de cada gen o grup de gens específics. Les propietats biològiques de cada organisme eucariòtic venen determinades en gran mesura per les proteïnes expressades, que en determinen la seva arquitectura, les activitats enzimàtiques, les seves interaccions amb l'ambient i moltes altres propietats

fisiològiques. Tot i així, en un moment determinat de la vida de la cèl·lula tan sols s'expressa una fracció dels RNAs i les proteïnes codificades al seu genoma. En moments diferents, el perfil dels productes gènics expressats pot diferir enormement, tant respecte a quines proteïnes s'expressen com a quins nivells ho fan (Griffiths *et al.*, 2008). És tota aquesta regulació que fa que un organisme pluricel·lular estigui compost per diferents teixits amb funcions específiques, el desenvolupament i la funcionalitat dels quals dependran del conjunt de proteïnes selectivament expressades per cada cèl·lula que el compona.

Desxifrar el transcriptoma d'una cèl·lula o teixit en un moment concret ajuda a entendre les bases de la seva funció biològica. Una aproximació àmpliament usada en l'anàlisi de transcriptomes és comparar els nivells d'expressió de l'RNA missatger en teixits específics o cèl·lules en unes condicions o tractament concret amb l'objectiu d'identificar els gens que estan diferencialment expressats amb funcions biològiques rellevants. A més a més, l'anàlisi de transcriptomes es pot fer servir per identificar nous biomarcadors que podrien ser utilitzats com a eines de diagnòstic.

### **1.2.1. Mecanismes de regulació de l'expressió gènica**

La regulació de l'expressió gènica té lloc a diferents nivells: durant el processament de l'RNA primari, durant el transport de l'RNA missatger del nucli al citoplasma cel·lular, durant la traducció de l'RNA a proteïna, durant el procés d'activació de la proteïna... però majoritàriament es produeix en la regulació de la transcripció del DNA a l'RNA. Tot i així, la regulació als diferents nivells no és independent, sinó que estan connectats i coordinats (Dahan *et al.*, 2011, Orphanides i Reinberg, 2002). Tradicionalment, la regulació de l'expressió gènica a nivell transcripcional ha estat el nivell que ha rebut la major atenció científica per ser el pas més important i per ser fàcil d'estudiar amb mètodes ja establerts (Mata *et al.*, 2005). La importància d'aquest nivell de regulació queda demostrada amb el 5-10% de gens codificants que el genoma dels animals dedica a regular la transcripció (Levine i Tjian, 2003).

També existeix un segon nivell de regulació de l'expressió gènica basat en les alteracions i modificacions de la conformació i estructura del DNA, i que no impliquen

cap canvi en la seqüència nucleotídica: els mecanismes epigenètics, com la metilació del DNA i la modificació d'histones, en el que l'empaquetament i compactació de la cromatina exerceixen un paper regulador per a l'expressió dels gens (Griffiths *et al.*, 2008, Bell i Spector, 2011).

En els últims anys, l'estudi de la regulació post transcripcional ha anat guanyant terreny des del descobriment, tot just farà 20 anys, de la primera molècula d'RNA de mida petita que no codificava per una proteïna, un microRNA (Lee *et al.*, 1993, Wightman *et al.*, 1993). Tot i així, no va ésser fins al 2000 que el món científic no es va adonar de la importància en la modulació de l'expressió gènica que tenien aquests RNAs de mida petita i que es trobaven presents en tot el regne animal (Pasquinelli *et al.*, 2000, Reinhart *et al.*, 2000). La biologia molecular d'aquests RNAs de mida petita s'ha anat descrivint als tres dominis de la vida que coneixem avui dia (eucariotes, bacteris i arqueobacteris, (Woese *et al.*, 1990)) i s'ha demostrat que aquest procés actua com a defensa cel·lular davant àcids nucleics foranis i també com a regulació post transcripcional de transcrits endògens, un procés que s'ha descrit com a interferència per RNAs de mida petita – RNAi (Agrawal *et al.*, 2003).

La regulació gènica duta a terme pels RNAs de mida petita es fa a través d'interaccions específiques de seqüència, on l'RNA petit s'uneix al transcrit diana, que pot ser des d'un RNA missatger fins a un RNA genòmic derivat d'un agent patogen, influenciant la seva integritat i funcionalitat. El coneixement d'aquest mecanisme s'ha anat estenent aquests últims anys, i actualment ja és sabut que aquest mecanisme és universal tot i que els seus usos fisiològics poden ser únics (tenOever, 2013). Els bacteris fan servir un sistema de defensa basat en RNAs curts de repeticions palindròmiques agrupades (crRNAs) per evadir o restringir la infecció per fags. Els PIWI-interacting RNAs (piRNAs) són un sistema de defensa contra els transposons i material genètic forani utilitzat pels organismes multicel·lulars. Els nematodes, els artròpodes i les plantes tenen un sistema immune antiviral innat (viRNAs). També hi ha un altre grup d'RNAs de mida petita conservats als eucariotes, semblants als viRNAs però pensats per guiar i reforçar la identitat cel·lular mitjançant la regulació post transcripcional d'RNAs missatgers endògens, els microRNAs.

#### 1.2.1.1. Els microRNAs

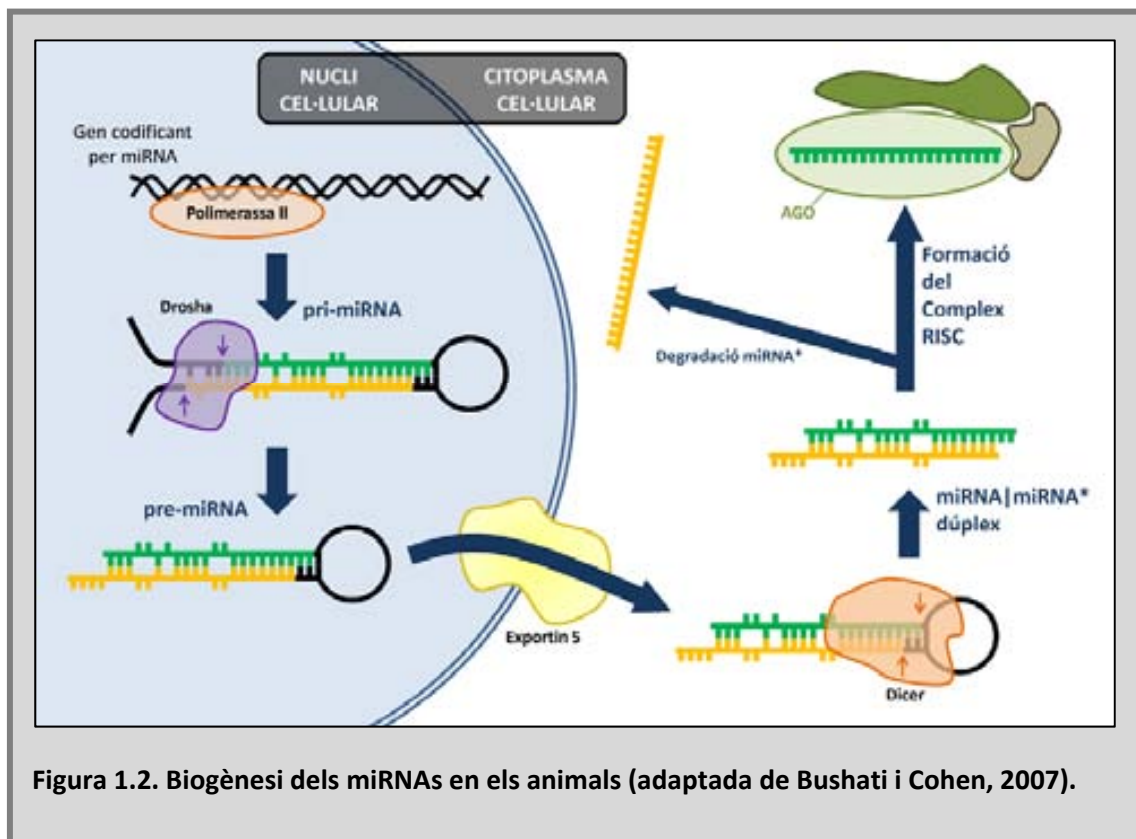
Els microRNAs (miRNAs) són RNAs no codificants de mida petita que acostumen a tenir entre 18 i 25 nucleòtids de llargada. Exerceixen una funció reguladora a nivell post transcripcional generalment inhibidora dels gens pels quals fa diana, inhabilitant l'RNA missatger o impedit la seva traducció. Participen en la regulació de quasi tots els processos cel·lulars investigats i els canvis en la seva expressió poden estar associats a processos patològics (Krol *et al.*, 2010).

Actualment, els microRNAs estan descrits als regnes dels animals, les plantes i els protozous, i també en alguns virus de DNA, i se sap que tenen un paper clau en la regulació d'un ampli ventall de processos biològics (Zheng *et al.*, 2013, Berezikov, 2011, Ambros, 2008). On estan més estudiats és en els animals i les plantes. En mamífers està predit que poden controlar fins a un 50%-60% del transcriptoma (Krol *et al.*, 2010, Friedman *et al.*, 2009).

El primer microRNA va ésser descobert en el nematode *Caenorhabditis elegans* mentre es caracteritzaven els gens que controlaven la temporalització del desenvolupament larvari. Es va descobrir que el gen *lin-4* no codificava per cap proteïna, però si transcrivía per un RNA de mida petita (Lee *et al.*, 1993). També es va observar que aquest RNA s'unia a la regió 3' UTR d'un altre gen, el *lin-14*, impedit la traducció del seu RNA missatger i disminuint així la quantitat de proteïna, afectant la seva expressió (Wightman *et al.*, 1993). Set anys més tard es va descobrir un altre gen a *C. elegans*, el *let-7*, que codificava per un altre RNA de mida petita i aquest també inhibia la traducció del gen *lin-14*, entre d'altres (Pasquinelli *et al.*, 2000, Reinhart *et al.*, 2000). Es va començar a pensar que aquests RNA petits podien formar part d'algun grup d'RNAs reguladors i que, per tant, n'hi podia haver més, a part de ser presents a altres organismes. Així, es van descriure els primers microRNAs i es va veure que n'hi havia de conservats entre espècies (Lee i Ambros, 2001, Lagos-Quintana *et al.*, 2001, Lau *et al.*, 2001).

### 1.2.1.1.1. Biogènesi i mecanismes d'actuació

Tant la biogènesi com els mecanismes d'interacció varien entre animals i plantes. En els animals, els miRNAs són processats inicialment com a transcrits primaris (pri-miRNAs) i pateixen un doble procés de maduració per esdevenir miRNAs madurs (Figura 1.2.). La primera digestió es produeix al nucli cel·lular i és duta a terme per la ribonucleasa III Drosha, que a partir del pri-miRNA genera el precursor del miRNA (pre-miRNA), de 70-80 parells de bases (Lee *et al.*, 2003). La segona digestió té lloc al citoplasma cel·lular després del transport del pre-miRNA per la proteïna transportadora Exportin-5 (Yi *et al.*, 2003). Aquest cop, la ribonucleasa III Dicer genera el miRNA dúplex (miRNA|miRNA\*) a partir del pre-miRNA (Lee *et al.*, 2002). Aquest microRNA dúplex presenta una cadena madura (miRNA) i una cadena complementària (miRNA\*). Una de les dues cadenes, generalment la cadena madura, s'uneix a una proteïna de la família Argonauta (AGO), entre d'altres, per formar el complex silenciador induït per RNAs (RISC). La cadena complementària se sol degradar, tot i que s'ha vist que en alguns casos també pot arribar a formar part del complex RISC com a cadena madura (Okamura *et al.*, 2008).



El miRNA incorporat dirigeix el complex RISC al seu RNA missatger diana i s'hi uneix per complementaritat de bases amb la cadena del miRNA. Aquesta complementaritat requereix ser perfecta en la seqüència que va dels 2 als 8 primers nucleòtids de la cadena madura del miRNA (lectura 5'-3'), zona anomenada *seed*. La resta de la cadena és complementària parcialment (Bartel, 2009). Finalment, l'RNA missatger es degrada o s'atura la seva traducció, disminuint els nivells de proteïna en qualsevol cas (Carthew i Sontheimer, 2009). A més a més, l'existència de variabilitat genètica als extrems 5' i 3' dels miRNAs madurs genera un llarg espectre de variacions de seqüències anomenades isoformes de miRNAs o isomiRs (Morin *et al.*, 2008). Els mecanismes involucrats en la generació dels isomiRs i la seva rellevància biològica han incrementat la complexitat dels mecanismes moleculars relacionats amb la regulació de l'expressió dels RNAs missatgers pels miRNAs.

#### **1.2.1.1.2. Conservació i especificitat**

Les diferències entre els processos de biogènesi i d'actuació dels miRNAs entre els animals i les plantes fan pensar que han evolucionat de manera independent entre els dos regnes, presentant miRNAs específics a cada regne. Tot i així, també es coneix que en els mamífers la majoria de miRNAs estan conservats entre espècies properes, i poden tenir miRNAs homòlegs en espècies més distants, proposant, doncs, que les funcions d'aquests es podrien haver conservat en l'evolució dels llinatges animals (Lim *et al.*, 2003). Tanmateix, Arteaga-Vázquez *et al.*, (2006) postulen que podrien existir miRNAs compartits entre els dos regnes, i juntament amb estudis que proposen l'existència de miRNAs a organismes eucariotes unicel·lulars (Zheng *et al.*, 2013, Huang *et al.*, 2012, Lin *et al.*, 2009) fan pensar en un aparell regulador eucariòtic ancestral comú, justificant les diferències existents en l'actualitat com a producte d'una pèrdua progressiva d'especificitat i no pas producte de la innovació (Tarver *et al.*, 2012).

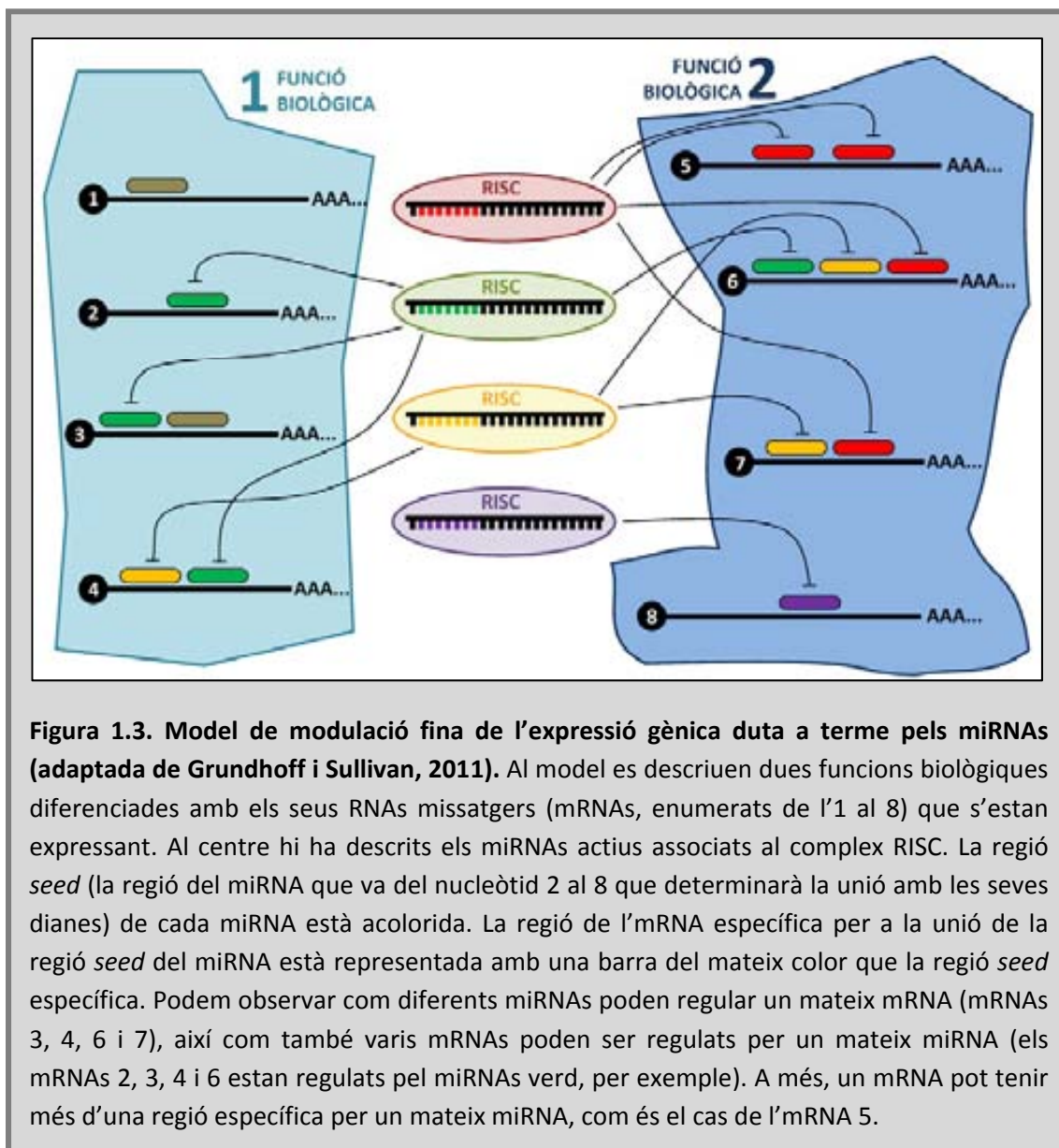
Els mètodes de seqüenciació massiva de nova generació, juntament amb les anàlisis computacionals van esdevenir l'aproximació de referència en el descobriment de nous miRNAs (Berezikov, 2011). Així, la descripció del perfil de miRNAs (miRNAome) en diferents teixits sota diferents estats fisiològics i patològics ha suscitat un gran interès durant els darrers anys (Xie *et al.*, 2011, Ribeiro-dos-Santos *et al.*, 2010, Shao *et al.*,

2010, Glazov *et al.*, 2008, Huang *et al.*, 2008). S'ha vist que, generalment, els miRNAs mantenen un patró d'expressió temporal i espacial específic (Berezikov, 2011). Així doncs, la determinació de patrons d'expressió d'RNAs missatgers alterats relacionats amb una malaltia, desordre, o tractament específic pot ser de gran ajuda en la identificació de miRNAs diferencialment expressats, i el seu ús com a nous biomarcadors (Patel i Sauter, 2011, Fendler *et al.*, 2011, Catto *et al.*, 2011, Chow *et al.*, 2010, Godwin *et al.*, 2010).

#### **1.2.1.1.3. Anotació i validació de la seva expressió**

L'auge que va experimentar aquest nou camp de recerca va ésser molt important. Per ordenar i establir criteris de nomenclatura i descripció dels miRNAs que s'anaven descobrint, es va elaborar un protocol d'anotació i validació experimental d'aquests (Ambros *et al.*, 2003) i una base de dades de miRNAs, anomenada miRBase (Griffiths-Jones, 2004). Aquests criteris d'anotació i nomenclatura han anat canviant amb el temps, a mesura que s'han anat descrivint miRNAs i s'han anat descobrint les particularitats en el regne animal i el regne vegetal (Kozomara i Griffiths-Jones, 2011, Meyers *et al.*, 2008, Griffiths-Jones *et al.*, 2008, Griffiths-Jones *et al.*, 2006). Actualment, a miRBase hi ha descrits més de 25.000 miRNAs en 189 espècies (miRBase v19, agost 2012). En la majoria d'organismes hi ha un nombre limitat de miRNAs comparat amb el nombre d'RNAs missatgers i proteïnes. En els genomes dels mamífers s'estima que hi ha codificats uns 1.000 miRNAs (Bartel, 2004). Si es creu, per exemple, que el genoma humà codifica per 30.000 RNAs missatgers (Pritchard *et al.*, 2012), un microRNA podria actuar en més d'un RNA missatger, i un RNA missatger podria ser regulat per més d'un miRNA, creant així una xarxa de regulació gènica complexa amb una modulació precisa (Figura 1.3., (Bartel, 2009)).

Tot i així, el nombre de miRNAs presents en un genoma podria ser més elevat del que s'espera ja que ja hi ha descrits 2.042 miRNAs a l'espècie humana. Tanmateix, només dues espècies superen els 1.000 miRNAs descrits: l'espècie humana i el ratolí. En el porc, tot i la seva importància en el sector productiu i biomèdic, tan sols hi ha descrits 306 miRNAs (miRBase v19, agost 2012). S'han descrit alguns miRNAs en diferents teixits de porc, com el múscul, la grassa, el cor, el fetge, el timus, l'intestí i els testicles



(Lian *et al.*, 2012, Sharbati *et al.*, 2010, Cho *et al.*, 2010, Reddy *et al.*, 2009, Nielsen *et al.*, 2010), i també en cultius cel·lulars *in vitro*, com la línia cel·lular PK-15, derivades de cèl·lules de l'epiteli renal porcí, o les cèl·lules dendrítiques porcines (Wu *et al.*, 2012, Anselmo *et al.*, 2011, Li *et al.*, 2011).

#### 1.2.1.1.4. Predicció de gens diana i anàlisi funcionals de miRNAs

Una segona etapa en l'estudi dels miRNAs va sorgir quan, a part de descriure'ls mitjançant les tècniques de seqüenciació massiva, van aparèixer metodologies per a predir la seva funcionalitat *in silico*. Eines com miRanda (Betel *et al.*, 2008, John *et al.*, 2004), PicTar (Lall *et al.*, 2006), TargetScan (Garcia *et al.*, 2011) i Diana microT



(Maragkakis *et al.*, 2009a, Maragkakis *et al.*, 2009b) han ajudat a trobar aquells RNAs missatgers pels quals un miRNA pot fer diana tenint en compte les peculiaritats de complementaritat de bases explicades anteriorment, sobretot en l'extrem 5' del miRNA madur (Bartel, 2009). Una vegada descrits els RNAs missatgers diana calia intentar dotar-los d'algun tipus de funcionalitat, assignant-los a vies o rutes metabòliques per donar algun tipus de significat biològic. En aquest sentit, eines com WebGestalt (Zhang *et al.*, 2005), que fa servir les bases de dades Gene Ontology (Ashburner *et al.*, 2000) i KEGG (Ogata *et al.*, 1999, Kanehisa i Goto, 2000), ofereix la possibilitat de treballar amb el conjunt de dianes per al miRNA estudiat donant informació sobre en quines rutes metabòliques i processos biològics participen. Tot i així, aquest protocol *in silico* és molt generalista i tan sols serveix per donar una visió primària per començar a especular sobre les funcions dels miRNAs que s'estan estudiant. Per confirmar la funció d'un miRNA és aconsellable fer estudis funcionals *in vivo*.

#### 1.2.1.1.5. Els miRNAs virals

Els miRNAs virals són aquells miRNAs codificats pel genoma d'un virus. Els primers miRNAs virals van ser descrits en un herpesvirus (Pfeffer *et al.*, 2004). Tots els virus que codifiquen miRNAs són virus de DNA genòmic. De fet, la majoria dels miRNAs virals descrits avui dia (95%) estan descrits en un herpesvirus (Grundhoff i Sullivan, 2011). Això pot ser degut a la grandària del genoma dels herpesvirus comparada amb els altres virus en els que s'han trobat miRNAs (poliomavirus, adenovirus, ascovirus, baculovirus). Una peculiaritat dels herpesvirus és que presenta dues fases diferenciades en el seu cicle vital, el cicle lític d'infecció i el cicle latent, on el virus es manté a la cèl·lula hoste sense causar infecció aparent. Per tant, un dels beneficis d'expressar miRNAs virals en aquests virus seria el fet de poder regular l'expressió gènica de l'hoste i també la seva, evitant així els factors proteics exposats a la resposta immune antigènica de l'hoste (Grundhoff i Sullivan, 2011, Skalsky i Cullen, 2010). Aquesta teoria guanya força si tenim en compte que la majoria d'herpesvirus estudiats expressen els miRNAs durant la fase de latència, inclús sent una fase en la que l'expressió gènica és molt restrictiva. Així, una de les funcions que se'ls atribueix és la regulació de l'expressió gènica viral, i més concretament, la regulació de la persistència

de la latència viral, que la podria aconseguir aprofitant-se del mecanisme conservat de regulació gènica de l'hoste, de manera que els miRNAs virals podrien ajudar a mantenir un ambient cel·lular adient (Skalsky i Cullen, 2010). Un objectiu en el futur és poder inhibir aquests miRNAs virals per poder eliminar la infecció latent dels herpesvirus (Grundhoff i Sullivan, 2011). Altres funcions podrien ser l'alteració del cicle cel·lular i l'esquivament de les defenses de l'hoste. S'ha vist que alguns miRNAs virals imiten alguns miRNAs de l'hoste interactuant amb transcrits de l'hoste mitjançant una complementarietat imperfecta (Skalsky i Cullen, 2010). Aquests últims anys, doncs, s'ha vist la importància de la regulació gènica a través dels miRNAs en processos d'interacció hoste-patògen, no només des d'una visió de l'hoste per combatre la infecció, sinó també des del punt de vista del virus per afavorir els interessos d'aquest.

Concretament en l'herpesvirus porcí tipus 1 (SHV-1) hi ha descrits un total de 13 miRNAs virals a miRbase, tots localitzats al transcrit gran de latència (LLT). Els dos estudis que els descriuen, un en cèl·lules dendrítiques (Anselmo *et al.*, 2011) i l'altre en la línia cel·lular PK-15 (Wu *et al.*, 2012), els associen a la fase de latència del virus precisament per la seva localització, tot i que està demostrat que tenen algun paper durant el procés d'infecció ja que s'ha validat la seva expressió en aquesta fase. Molts d'aquests miRNAs virals també s'han associat als principals gens reguladors de l'SHV-1, com el gen regulador primerenc *EPO*. Aquest gen transcriu la proteïna primerenca 0, encarregada d'activar l'expressió gènica dels promotors virals *IE180*, *UL23* i *US4*. Concretament l'*IE180* és l'únic gen immediat primerenc que codifica l'SHV-1 (Pomeranz *et al.*, 2005), un potent activador transcripcional.

### **1.2.2. Mesura de l'expressió gènica**

Històricament, la mesura de l'expressió gènica a nivell del transcriptoma (RNA missatger) ha anat evolucionant amb el desenvolupament de noves tècniques que han anat cobrint les necessitats científiques de cada moment. La mesura de l'expressió gènica es va començar a realitzar de manera individual per a cada gen estudiat gràcies a la tècnica d'hibridació *Northern blot* (Alwine *et al.*, 1977), que mesura el nivell d'expressió dels transcrits amb la intensitat relativa que genera la banda hibridada. Altres tècniques que també es varen emprar per comparar perfils d'expressió o per

determinar l'expressió de gens foren l' *Expression Sequence Tags* (ESTs, Adams *et al.*, 1991), el *Differential Display PCR* (DD-PCR, Liang i Pardee, 1992), el Ribonuclease Protection Assay (RPA, Gilman *et al.*, 1989) o el *Serial Analysis of Gene Expression* (SAGE, Velculescu *et al.*, 1995). Aquesta darrera tècnica ja permetia l'estudi de milers de transcrits utilitzant la seqüenciació del DNA. Una altra tecnologia molt utilitzada que també permet estudiar l'expressió de gens a gran escala són els Microarrays (Schena *et al.*, 1995).

Una mica més tard va ésser desenvolupada una variant de la reacció en cadena de la polimerasa (PCR – *Polymerase Chain Reaction*) que permet amplificar i quantificar en termes absoluts o relatius el producte d'amplificació, per això s'anomena PCR quantitativa (qPCR, Heid *et al.*, 1996). Com a novetat, la qPCR utilitza un fluoròfor com a marcador, que en un termociclador dotat de sensors per mesurar la fluorescència, després de ser excitat a la longitud d'ona adequada, permet mesurar la quantitat de producte generat. Aquesta mesura es pren després de cada cicle d'amplificació, motiu pel qual la tècnica també se l'anomena PCR a Temps Real (*Real Time PCR*). Aquesta tecnologia també es pot emprar amb l'RNA, aplicant una reacció de transcripció reversa (RT) prèvia a la qPCR. La transcripció reversa és una tècnica emprada per a generar DNA complementari (cDNA) a partir de l'RNA amb l'enzim transcriptasa reversa. Així, l'RT-qPCR ha permès mesurar també els nivells d'expressió de l'RNA a través del seu cDNA, una tècnica que és més acurada i específica per mesurar el nivell d'expressió dels RNAs missatgers.

L'aparició de les tècniques seqüenciació massiva, durant la primera meitat de la dècada dels 2.000, va obrir la porta a la seqüenciació i estudi del genoma complet d'una espècie com l'humana i també l'estudi i anàlisi del transcriptoma a nivell del genoma complet, també conegut com seqüenciació d'RNA (RNA-seq, Mardis, 2008). Generar milions de seqüències ha suposat que les tècniques de seqüenciació de nova generació siguin semi-quantitatives i et permetin fer estudis d'expressió gènica aproximats. Tot i així, els resultats d'expressió gènica obtinguts amb aquestes tècniques encara necessiten ser validats individualment per a cada gen, on la qPCR és la tècnica més utilitzada per a dur a terme aquestes validacions.

### **1.2.2.1. Tècniques de seqüenciació massiva**

#### **1.2.2.1.1. Primera generació de seqüenciació: el mètode de Sanger**

El concepte de seqüenciació del DNA va ésser introduït per Sanger el 1975 descrivint un mètode ràpid de seqüenciació basat en la síntesi de DNA amb uns encebadors i l'enzim DNA polimerasa (Sanger i Coulson, 1975). El 1977 va publicar la tècnica enzimàtica de seqüenciació del DNA en la que utilitzava dideoxinucleòtids, i provocava la finalització de l'amplificació de la cadena quan se n'incorporava un (Sanger *et al.*, 1977). D'aquesta manera, realitzant l'amplificació en quatre reaccions separades (un per a cada dideoxinucleòtid) i la posterior electroforesi podies esbrinar la seqüència de la cadena amplificada. Amb l'automatització de la tècnica va aparèixer el primer seqüenciador, l'ABI Prism 310, comercialitzat per Applied Biosystems (1996). Amb la posterior evolució del mètode, utilitzant capil·lars amb una matriu de polímer en comptes d'un gel, va aparèixer el seqüenciador ABI Prism 3700 d'Applied Biosystems (1998). Aquest seqüenciador ja utilitzava 96 capil·lars simultàniament i és el que es va fer servir per la seqüenciació del genoma humà (Yamey, 2000). Des d'ençà, l'evolució de les tècniques de seqüenciació del DNA han evolucionat de manera vertiginosa fent servir mètodes que realment són millors del mètode inicial dideoxi ideat per Sanger (Pareek *et al.*, 2011).

#### **1.2.2.1.2. Segona generació de seqüenciació: seqüenciació en paral·lel massiva amb pre-amplificació**

L'avenç dels darrers anys en les tècniques de seqüenciació, anomenades tècniques de seqüenciació massiva (HTS – *High Throughput Sequencing*) o tècniques de seqüenciació de nova generació (NGS – *Next Generation Sequencing*) ha permès obtenir milions de seqüències de manera més econòmica i reduir el temps de treball. Actualment hi ha quatre principals tecnologies associades a la seqüenciació en paral·lel massiva amb una característica en comú: la necessitat de realitzar una amplificació prèvia a la seqüenciació (Taula 1.1).

#### ***Piroseqüenciació***

El primer aparell NGS que va sortir al mercat va ésser el seqüenciador GS 20 de 454 Life Sciences (2005). La tècnica desenvolupada utilitza la PCR en emulsió (aigua en oli)

combinada amb la piroseqüenciació (Margulies *et al.*, 2005). Més endavant, l'empresa Roche va adquirir 454 Life Sciences traient al mercat el seqüenciador 454 GS FLX. El seqüenciador 454 GS FLX treballa amb el principi de la seqüenciació per síntesi de DNA, ja que fan servir un enzim DNA polimerasa. Durant la piroseqüenciació, cada nucleòtid incorporat per la DNA polimerasa allibera un pirofosfat, que inicia una sèrie de reaccions que finalment acaben produint llum gràcies a l'enzim luciferasa (Figura 1.4.). Justament aquesta quantitat de llum emesa és la que es detecta i és proporcional al nombre de nucleòtids incorporats. Durant la síntesi del DNA, cada dNTP és alliberat de manera seqüencial, així s'associa cada pic de llum al nucleòtid alliberat en aquell moment (Mardis, 2008).

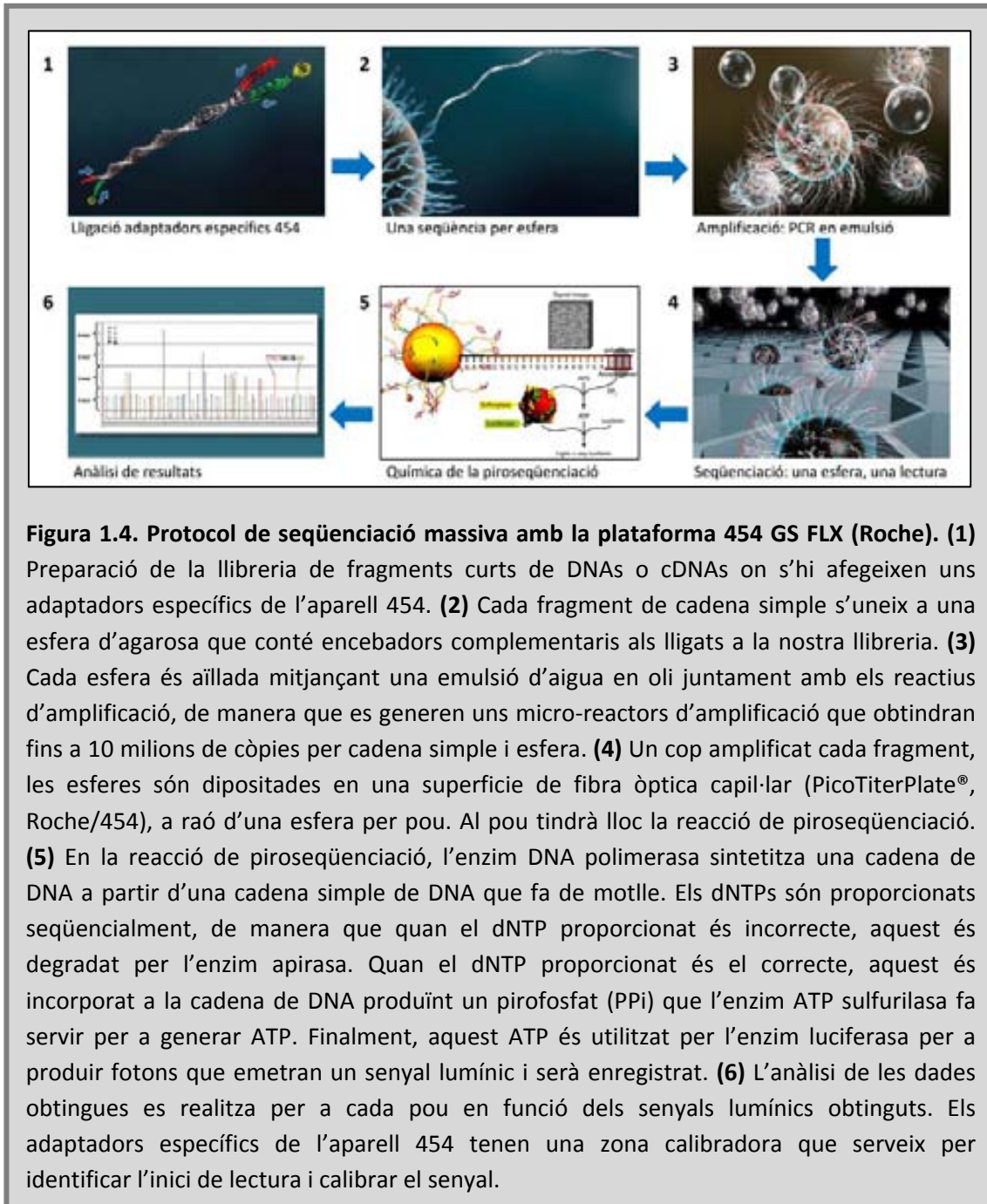
En la present tesi s'ha implementat la tècnica de la piroseqüenciació per a generar llibreries d'RNA de mida petita en el ronyó de diferents races porcines amb l'objectiu de descriure el perfil de miRNAs presents, i també per a detectar nous miRNAs.

#### ***Fluoròfor finalitzador reversible***

Aquesta tecnologia també està basada en la seqüenciació per síntesi. En aquest cas, però, cada dNTP té un fluoròfor associat determinat a l'extrem 3' i això fa que la DNA polimerasa tan sols pugui afegir un nucleòtid a la seqüència. Els dNTPs, al tenir un fluoròfor específic, poden ser alliberats simultàniament. Un cop afegit el nucleòtid pertinent a la seqüència, el làser llegeix la fluorescència emesa pel nucleòtid incorporat i tot seguit s'elimina el fluoròfor de manera química i així allibera l'extrem 3', deixant-lo preparat per al proper cicle d'amplificació. Aquest és el sistema dissenyat per Solexa que utilitza l'aparell *Genome Analyzer* comercialitzat per Illumina (Shendure i Ji, 2008).

#### ***Seqüenciació per lligacions encadenades***

Aquesta tecnologia va ésser desenvolupada per Applied Biosystems i és la que fa servir la plataforma SOLiD™ (2007). El principi que utilitza aquesta metodologia és la seqüenciació per lligació, ja que usa una DNA lligasa. Està basat en la lligació seqüencial mitjançant sondes d'oligonucleòtids a través de la PCR en emulsió (Metzker, 2010).



### Seqüenciació per síntesi amb sensibilitat iònica

Una altra tecnologia comercialitzada per l'empresa Life Technologies (actualment l'empresa Applied Biosystems en forma part) és el sistema *Ion Torrent Semiconductor Technology* (*PostLight™ Sequencing Technology*), amb la plataforma Ion Torrent (2010). El principi tecnològic d'aquest sistema està basat en un procés químic tan senzill com el canvi de pH que es produeix quan un nucleòtid és incorporat a la

seqüència que s'està amplificant a causa de l'ió d'hidrogen que s'allibera (Rothberg *et al.*, 2011, Pourmand *et al.*, 2006). En aquest cas, cada nucleòtid és alliberat de manera seqüencial. Aquest canvi de detecció òptica per elèctrica presenta certs avantatges: no és necessari fer servir nucleòtids modificats, un estalvi econòmic al no utilitzar tecnologia òptica i un menor temps de processament de les dades, ja que no s'acumulen grans quantitats de dades d'imatge de fluorescència.

Aquesta tecnologia s'ha emprat en aquesta tesi per a crear les llibreries en l'estudi de la infecció experimental *in vitro* i *in vivo* amb el virus de la malaltia d'Aujesky.

#### **1.2.2.1.3. Tercera generació de seqüenciació: seqüenciació en paral·lel massiva de molècula única**

La segona onada d'aparells de seqüenciació massiva, o aparells de seqüenciació de tercera generació, presenta l'avantatge de treballar seqüenciant molècules de DNA que no han estat amplificades prèviament, reduint així el biaix produït pels errors de la DNA polimerasa i disminuint també la distorsió dels nivells d'abundància relativa, fent que els resultats siguin més fiables a nivell d'expressió gènica (Stranneheim i Lundeberg, 2012). La varietat de tecnologies desenvolupades és molt gran, algunes de les quals ja s'allunyen força del mètode de seqüenciació ideat per Sanger (Taula 1.1.).

##### ***Fluoròfor finalitzador únic reversible***

Tecnologia utilitzada per la primera plataforma d'aquestes característiques, l'aparell *Heliscope™ Single Molecule Sequencer* (2008), que va ésser comercialitzat per Helicos Biosciences. El sistema utilitza nucleòtids marcats de manera reversible en una aproximació de seqüenciació per síntesi en un DNA immobilitzat en una superfície (Bowers *et al.*, 2009).

##### ***Seqüenciació amb DNA polimerasa immobilitzada***

Un altre aparell comercialitzat per Pacific Biosciences i que utilitza el mètode de seqüenciació anomenat SMRT (*Single-Nucleotide Real Time sequencing*, 2009) és el PACBIO RS. En aquest cas també es fan servir nucleòtids marcats amb un fluoròfor, però no estan adjunts a la base nucleotídica, sinó al fosfat extern, fet que permet que

la DNA polimerasa no s'hagi d'anar aturant a cada incorporació nucleotídica. Això implica que aquesta metodologia és fins a vint mil cops més ràpida que els aparells de segona generació, i també té un cost més baix. En aquest cas, la DNA polimerasa està immobilitzada al terra d'una nanoestructura anomenada *Zero-Mode Waveguide* (ZMW), que ofereix una lectura òptica molt confinada, minimitzant el soroll de fons (Korlach *et al.*, 2010, Eid *et al.*, 2009).

### **Seqüenciació en nanoporus**

D'altres tecnologies s'estan fent un lloc dins les plataformes de seqüenciació massiva d'última generació. Aquest és el cas de la tecnologia de seqüenciació en nanoporus (Branton *et al.*, 2008). Aquest sistema no fa servir nucleòtids etiquetats amb fluoròfors, sinó que està basat en la modulació de la corrent iònica que hi ha dins un nanoporus proteic quan hi passa un nucleòtid, podent-lo identificar. Aquesta tecnologia encara està en desenvolupament tot i que l'empresa Oxford Nanopore Technologies ja ha presentat el dispositiu GridION™ que serà comercialitzat aquest 2013.

Els avantatges que presenten els aparells de seqüenciació massiva de tercera generació respecte els de segona generació són: un major rendiment, una lectura de seqüències més llarga i precisa, una quantitat de material de partida molt menor i un menor cost (Pareek *et al.*, 2011).

#### **1.1.1.1.1. Aplicació de les tècniques de seqüenciació massiva a l'estudi de miRNAs**

L'estudi del perfil dels miRNAs i l'anàlisi de diferències en l'expressió d'aquests en diferents situacions, així com el descobriment de nous miRNAs, ha generat molt interès durant els últims anys. Les plataformes de seqüenciació masiva, amb protocols de treball específics per a generar llibreries d'RNA, han jugat un paper clau en la determinació d'aquests RNAs no codificants de mida petita. Nombrosos estudis han utilitzat aquestes tècniques per a descriure perfils de miRNAs en diferents teixits del porc, com al múscul esquelètic (Nielsen *et al.*, 2010), al greix dorsal (Chen *et al.*, 2012)



**Taula 1.1.** Síntesi de l'evolució de les plataformes de seqüenciació massiva.

	<b>Plataforma</b>	<b>Principi de seqüenciació</b>	<b>Preparació de la mostra</b>	<b>Principi de detecció</b>	<b>Rendiment per mostra</b>
1a generació	Mètode de Sanger	Finalització enzimàtica de la cadena	Clonació/PCR	Òptica	+
2a generació: seqüenciació en paral·lel massiva amb pre-amplificació	454 GS FLX/ Roche	Piroseqüenciació	PCR en emulsió	Òptica	++
	Genome Analyzer/ Illumina	Seqüenciació per síntesi amb fluoròfor finalitzador reversible	Amplificació pont	Òptica	+++
	SOLID™/ ABI	Seqüenciació per lligació	PCR en emulsió	Òptica	+++
	Ion Torrent/ Life Technologies	Seqüenciació per síntesi amb sensibilitat iònica	PCR en emulsió	Electrònica	++
3a generació: seqüenciació en paral·lel massiva de molècula única	Heliscope™/ Helicos Biosciences	Seqüenciació per síntesi amb fluoròfor finalitzador únic reversible	No amplificació	Òptica	++
	PACBIO RS/ Pacific Biosciences	Seqüenciació amb DNA polimerasa immobilitzada (SMRT)	No amplificació	Òptica	++
	GridION™/ Oxford Nanopore Technologies	Canvis de corrent iònica en un nanoporus	No amplificació	Electrònica	+++

a l'intestí (Sharbati *et al.*, 2010) o als testicles (Lian *et al.*, 2012). També s'han utilitzat per a observar canvis d'expressió en el perfil dels miRNAs durant el desenvolupament d'una patologia, malaltia o procés fisiològic, com en la miogènesi (Hou *et al.*, 2012) o en un cas de pleuropneumònia bacteriana (Podolska *et al.*, 2012).

Així doncs, en la present tesi s'han utilitzat les metodologies de la piroseqüenciació (454 GS FLX/Roche) i la seqüenciació per síntesi amb sensibilitat iònica (Ion Torrent/Life Technologies) per a descriure el perfil d'expressió dels miRNAs en el ronyó del porc i també en una infecció experimental amb el virus de la malaltia d'Aujeszky. L'estudi de l'expressió dels miRNAs i la validació dels nous miRNAs descrits s'ha avaluat mitjançant la tècnica de l'RT-qPCR.

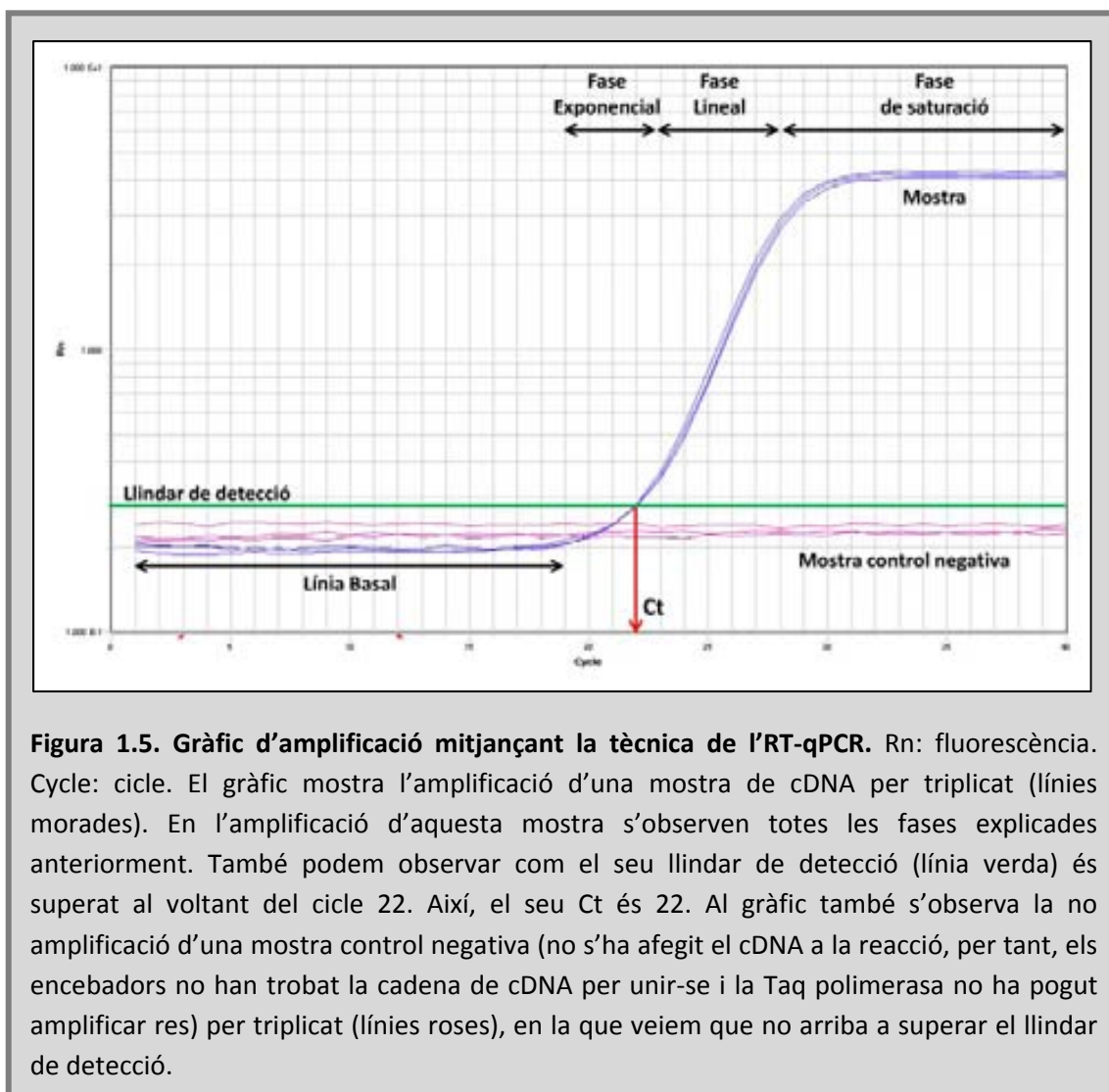
#### **1.1.1.2. La tècnica de la reacció en cadena de la polimerasa a temps real mitjançant transcripció reversa (RT-qPCR)**

##### **1.1.1.2.1. Metodologia de l'RT-qPCR**

La PCR quantitativa a temps real amb transcripció reversa (RT-qPCR) ha esdevingut una de les tècniques d'elecció en la detecció i estudi de l'mRNA i altres tipus d'RNA (Bustin, 2000). La necessitat de quantificar l'RNA i l'aparició de les tècniques de seqüenciació massiva han convertit l'RT-qPCR en una eina d'ús generalitzat en molts camps, com la genòmica funcional, la biologia molecular, la virologia i la microbiologia, la forense, i també en la biotecnologia. Alguns factors han contribuït a l'auge en la utilització d'aquesta tècnica, com (i) la seva homogeneïtat en el disseny, (ii) evita la necessitat de la manipulació post-PCR, (iii) té un rang dinàmic ample ( $> 7$  ordres de magnitud ( $10^7$ )) que permet la comparació entre RNAs que poden diferir força en la seva abundància i (iv) la tècnica afegeix un potencial quantitatiu a la tècnica de la PCR, que fa que el disseny sigui qualitatiu i quantitatiu (Bustin *et al.*, 2005, Ginzinger, 2002). És un mètode molt sensible per a la detecció i quantificació dels nivells d'expressió gènica, inclús per a trobar petits canvis en els nivells d'expressió de l'mRNA, especialment en situacions on la concentració d'RNA no és la desitjada, ja sigui per transcrits poc abundants, teixits amb poca concentració d'RNA o amb limitació de volum mostra (Pfaffl, 2004).

La reacció de l'RT-qPCR manté les mateixes bases que la PCR convencional. Primerament, l'RNA és convertit a cDNA a través de la reacció de la transcripció reversa. Després, mitjançant l'ús de dos encebadors que flanquejaren la seqüència de cDNA, aquest serà utilitzat com a motlle del material a amplificar per una DNA polimerasa termoestable, que durant la reacció utilitzarà els dNTPs per anar amplificant la seqüència. També s'utilitzarà un tampó de la reacció que conté ions de magnesi. La reacció es duu a terme mitjançant cicles de temperatura. Cada cicle està format per tres fases; una primera fase de desnaturalització, a temperatura elevada (95°C), que permetrà que la doble cadena de cDNA se separi; una segona fase d'hibridació, on la temperatura disminueix per permetre la unió dels encebadors amb les cadenes de cDNA; i una tercera fase d'elongació a una temperatura òptima per l'enzim que realitza l'amplificació de la cadena. L'enzim utilitzat és la Taq polimerasa, la temperatura òptima de la qual és 72°C. Actualment, però, existeixen Taq polimerases que treballen a temperatures més baixes (60°C) i permeten fer la hibridació i l'elongació a una mateixa temperatura. Per poder quantificar el producte amplificat a cada cicle cal utilitzar un agent fluorescent. A mesura que es vagi sintetitzant producte, la fluorescència anirà augmentant i serà proporcional a la quantitat de producte amplificat. El patró d'amplificació que segueix una RT-qPCR és el que es mostra a la Figura 1.5.

La primera fase és la que el sensor no detecta un augment de fluorescència i manté una línia basal que és el soroll de fons de la reacció. Arriba un moment que la quantitat de producte que s'amplifica ja és detectable i s'inicia la fase exponencial de l'amplificació. És en aquesta fase on realitzarem la quantificació del producte, ja que l'amplificació és precisa i es comporta de manera exponencial, és a dir, les reaccions es comporten igual ( $2^n$ ). Per fer el càlcul de la quantificació utilitzarem el valor anomenat Ct (Cicle Threshold), que és el número de cicle d'amplificació que supera el llindar establert de fluorescència emesa. Així, el Ct és inversament proporcional al nivell d'expressió: quan més gran sigui el Ct, més baixa serà l'expressió del producte amplificat perquè haurà trigat més cicles a creuar el llindar de detecció. Seguida la fase exponencial arriba la fase lineal, en que la reacció d'amplificació perd eficàcia i es perd la correlació entre el nombre de cicles i el producte amplificat. Finalment arribem a



l'última fase, anomenada fase de saturació o altiplà, on la reacció d'amplificació deixa de produir-se per diverses raons, com la manca d'enzim o de dNTPs, o la inhibició per excés de DNA.

Hi ha diferents maneres d'utilitzar els agents fluorescents en una RT-qPCR: amb agents no específics intercal·lants o amb sondes específiques. Els agents intercal·lants, com el SYBR GREEN, s'uneixen inespecíficament a la doble cadena de DNA, independentment de si la cadena amplificada és del producte que estem quantificant o d'altres productes que es poden estar amplificant per contaminacions, inespecificitats o dímers d'encebadors. El control de l'especificitat es realitza mitjançant una corba de dissociació al final de l'RT-qPCR, per assegurar que el producte amplificat és el correcte

i no hi ha altres amplificacions. Aquesta corba s'obté augmentant gradualment la temperatura de manera el senyal de fluorescència anirà disminuint. Arribarà un moment que la doble cadena de DNA es desnaturalitzarà i s'alliberarà tot l'agent intercal·lant, produint una baixada notòria de la fluorescència. Si es realitza la primera derivada negativa de la correlació temperatura:fluorescència, mostrarà un pic que serà la temperatura de fusió del producte amplificat. Si hi ha més d'un pic ens estarà indicant que hi ha més d'un producte amplificat.

Les sondes específiques tenen diferents característiques. Hi ha sondes amb la cadena nucleotídica modificada químicament per millorar les propietats d'unió a la cadena diana, com els encebadors LNA (Locked Nucleic Acids - LNA<sup>TM</sup>, Exiqon) o els LUX (Lux Upon eXtension - LUX<sup>TM</sup>, Invitrogen). Una altra característica de les sondes específiques és la forma en que emeten fluorescència i quants fluoròfors fan servir. En aquest sentit, una de les sondes més utilitzades són les sondes TaqMan<sup>®</sup> (Life Technologies), que emeten fluorescència després de la hidròlisi quan el fluoròfor s'allibera de la sonda i s'allunya del seu agent extintor que absorbia la fluorescència. Una altra manera d'emetre fluorescència és quan té lloc la hibridació, mitjançant dos encebadors marcats cadascun amb un fluoròfor, un donador i l'altre receptor (Hybprobe, Roche). També existeixen sondes en forma de forqueta que emeten la fluorescència quan s'hibriden a la cadena diana (Molecular Beacons; LUX<sup>TM</sup>, Invitrogen) (Kubista *et al.*, 2006).

#### **1.1.1.2.2. Estratègies de quantificació de l'expressió gènica mitjançant l'RT-qPCR**

La mesura de l'expressió gènica es pot realitzar en termes absoluts o relatius. L'estratègia per a poder calcular l'expressió gènica amb quantificació absoluta és relacionar el senyal d'amplificació de la mostra problema amb la concentració de DNA, emprant una corba de calibratge a través d'un banc de dilucions seriades on es coneix el nombre de molècules presents a cada dilució. La dificultat d'obtenir mostres amb concentracions conegudes del gen estudiat, així com moltes vegades no cal saber el nombre de còpies exactes però sí quantes vegades s'expressa de més una mostra respecte una altra, fa que la quantificació relativa sigui l'aproximació més utilitzada

(Pfaffl, 2004). Així doncs, en la quantificació relativa es compara l'expressió de la mostra problema respecte l'expressió d'un gen control, que serveix de referència. Les unitats que s'utilitzen per expressar les quantitats relatives són irrelevantes doncs. L'expressió del gen de referència ha de ser estable per a les mostres analitzades i no s'ha de veure afectada per cap tipus de paràmetre que s'estigui estudiant (tractament, malaltia, condició ambiental o fisiològica...). La normalització de les dades és essencial per controlar factors de variació com la quantitat d'RNA inicial empleada en l'RT, la degradació d'aquest durant el processament de la mostra, les diferències en la qualitat de l'RNA utilitzat, variacions durant la realització de la tècnica com errors de pipeteig...

Existeixen diferents estratègies de quantificació relativa de l'expressió gènica. Al llarg dels anys s'han establert diferents models matemàtics basats en la comparació del valor de Ct. El més popular és el conegut com a mètode del  $2^{-\Delta\Delta Ct}$  (Schmittgen i Livak, 2008, Livak i Schmittgen, 2001), que compara la concentració entre mostres calculant la diferència de Ct entre un gen diana i un gen de referència i el  $\Delta Ct$  de diferents mostres es compara directament. Aquest mètode assumeix que les eficiències d'amplificació entre el gen problema i els gens de referència són idèntiques o quasi idèntiques i molt properes al 100%. El model de Pfaffl, també conegut com a mètode de correcció de l'eficiència, que incorpora al càlcul el valor real d'eficiència dels gens estudiats, és també molt utilitzat (Pfaffl, 2001). Finalment, sempre hi ha la possibilitat d'introduir una corba estàndard en tots els assajos experimentals, mètode conegut com a quantificació relativa amb corba estàndard (Applied Biosystems 1997, Technical Bulletin #2), a partir de la qual s'obtindrà un valor anomenat *quantity* que simula un valor de quantitat relatiu per a cada mostra. Aquest últim mètode, que té en compte les petites variacions d'eficiència que pugui haver-hi entre els diferents experiments realitzats per un mateix assaig, ha estat l'utilitzat en les validacions de miRNAs mitjançant RT-qPCR en els treballs d'aquesta tesi.

#### **1.1.1.2.3. Adaptació de l'RT-qPCR per a l'anàlisi de miRNAs**

Tal i com s'ha esmentat anteriorment, la tècnica de l'RT-qPCR és utilitzada per validar el perfil d'expressió que s'ha observat dels miRNAs mitjançant tècniques d'estudi del genoma complet o el transcriptoma, com els microarrays o la seqüenciació massiva.

Per a realitzar una RT-qPCR que generi uns resultats creïbles i reproduïbles, cal tenir en compte una sèrie de paràmetres i, en el cas dels miRNAs, alguns prenen més rellevància degut a les seves característiques.

Un factor important és l'extracció d'RNA i la seva qualitat. Per valorar la qualitat de l'RNA es pot utilitzar el nombre RIN (*RNA integrity number*), un valor que atorga l'aparell Agilent 2100 Bioanalyzer (Agilent Technologies). Per a una bona qualitat de l'RNA s'aconsella que el valor RIN no sigui inferior a set (Schroeder *et al.*, 2006).

Un altre factor que pren una major importància en els cas dels miRNAs és el pas d'RNA a cDNA. La síntesi del cDNA és una mica més complexa en els cas dels miRNAs ja que al ser RNAs de cadena curta (18-25 nucleòtids) és més complicat fer un bon disseny d'encebadors. Una aproximació que es fa servir és l'ús d'encebadors específics per miRNAs, que poden ser lineals o acabats en forqueta a l'extrem 5' (Chen *et al.*, 2005). L'avantatge dels encebadors que acaben en forqueta és que només el miRNA madur s'hi podrà unir i, per tant, evitarà l'amplificació de pre-miRNAs i pri-miRNAs, que també contenen la seqüència madura del miRNA i els encebadors lineals també els amplifiquen (Benes i Castoldi, 2010). El desavantatge d'aquesta aproximació és que al ser específica del miRNA, el rendiment que es pot obtenir de la mostra és baix si es vol treballar amb varis miRNAs. Una altra aproximació és l'ús d'un encebador universal on prèviament s'ha elongat l'extrem 3' dels miRNAs amb una cua d'adenines amb l'enzim Poly(A) polimerasa (PAP) (Shi i Chiang, 2005). D'aquesta manera es pot dissenyar un encebador que estarà compost per una seqüència de timines a l'extrem 5' (complementària a la cua d'adenines) seguida d'una seqüència universal degenerada que permetrà unir l'encebador a tots els miRNAs que tinguin una cua d'adenines. Hi ha altres aproximacions no tant utilitzades com circularitzar el miRNA amb una lligasa (Kumar *et al.*, 2011) o unir un adaptador al miRNA també amb una lligasa (Benes i Castoldi, 2010).

El disseny dels encebadors per fer la qPCR també és important en l'estudi de l'expressió dels miRNAs, ja que aquests presenten una elevada heterogeneïtat en el contingut de G/C i fa que la temperatura de fusió entre el miRNA i el seu encebador

específic pugui ser molt dispar. El disseny dels encebadors específics va lligat a la manera com s'ha sintetitzat el cDNA i també al mètode utilitzat per detectar l'amplicó. Un disseny força comú és l'ús d'un encebador específic del miRNA i un altre d'universal en la qPCR i detectar l'amplificació amb sondes TaqMan® (Life Technologies). Malgrat ser sondes específiques, si en la síntesi del cDNA no es fa servir un encebador específic, no s'assegurarà que la sonda TaqMan® s'uneixi a l'amplicó desitjat. També es pot detectar l'amplicó amb el l'agent intercal·lant SYBR Green, més utilitzat quan es fa servir l'aproximació de la poliadenilació en la síntesi del cDNA, però menys específic que les sondes TaqMan®. Un avantatge del SYBR Green, però, és que permet fer un control d'especificitat realitzant la corba de dissociació. Una altra aproximació molt utilitzada és la tècnica *Universal RT microRNA PCR* (Exiqon), que combina la poliadenilació en l'RT i l'ús d'encebadors específics i modificats (Locked Nucleic Acid – LNA™, Exiqon), augmentant la temperatura de fusió i, així, l'especificitat. L'amplicó és detectat amb SYBR Green. Tot i així, els encebadors LNA™ presenten una amplificació poc acurada ja que fan de mal motlle per a les DNA polimerases (Veedu *et al.*, 2007). Finalment hi ha una altra aproximació que utilitza la poliadenilació amb un encebador universal per a realitzar l'RT i dos encebadors específics del miRNA amb detecció de l'amplicó mitjançant SYBR Green. El problema de la temperatura de fusió és solventat afegint una cua de nucleòtids a cada encebador, de manera que augmenta l'especificitat i l'eficiència en la reacció d'amplificació (Balcells *et al.*, 2011). En els treballs experimentals duts a terme en aquesta tesi s'han utilitzat les aproximacions *Universal RT microRNA PCR* (Exiqon) i els encebadors específics de miRNA amb cua de nucleòtids.

#### **1.1.1.2.4. L'ús de microRNAs com a gens de referència**

L'últim paràmetre que cal tenir en compte a l'hora de validar l'expressió dels miRNAs mitjançant quantificació relativa a través de l'RT-qPCR és la selecció dels gens de referència per la normalització de les dades obtingudes. Originàriament les estratègies de normalització contemplaven l'ús d'un únic gen de referència. Aquesta idea, però, ha evolucionat a diferents aproximacions en la que es fan servir més d'un gen de referència, des del mètode de normalització amb una mitjana global (Mestdagh *et al.*, 2009) fins a l'ús de varis gens de referència (Vandesompele *et al.*, 2002).



Els gens de referència més utilitzats en estudis d'expressió de miRNAs són els RNAs ribosomals com el 5S RNA (Gu *et al.*, 2011, Shen *et al.*, 2011, Peltier i Latham, 2008) i els RNAs nuclears com el RNU6B (Mestdagh *et al.*, 2009, Gu *et al.*, 2011, Shen *et al.*, 2011, Peltier i Latham, 2008, Sharkey *et al.*, 2012, Wotschofsky *et al.*, 2011, Feng *et al.*, 2010). Tot i ser molt important que els gens de referència utilitzats siguin de la mateixa naturalesa que el subjecte d'estudi i tinguin la mateixa llargada que aquest per garantir una mateixa eficiència durant l'extracció de l'RNA i la transcripció reversa (Schaefer *et al.*, 2010), l'ús dels miRNAs com a gens de referència encara no és molt popular. És per aquest propòsit que es va creure convenient fer un estudi per avaluar l'estabilitat dels miRNAs en diferents teixits i diferents races de porc per a poder ser utilitzats com a gens de referència en estudis de quantificació relativa.

#### **1.1.1.2.5. La PCR digital, la tercera generació de la PCR**

Si la PCR a temps real es considera la segona generació de l'ús de la PCR convencional, recentment està agafant protagonisme la PCR digital (dPCR). Tot i que la tècnica ja es va descriure el 1997, les primeres plataformes comercials no van començar a sortir al mercat anys més tard (Blow, 2007). Aquesta tècnica es basa en la quantificació absoluta de l'amplificació de la mostra diana. El principi de la tècnica es basa en dividir la mostra en milers de reaccions individuals per així amplificar cada molècula de la mostra en un pou diferent. Al final de la reacció, mitjançant fluorescència, es determina si la molècula s'ha amplificat i s'assumeix una distribució de Poisson per a la mostra dividida, assignant 0 o 1 en funció si hi ha hagut amplificació o no, assumint que a cada pou hi haurà 0 o 1 molècules. Amb aquesta aproximació es desvincula el nombre de cicles d'amplificació de la quantitat de mostra i no cal fer servir gens de referència. Els controls negatius de la tècnica, però, prenen força per garantir que les amplificacions positives no són falsos positius. La precisió de la quantificació absoluta permet la detecció de lleugeres diferències en la variació del nombre de còpies d'un gen o detectar diferències en l'expressió d'mRNA de fins a 1,2x vegades. Actualment diverses plataformes ja s'estan comercialitzant (qdPCR 37K™ IFC, Fluidigm; QuantStudio™ 3D, Life technologies; QX100, Bio-Rad; RainDrop™, RainDance Technologies).



## 2. OBJECTIUS

---

El treball realitzat en la present tesi titulada

### *CARACTERITZACIÓ DE MICRORNAS D'INTERÈS EN L'ESPÈCIE PORCINA*

s'emmarca dins el projecte MIRPIGVIR (AGL2007-66371-C02 i AGL2010-22358-C02-01, finançat pel ministeri de Ciència i Innovació del govern de l'Estat espanyol), els objectius principals del qual han estat la caracterització de microRNAs virals i porcins i l'estudi de la seva funció en les interaccions hoste - patogen durant la infecció amb el virus de la malaltia d'Aujeszky (SHV-1).

Els objectius concrets de la tesi han estat:

1. Caracterització del perfil de microRNAs de ronyó de diferents races porcines i estudi de la seva expressió.
2. Descobriments i caracterització de nous microRNAs porcins.
3. Descobriments i caracterització de microRNAs virals durant una infecció amb el virus de la malaltia d'Aujeszky.
4. Estudi i detecció de microRNAs (porcins i virals) amb una possible relació amb el procés d'infecció del virus de la malaltia d'Aujeszky.
5. Estudi dels microRNAs com a gens de referència en estudis de quantificació relativa mitjançant RT-qPCR.

L'assoliment dels objectius es presenta en format d'articles publicats o en procés d'acceptació:

**miRNA Expression Profile Analysis in Kidney of Different Porcine Breeds**

Timoneda O, Balcells I, Núñez JI, Egea R, Vera G, Castelló A, Tomàs A, Sánchez A  
2013 PLoS ONE 8(1): e55402. doi:10.1371/journal.pone.0055402

*Objectius 1 i 2*

**The Role of Viral and Host microRNAs in the Aujeszky's Disease Virus During the Infection Process**

Timoneda O, Núñez F, Balcells I, Muñoz M, Castelló A, Vera G, Pérez LJ, Egea R, Mir G, Córdoba S, Rosell R, Segalés J, Tomàs A, Sánchez A, Núñez JI  
(manuscrit en preparació per a ser presentat a PLoS ONE)

*Objectius 3 i 4*

**Determination of Reference microRNAs for Relative Quantification in Porcine Tissues**

Timoneda O, Balcells I, Córdoba S, Castelló A, Sánchez A  
2012 PLoS ONE 7(9): e44413. doi:10.1371/journal.pone.0044413

*Objectiu 5*

# 3. ARTICLES I ESTUDIS

---

## 3.1. Estudi del perfil d'expressió dels microRNAs al ronyó de diferents races porcines

### **miRNA Expression Profile Analysis in Kidney of Different Porcine Breeds**

Timoneda O, Balcells I, Núñez JI, Egea R, Vera G, Castelló A, Tomàs A, Sánchez A

2013 PLoS ONE 8(1): e55402. doi:10.1371/journal.pone.0055402



# miRNA Expression Profile Analysis in Kidney of Different Porcine Breeds

Oriol Timoneda<sup>1\*</sup>, Ingrid Balcells<sup>1</sup>, Jose Ignacio Núñez<sup>2</sup>, Raquel Egea<sup>1</sup>, Gonzalo Vera<sup>1</sup>, Anna Castelló<sup>1</sup>, Anna Tomàs<sup>3</sup>, Armand Sánchez<sup>1</sup>

**1** Departament de Genètica Animal, Centre de Recerca en AgriGenòmica (CRAG), Universitat Autònoma de Barcelona, Campus UAB, Bellaterra, Spain, **2** Centre de Recerca en Sanitat Animal (CRESA), UAB, Bellaterra, Spain, **3** Program Infection and Immunity, FISIB, Bunyola, Spain

## Abstract

microRNAs (miRNAs) are important post-transcriptional regulators in eukaryotes that target mRNAs repressing their expression. The uncertain process of pig domestication, with different origin focuses, and the selection process that commercial breeds suffered, have generated a wide spectrum of breeds with clear genetic and phenotypic variability. The aim of this work was to define the miRNAs expression profile in kidney of several porcine breeds. Small RNA libraries from kidney were elaborated and high-throughput sequenced with the 454 Genome Sequencer FLX (Roche). Pigs used were classified into three groups: the European origin group (Iberian breed and European Wild Boar ancestor), European commercial breeds (Landrace, Large White and Piétrain breeds) and breeds with Asian origin (Meishan and Vietnamese breeds). A total of 229 miRNAs were described in the pig kidney miRNA profile, including 110 miRNAs out of the 257 previously described pig miRNAs and 119 orthologous miRNAs. The most expressed miRNAs in pig kidney microRNAome were Hsa-miR-200b-3p, Ssc-miR-125b and Ssc-miR-23b. Moreover, 5 novel porcine miRNAs and 3 orthologous miRNAs could be validated through RT-qPCR. miRNA sequence variation was determined in 116 miRNAs, evidencing the presence of isomiRs. 125 miRNAs were differentially expressed between breed groups. The identification of breed-specific miRNAs, which could be potentially associated to certain phenotypes, is becoming a new tool for the study of the genetic variability underlying complex traits and furthermore, it adds a new layer of complexity to the interesting process of pig evolution.

**Citation:** Timoneda O, Balcells I, Núñez JI, Egea R, Vera G, et al. (2013) miRNA Expression Profile Analysis in Kidney of Different Porcine Breeds. PLoS ONE 8(1): e55402. doi:10.1371/journal.pone.0055402

**Editor:** Yi Xing, University of Iowa, United States of America

**Received:** October 26, 2012; **Accepted:** December 23, 2012; **Published:** January 25, 2013

**Copyright:** © 2013 Timoneda et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the projects AGL2007-66371-C02-01 and AGL2010-22358-C02-01 and by the Consolider-Ingenio 2010 program (CSD2007-00036) from Ministerio de Ciencia e Innovación. OT is recipient of FPI PhD fellowship from Spanish Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: oriol.timoneda@cragenomics.es

## Introduction

MicroRNAs (miRNAs) are small regulatory RNAs that play important roles in the regulation of gene expression [1,2]. These single-stranded non-coding RNAs that are approximately 22 nucleotides long are involved in post-transcriptional regulation mechanisms acting mainly through down-regulation of target messenger RNAs (mRNAs) in a wide range of biological and pathological processes [3,4]. Thus, miRNAs inhibit gene expression by blocking protein translation or inducing mRNA degradation [5–7]. mRNAs can be regulated by several miRNAs, and each miRNA can target hundreds of mRNAs in different binding sites [8,9]. The recent evidence about the genetic variability at both 5' and 3' ends of the mature miRNA sequence generates a large spectrum of sequence variants called miRNA isoforms or isomiRs [10–12]. Mechanisms involved in isomiRs generation and their biological relevance has increased the complexity of molecular mechanisms related to regulation of mRNA expression mediated by miRNAs.

Nowadays, a large number of miRNAs have been reported in animals, plants and viruses, with up to 18,226 entries in the miRBase database (v18, November 2011, URL: <http://www.mirbase.org/>) [13–15]. In mammalian genomes, the number of encoded miRNAs has been predicted to be up to 1000 miRNAs,

comprising about 3% of all protein-coding genes [5]. In mammals, nearly all miRNAs are conserved in closely related species [16], and may have homologous in distant species, suggesting that miRNA functions could also be conserved throughout the evolution of animal lineages [17]. Several studies showed that variability in miRNA sequences has been lost over time since miRNAs have been described in unicellular eukaryotes, showing its deeper evolutionary history among eukaryotes [18].

The pig is an important livestock species, not only for its production in industry, but also as a suitable animal model for comparative genomics and biomedical studies [19]. However, the number of porcine miRNAs available in public databases is still poor, with only 257 porcine miRNAs described in the pig genome compared to the completely sequenced human (1,921), mice (1,157), bovine (676), poultry (544), equine (360) or canine (289) genomes (miRBase v18). The description of the microRNAome (miRNAome) held by different tissues under different pathophysiological states has become of great interest in the last years [20–24]. Determining altered patterns of miRNA expression related to disease or specific treatments or conditions would be useful in order to identify differentially expressed miRNAs that could be used as novel biomarkers [25–30]. In pigs, several miRNAomes have been described in different tissues such as muscle, fat, heart,

liver, thymus, intestine and testes [31–35], or using *in vitro* cells models such as porcine PK-15 cells (derived from porcine kidney epithelial cells) and porcine dendritic cells [36–38]. However, the porcine kidney miRNAome has not yet been described although it is an essential organ involved in functions like blood filtering, gluconeogenesis and in the secretion of important hormones like erythropoietin, renine and vitamin D.

The process of pig domestication has been very complex. Independent geographical origins have been described [39] which have generated multiple phenotypically different breeds [40,41]. In particular, high differences are found regarding reproductive and meat production and quality traits between Asian and European breeds. In addition, pig has become an important production animal for human, and, consequently, it has been strongly selected for traits of economical interest. As a result of this high selection, many commercial breeds have been generated. Different consuming necessities in pig industry have expanded the large phenotypic variability due to different selected traits. Genetic diversity between pig breeds has been further studied [42–44] showing variability in gene expression and elucidating the artificial selection performed by porcine industry during pig domestication [45]. However, the genetic variability held at miRNA sequences, which could imply changes at post-transcriptional level, has not yet been explored although it may lead to great phenotypic differences and/or pathological disorders.

Studies with mitochondrial DNA (mtDNA) assume significant differences between European and Asian pigs prior to domestication and several studies describe the phylogenetic relationships between European and Asian pig breeds, determining different groups into European diversity and other groups into Asian diversity [39,44,46]. Other studies confirmed some introgression events of Asian pig breeds in most European commercial breeds and evidence that local Spanish breeds have never been introgressed with Asian neither with European commercial populations [47–49]. In this sense, studying miRNAs differentially expressed between European and Asian pig breeds could contribute to understand their phenotypical differences.

In the present work, the miRNAome of several pig breeds belonging to the pure European and Asian branches as well as European commercial breeds that have been introgressed with Asian genes has been undertaken. The main goals of this study were to describe the pig kidney miRNAome through high throughput sequencing (HTS) technology, to study the miRNA sequence variability among porcine breeds and to determinate new porcine miRNAs.

## Materials and Methods

### Samples and RNA preparation

Kidney samples were collected from twenty pigs belonging to six different breeds and the European ancestor (Wild boar). Animals were classified into three main groups based on breed origin; (1) European breeds: Wild boar (WB,  $n = 2$ ) and Iberian (IB,  $n = 4$ ), (2) Asian breeds: Meishan (ME,  $n = 3$ ) and Vietnamese (VT,  $n = 3$ ), and (3) European commercial breeds with strong influence from Asian breeds: Landrace (LD,  $n = 3$ ), Large White (LW,  $n = 2$ ) and Piétrain (PT,  $n = 3$ ). All samples were taken from slaughterhouse (PRIMAYOR, Mollerussa, Spain) under veterinary supervision and they were immediately snap-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until use. Total RNA was isolated using TRIzol<sup>®</sup> reagent (Invitrogen, Carlsbad, USA) following the manufacturer's recommendations. Total RNA was quantified using ND 1000 Nanodrop<sup>®</sup> Spectrophotometer (Thermo Scientific, Wilmington, USA) and its quality was assessed on an Agilent

2100 Bioanalyzer using the RNA 6000 Nano kits (Agilent Technologies, Santa Clara, USA). RNA quality threshold was set to RNA Integrity Number (RIN) above seven. For RT-qPCR validations, additional kidney samples were added ( $n = 29$ : 2 WB, 4 IB, 4 LD, 4 LW, 4 PT, 4 ME and 7 VT).

### Small RNA library construction and high throughput sequencing

Small RNA fraction (15–30 nt) was excised and isolated from denaturing 12.5% polyacrilamide gel electrophoresis (PAGE) using miSpike<sup>™</sup> (IDT<sup>®</sup>, Coralville, USA) as internal size marker. For each breed, 60  $\mu\text{g}$  of total RNA were loaded on separate gels to avoid cross-contamination. Gels were stained with GelStar<sup>®</sup> Acid Nucleic Gel Stain (Lonza, Basel, Switzerland) for UV visualization. Excised small RNA fraction were purified using Performa<sup>®</sup> DTR gel filtration cartridges (EdgeBio, Gaithersburg, USA) and, then, isolated small RNAs were pooled by breed to construct small RNA libraries. Briefly, 3' and 5' linkers from miRCat<sup>™</sup> kit (IDT, Coralville, USA) were ligated at both ends of the small RNAs in two separated reactions using a T4 RNA ligase without ATP (Fermentas, Germany) and T4 RNA ligase with ATP (Ambion, Austin, USA), respectively. Between 3' and 5' primer ligations, the 60 nt RNAs were purified by PAGE to eliminate unligated products. Then, linked products were used to perform a reverse transcription reaction using the SuperScript<sup>™</sup> III Reverse Transcriptase kit (Invitrogen<sup>™</sup>, Carlsbad, USA) and the cDNA obtained was amplified with the Platinum<sup>®</sup> Taq DNA Polymerase High Fidelity kit (Invitrogen<sup>™</sup>, Carlsbad, USA). PCRs were done with primers complementary to 3' and 5' linkers and, in addition, they included multiplex identifiers at the 5' end (a five nucleotide sequence tag) to allow differentiation between libraries (Table S1). Obtained PCR products were purified using the QIAquick PCR Purification Kit (Qiagen<sup>®</sup>, Germany). Libraries were quantified with Qubit<sup>™</sup> fluorometer, Quant-IT<sup>™</sup> (Invitrogen<sup>™</sup>, Carlsbad, USA) and prepared to a  $10^8$  DNA molecules/ $\mu\text{L}$  concentration for sequencing by 454 Genome Sequencer FLX Titanium (Roche, Germany) following the manufacturer's protocol at DNA sequencing facilities at CRAG (Bellaterra, Spain).

### Sequence annotation for miRNA expression profile and validation of selected miRNAs through Reverse Transcription quantitative real time PCR (RT-qPCR)

Primers were trimmed from sequences and only those sequences between 15 and 29 nucleotides and with total number of sequences  $\geq 3$  were kept for further analysis. Sequences were compared to all available miRNA sequences (miRBase v18) using local Blast. Parameters were set to 100% identity and up to 4 mismatches allowed at the end of the sequences to assume variability on 3' and 5' ends [31].

Differences in miRNA expression between breed groups (European breeds: WB and IB, Asian breeds: ME and VT, and European commercial breeds: LD, LW and PT) were assessed using the miRNAs annotated through miRBase database. Total number of sequences obtained for each miRNA was normalised by library size (in counts per thousand) and, then, averaged by group. Fold changes (FC) between groups were calculated using normalised data. According to sequence count, nine miRNAs highly expressed (Hsa-miR-200b-3p, Hsa-miR-200c-3p, Ssc-miR-126, Ssc-miR-126\*, Ssc-miR-99a, Ssc-miR-532-5p, Ssc-miR-92a, Ssc-miR-26a and Bta-miR-193b,  $n > 100$ ) and four miRNAs lowly expressed (Ssc-miR-423-5p, Ssc-miR-29c, Ssc-miR-486 and Ssc-let-7f,  $n < 100$ ) in the kidney miRNAome were selected to measure



their expression levels by RT-qPCR. Hsa-let-7a, Hsa-miR-25 and Hsa-miR-93 were used as reference miRNAs based on literature [50,51]. Primers used were designed by Exiqon® (Denmark). For low expressed miRNAs and Bta-miR-193b primers were designed following the methodology suggested by Balcells et al [52] (Table 1), where it is described that miR-specific quantitative RT-qPCR with DNA primers is a highly specific, sensitive and accurate method for microRNA quantification.

Reverse transcription (RT) reactions were performed in duplicate with the universal cDNA synthesis kit (Exiqon, Denmark) using 300 ng of total RNA following the manufacturer's instructions. Non template controls (NTC) and minus poly(A) polymerase controls for each tissue were included. qPCRs were performed in duplicate using miRCURY™ LNA™ Universal RT microRNA PCR kit (Exiqon, Vedbaek, Denmark) on a 7900HT Sequence Detection System (Applied Biosystems, Warrington, UK). Each qPCR was done in 20 µL total volume including 10 µL SYBR Green master mix (Exiqon, Vedbaek, Denmark), 0.5 µL of each LNA primer set (Exiqon, Denmark) and 5 µL of a 1:20 dilution of the cDNA. To assess qPCR efficiency, standard curves with 10-fold serial dilutions of a pool of equal amounts of cDNA from all samples were included in each assay. Thermal profile was set as follows: 95°C for 10 min and 40 cycles at 95°C for 15 sec and 60°C for 60 sec. NTC and minus poly(A) polymerase controls were included. Melting curve analysis was included at the end of the qPCR to detect unspecific amplifications.

Quantities from each sample were obtained from the calibration (standard) curve added in each RT-qPCR reaction. GeNorm v.3.5 software [53] was used to examine the stability of the reference miRNAs ( $M < 1.5$ ) and to obtain a normalization factor (NF). The quantity obtained from each miRNA and sample was normalised by the NF and FCs were calculated in relation to the lowest normalised value. FCs were  $\log_2$  transformed and expression data were analysed by a two-way analysis of variance (ANOVA) with the General Linear Models procedure of the Statistical Package for the Social Scientists (IBM® SPSS® Statistics 19; IBM Corporation, Armonk, USA) including the RT and breed as fixed factors. Significance threshold was set at  $\alpha < 0.05$ . Scheffé test was used to determine significant differential expression between breed groups.

### Sequence analysis for novel miRNAs discovery and validation through RT-qPCR

All primer trimmed sequences were mapped in the pig genome retrieved from Ensembl Genome Browser (Ensembl release 62, [ftp://ftp.ensembl.org/pub/release-62/fasta/sus\\_scrofa/dna/](ftp://ftp.ensembl.org/pub/release-62/fasta/sus_scrofa/dna/), Sscrofa9.62) considering 100% of alignment and identity (perfect match). Sequences that mapped only once in the genome and with a length from 19 to 23 nt were selected. Among these, only sequences with unknown annotation at Ensembl genome browser and with copy number (CN) higher than 2 were considered. They were clustered taking into account only the position in the genome. Hence, sequences positioned in the same region were grouped and the sequence with higher CN was selected as the reference sequence for each cluster. Fifteen clusters were considered novel miRNAs to be validated. Flanking regions (50 nt) of the selected reference sequences were used to predict pre-miRNA folding structure using MFold software [54] following the guidelines reported by Ambros et al. [55]. At the end, 8 clusters were selected for RT-qPCR validation. RT reactions were performed in duplicate using total RNA as previously described by Balcells et al [52]. Briefly, 600 ng of total RNA in a final volume of 20 µL including 2 µL of 10x poly(A) polymerase buffer, 0.1 mM of ATP, 0.1 mM of each dNTP, 1 µM of RT-primer, 200 U of M-MuLV Reverse Transcriptase (New England Biolabs, USA) and 2 U of poly(A) polymerase (New England Biolabs, USA) was incubated at 42°C for 1 hour and at 95°C for 5 minutes for enzyme inactivation. Non template controls (NTC) and minus poly(A) polymerase controls for each tissue were included.

DNA primers for each miRNA were designed following the methodology suggested by Balcells et al [52] (Table 1). qPCR reactions were performed in duplicate in 20 µL final volume including 10 µL FastStart Universal SYBR Green Master (Roche, Germany), 125–500 nM of each primer (Table 1) and 5 µL of a 1:100 dilution of the cDNA. Standard curves were generated by 10 fold serial dilutions of a pool of all cDNAs in order to calculate the qPCR efficiency. For Cluster 2 (Cl-2), standard curve was performed by using 2-fold serial dilutions. qPCR settings and data analysis was performed as explained above. Hsa-let-7a, Hsa-miR-25 and Hsa-miR-103 were used as reference miRNAs [50,51].

**Table 1.** Primers used for the RT-qPCR validation design.

miRNA/Cluster	Forward primer (5'-3')	Reverse primer (5'-3')	Primer conc. (nM each)
Bta-miR-193b	AGAACTGGCCCAAAAGTC	TCCAGTTTTTTTTTTTTTAGCGG	250
Ssc-miR-423-5p	TGAGGGGACAGAGAGCGA	GTCCAGTTTTTTTTTTTTAAAGTC	250
Ssc-miR-29c	GCACTAGCACCATTTGAAATC	GGTCCAGTTTTTTTTTTTTAACC	250
Ssc-miR-486	GCACTCCTGTACTGAGCTG	GGTCCAGTTTTTTTTTTTTCTCG	250
Ssc-let-7f	CGCAGTGAGGTAGTAGATTG	AGGTCCAGTTTTTTTTTTTTAACT	250
Cl-2	CAGCTGCTATGCCAACA	CCAGTTTTTTTTTTTTTGCCAA	250
Cl-5	GCTGTAACGCAACTCCA	TCCAGTTTTTTTTTTTTTCCAC	500
Cl-15	GCGACCACTCTTGGT	TCCAGTTTTTTTTTTTTTCATGG	No amplification
Cl-16	CAGTTGGTGACCAGGTG	AGTTTTTTTTTTTTTCCCTGA	No amplification
Cl-24	GCTGCATTCTCGGCTG	GGTCCAGTTTTTTTTTTTTAAATTAAG	No amplification
Cl-25	CAGCTGGTGTGTGAATC	GTTTTTTTTTTTTTTCCGGCT	125
Cl-29	GGTTGGGTACTACTGGAA	GTCCAGTTTTTTTTTTTTTAGCT	250
Cl-38	GTCTCCGTTTGCCTGT	CCAGTTTTTTTTTTTTTTCAGCA	250

doi:10.1371/journal.pone.0055402.t001

## Target prediction and functional analysis

DIANA – microT v3.0 web server [56,57] was used to identify *in silico* potential mRNA targets for differentially expressed miRNAs. Porcine genes are not included in the current version of DIANA – microT v3.0 and predictions were based on the human mRNA: miRNA interactions assuming sequence conservation. *In silico* functional annotation of putative mRNA target genes for each miRNA were analyzed with WEB-based Gene Set Analysis Toolkit (WebGestalt, [58]). Predicted miRNA targets were functionally annotated through the biological process information supported by Gene Ontology (GO, [59]) and the pathways in which they were involved were described by using the Kyoto Encyclopedia of Genes and Genomes database (KEGG, [60,61]). Over or under represented functional categories were identified with hypergeometric test corrected by the multiple test adjustment proposed by Benjamini & Hochberg [62]. Significant threshold was set at  $\alpha < 0.05$ .

## Results

### Characterisation of miRNA expression profile in porcine kidney

Small RNA libraries from kidney of six porcine breeds (Iberian, Landrace, Large White, Piétrain, Meishan and Vietnamese) and the European ancestor (Wild Boar) were sequenced in a high throughput 454 GS FLX sequencer (Roche). A total of 115,305 reads (corresponding to 17,913 unique sequences) containing short inserts of length ranging from 15 to 29 nt (corresponding to miRNA size) were obtained (Table 2). The total number of reads obtained per breed ranged from 13,704 to 23,815 counts; however, the Iberian breed yielded a much lower number of total counts (4,022). If we consider the number of unique sequences obtained per breed, as an average, 17% of total counts corresponded to unique sequences except for Iberian and Vietnamese breeds that had 39% and 38% of unique sequences, respectively.

Alignment of sequences to miRBase database revealed that 77.7% (89,594) of the total sequences yielded a positive match to a known miRNA. In total, 229 different miRNAs were found to be expressed in porcine kidney. Among them, 110 miRNAs corresponded to previously described porcine miRNAs and 119 to orthologous miRNAs. The percentage of detected miRNAs per breed varied from 55 to 83%, being the Iberian breed the library with less detected miRNAs and the Large White the most, in agreement with the total number of sequences obtained in these libraries. The proportion of porcine miRNAs and orthologous miRNAs was similar in all libraries, around 50% both. Sixty-two miRNAs had a CN below 3 and were excluded from the analysis. Overall, the porcine kidney miRNA profile was estimated to be of 167 miRNAs (Table S2). The most expressed miRNAs (CN > 350, 0.30%) in porcine kidney are listed in Table 3. The most abundant miRNA was Hsa-miR-200b-3p (27,097, representing 23.50% of all porcine kidney miRNAs), followed by Ssc-miR-125b (8,809; 7.64%), Ssc-miR-23b (5,412; 4.69%), Ssc-miR-126 (5,274; 4.57%) and Ssc-miR-23a (5,156; 4.47%). The miRNA expression pattern was similar between breeds; however, some differences were found. As an example, Ssc-miR-125b was the most expressed miRNA in Iberian breed, followed by Ssc-miR-99a and Hsa-miR-200b-3p, and Ssc-miR-192 and Hsa-miR-200c-3p were the fourth and fifth most expressed miRNAs in Large White breed, respectively. In addition, Hsa-miR-200c-3p appeared as the fifth most expressed miRNA in Meishan breed.

### Sequence miRNA variation

All miRNA sequences showed variability at both ends although it was higher in the 3' end. In some cases, nucleotide variability could be also detected inside the miRNA sequence. In general terms, for each miRNA there was one sequence more expressed called reference miRNA sequence and, then, there was a number of sequence variations with less number of copies (Figure 1). Nucleotide variation was found in 70% of the miRNAs expressed in kidney and only a small proportion of kidney miRNAs did not show isomiRs. Nevertheless, those miRNAs without sequence variations were low expressed, from 3 to 38 total reads. Overall, 116 miRNAs presented isomiRs with an average rate of 10 isomiRs per miRNA. The miRNA with more variants was Hsa-miR-200b-3p with 123 variants, followed by Ssc-miR-23b and Ssc-miR-125b, with 59 and 51 variants, respectively. These three miRNAs are, in fact, the most expressed miRNAs in porcine kidney (Table 3). In this sense, variants expression followed two main patterns: according to those miRNAs with more than 1,000 total reads, they were distributed in those miRNAs with a strong predominant isomiR, such as Hsa-miR-200b-3p, Ssc-miR-125b, Ssc-miR-23b, Ssc-miR-23a, Ssc-miR-192, Ssc-miR-10b, Ssc-miR-126\* and Ssc-miR-10a, and those miRNAs where there is not a really strong predominant isomiR, like Ssc-miR-126, Ssc-miR-99a, Hsa-miR-200c-3p, Ssc-miR-30d and Ssc-miR-125a (Table S3). Furthermore, the most expressed isoform was, in most cases, the same in all breeds; it only differed between breeds in Ssc-miR-99a and Hsa-miR-200c-3p.

Interestingly, reference miRNA sequence for each miRNA was not always the same as the one described in miRBase database. This was the case of 72 miRNAs (Table S4). Sequence differences were mainly found at 3' end of the miRNA (in 90% of cases) and consisted, basically, in the addition or deletion of one or two nucleotides. In Table 4 are shown those most expressed miRNA sequences (CN > 100) differing from the miRBase reference sequence. These differences in miRNA sequences are not influenced by the miRNA abundance, as total reads in these 72 miRNAs include high expressed miRNAs (Ssc-miR-23a, 5,156 reads) and low expressed miRNAs (Mmu-miR-5100, 3 reads). Moreover, counts of these 72 miRNAs represent the 21% of total counts of kidney miRNA profile, a considerable amount.

Furthermore, the most abundant reference sequence for a specific miRNA varied also between breeds. We found 76 miRNAs for which the most abundant sequence in each breed was not the same (data not shown). Nonetheless, differences in isomiR predominance between breeds did not follow any pattern related to breed origin. In addition, for some miRNAs, the described sequence in miRBase was expressed at a low level or even could not be detected in the porcine kidney samples analysed in this study. This was the case of Bta-miR-193b, Ssc-miR-374a, Ssc-miR-145, Ssc-miR-362, and Hsa-miR-324-3p, that had a CN of 21, 13, 3, 0 and 0, respectively.

### Between-breed kidney miRNA differential expression study

Differential miRNA expression analysis was performed between breed groups: EU, EA and AS. miRNAs were considered differentially expressed (DE) when fold change difference between groups was greater than 1.5 or when a miRNA was not expressed in one group although *p-values* were not significant (*p-value* < 0.05). In this sense, 42 miRNAs (25%) were equally expressed among analysed groups whereas 125 miRNAs (75%) were DE between breed groups (Table S5). About 77% of the described miRNAs were present in all groups. Focusing on miRNAs that were differential expressed from one group to the remaining groups, 26

**Table 2.** Summary of sequence alignment with miRBase database<sup>1</sup>.

Library	Item	Total	miRBase homology*	<i>Sus scrofa</i> homology*	Other species homology*	No homology in miRBase
Iberian (IB)	Total counts	4,022	2,748	1,897	851	1,274
	Unique sequences	1,575	512	289	223	1,063
	miRNAs		125	62	63	
Wild Boar (WB)	Total counts	18,391	14,787	7,146	7,641	3,604
	Unique sequences	3,480	1,295	726	569	2,185
	miRNAs		183	92	91	
Landrace (LD)	Total counts	13,704	11,131	4,897	6,234	2,573
	Unique sequences	2,670	982	545	437	1,688
	miRNAs		165	86	79	
Large White (LW)	Total counts	23,815	19,728	10,609	9,119	4,087
	Unique sequences	4,078	1,492	865	627	2,586
	miRNAs		191	99	92	
Piétrain (PT)	Total counts	20,215	16,971	9,890	7,081	3,244
	Unique sequences	3,307	1,351	781	570	1,956
	miRNAs		183	93	90	
Meishan (ME)	Total counts	17,407	14,990	5,380	9,610	2,417
	Unique sequences	2,607	1,079	597	482	1,528
	miRNAs		165	88	77	
Vietnamese (VT)	Total counts	17,751	9,239	5,105	4,134	8,512
	Unique sequences	6,695	1,022	577	445	5,673
	miRNAs		157	82	75	
<b>TOTAL</b>	Total counts	<b>115,305</b>	<b>89,594</b>	<b>44,924</b>	<b>44,670</b>	<b>25,711</b>
	Unique sequences	<b>17,913</b>	<b>3,624</b>	<b>2,007</b>	<b>1,617</b>	<b>14,289</b>
	miRNAs		<b>229</b>	<b>110</b>	<b>119</b>	

<sup>1</sup>: miRBase database (v18, November 2011).

\*Mismatches or gaps have not been allowed into the sequence alignment. Up to 4 mismatches allowed in the ends of the sequences.

doi:10.1371/journal.pone.0055402.t002

miRNAs were up regulated in EU, 25 in EA and 7 in AS. Conversely, 16 miRNAs were down regulated only in EU, 9 in EA and 31 in AS (Table S5). Looking at the most expressed miRNAs, Ssc-miR-125b and Ssc-miR-99a were up regulated in EU, Hsa-miR-200c-3p and Ssc-miR-192 were up regulated in EA and Hsa-miR-200b-3p, Ssc-miR-23a and Ssc-miR-23b were up regulated in AS. Several miRNAs displayed high fold changes ( $FC \geq 5$ ) in at least one comparison (Table S5); however, their CN tended to be low, suggesting that FCs were overestimated. Additionally, looking at normalised counts for each breed (Table S6), there were some miRNAs with low frequency which their expression was higher in some breed, such as Hsa-miR-193b-5p, Hsa-miR-140-5p, Ssc-miR-19b, Ssc-miR-423-5p and Ssc-miR-24 in Iberian breed, Ssc-miR-151-5p, Hsa-miR-4454, Ssc-miR-199a-3p and Ssc-miR-374b-5p in Landrace breed, Ssc-miR-486 in Piétrain breed and Ssc-let-7f and Hsa-let-7i-5p in Meishan breed. Thirteen miRNAs displaying expression differences between breeds by HTS were selected to be validated through RT-qPCR (Table 5). For RT-qPCR validations, additional kidney samples were added ( $n = 29$ : 2 WB, 4 IB, 4 LD, 4 LW, 4 PT, 4 ME and 7 VT). Hsa-miR-200b-3p, Hsa-miR-200c-3p and Bta-miR-193b were chosen to be orthologous miRNAs and thus, not described in pig yet. Ssc-miR-99a, Bta-miR-193b and Ssc-miR-423-5p were selected to be up regulated in EU, Hsa-miR-200b-3p, Ssc-miR-532-5p and Ssc-let-7f to be up regulated in AS and, Hsa-miR-200c-3p, Ssc-miR-92a, Ssc-miR-26a, Ssc-miR-486 and Ssc-miR-29c to be up regulated in

EA. Finally, Ssc-miR-126 and Ssc-miR-126\* were also included to be suggestively up regulated in EA. These miRNAs were mapped in the pig genome sequence (Sscrofa 9.62) and miRNA folding was predicted for Hsa-miR-200b-3p and Hsa-miR-200c-3p using MFold software [54] following the guidelines described by Ambros et al. [55]. Unfortunately, Bta-miR-193b could not be mapped in the pig genome and, consequently, its miRNA folding could not be predicted. RT-qPCR efficiencies obtained were high (ranging from 90% to 110%) and standard curves correlations were at least of 0.99, allowing us to perform a good gene expression quantification [63]. Eight miRNAs (Hsa-miR-200b-3p, Hsa-miR-200c-3p, Ssc-miR-126, Ssc-miR-126\*, Ssc-miR-99a, Bta-miR-193b, Ssc-miR-486 and Ssc-let-7f) were differentially expressed in at least one comparison by RT-qPCR ( $p\text{-value} < 0.05$ ) (Table 6). Relative quantification from RT-qPCR data determined that Hsa-miR-200b-3p, Bta-miR-193b and Ssc-let-7f were up regulated in EU while Ssc-miR-126, Ssc-miR-126\* and Ssc-miR-99a were down regulated in AS. On the other hand, Hsa-miR-200c-3p and Ssc-miR-486 were down and up regulated in EA, respectively. Correlation between HTS and qPCR data was medium to low ( $0.52 > r^2 > 0.003$ ). Therefore, not all the results obtained with both techniques are in agreement.

**Table 3.** Most abundant miRNA profile in pig kidney (CN>350).

miRNA name	Total counts (%)	IsomiRs	IB	WB	LD	LW	PT	ME	VT
Hsa-miR-200b-3p	27,097 (23.50%)	123	239	4,924	4,298	5,163	3,683	6,987	1,803
Ssc-miR-125b	8,809 (7.64%)	51	567	1,245	933	2,170	2,322	798	774
Ssc-miR-23b	5,412 (4.69%)	59	116	992	638	935	1,090	673	968
Ssc-miR-126	5,274 (4.57%)	49	158	928	419	1,308	1,186	660	615
<u>φSsc-miR-23a</u>	5,156 (4.47%)	51	123	828	535	801	1,015	598	1,256
Ssc-miR-192	3,863 (3.35%)	41	89	723	194	1,105	810	484	458
<u>φSsc-miR-99a</u>	3,781 (3.28%)	32	339	615	375	884	888	234	446
Hsa-miR-200c-3p	3,478 (3.02%)	32	42	367	309	1,076	762	631	291
Ssc-miR-10b	2,846 (2.47%)	31	69	589	271	504	519	561	333
Ssc-miR-126*	2,796 (2.43%)	26	60	440	387	503	599	418	389
<u>φSsc-miR-30d</u>	1,977 (1.72%)	31	82	223	142	757	434	238	101
<u>φSsc-miR-125a</u>	1,369 (1.19%)	29	103	192	194	308	288	173	111
Ssc-miR-10a	1,317 (1.14%)	19	43	287	176	219	188	284	120
Ssc-miR-365-3p	986 (0.86%)	16	10	168	173	225	246	100	64
Ssc-miR-92a	797 (0.69%)	17	12	96	105	217	125	145	97
Ssc-miR-204	748 (0.65%)	14	10	61	271	173	133	58	42
Ssc-miR-378	743 (0.64%)	20	67	127	52	179	148	103	67
Ssc-miR-26a	598 (0.52%)	17	3	94	76	159	134	91	41
Bta-miR-200a	485 (0.42%)	9	18	87	99	95	58	101	27
<u>φBta-miR-193b</u>	473 (0.41%)	26	69	73	83	102	78	36	32
<u>φSsc-miR-30e-5p</u>	461 (0.40%)	14	27	57	43	134	93	56	51
Ssc-miR-100	405 (0.35%)	6	20	55	72	92	83	39	44
Ssc-miR-99b	350 (0.30%)	8	15	43	47	102	73	28	42

IB: Iberian breed, WB: European Wild Boar, LD: Landrace breed, LW: Large White breed, PT: Piétrain breed, ME: Meishan breed, VT: Vietnamese breed. miRNA name represents the most expressed miRNA sequence in the cluster.

Underlined miRNAs correspond to miRNAs that had not been previously described.

φ: miRNA clusters whose predominant isomiR is different from that described in miRBase database (v18, November 2011).

doi:10.1371/journal.pone.0055402.t003

### Validation of novel miRNAs through RT-qPCR

Unannotated sequences (486 unique sequences) that mapped in a unique region within the pig genome (Sscrofa 9.62) were clustered taking into account their position in the genome. After filtering by CN>2, 15 clusters remained for further analysis (Table S7). A compatible pre-miRNA folding structure could be successfully predicted *in silico* for 8 clusters (Cl-2, Cl-5, Cl-15, Cl-16, Cl-24, Cl-25, Cl-29 and Cl-38, Figure 2). Then, a validation protocol through RT-qPCR was designed for them. Five clusters (Cl-2, Cl-5, Cl-25, Cl-29 and Cl-38) were successfully amplified and, therefore, they have been considered as novel porcine miRNAs (Table 5). RT-qPCR efficiencies were high (ranging from 90% to 110%) and standard curves correlations were at least of 0.99. In addition, expression data showed that 4 clusters (Cl-5, Cl-25, Cl-29 and Cl-38) were differentially expressed ( $p$ -value<0.05) between breed groups (Table 6). Cl-5, Cl-29 and Cl-38 were down regulated in AS whereas Cl-25 was up regulated in EU.

To propose a nomenclature for the novel miRNAs, they were aligned to miRBase database accepting internal mismatches and some variation at the ends (Table 6). Cl-5 had a positive match with Hsa-miR-194-5p ( $p$ -value = 0.0004) with two nucleotide deletions at the ends, and also with Ssc-miR-194 ( $p$ -value = 0.002) although in this alignment one internal mismatch was detected in addition of the two nucleotide deletions at the ends. Cl-25 had homology with Hsa-miR-551a and Cl-29 did not match to any orthologous miRNA in miRBase database. Cl-38 matched with

Bta-miR-1468 ( $p$ -value = 0.0003) with one nucleotide deletion in the 5' end. Cl-2 was not differentially expressed through RT-qPCR but it matched with Hsa-miR-31-3p ( $p$ -value = 0.001) having two nucleotide deletion at the ends.

### Target prediction and functional analysis of differentially expressed porcine miRNAs

Target genes for the eight differentially expressed miRNAs (Hsa-miR-200b-3p, Hsa-miR-200c-3p, Ssc-miR-126, Ssc-miR-126\*, Ssc-miR-99a, Bta-miR-193b, Ssc-miR-486 and Ssc-let-7f) were predicted *in silico*. A total of 1,884 target genes were identified (Table S8) and they were functionally analyzed through KEGG pathways database where 96 biological pathways were determined (data not shown). Interestingly, Hsa-miR-200b-3p and Hsa-miR-200c-3p, from the same miRNA family, practically share the same pathways, related to cellular processes, signal transduction and some biological processes, like immune, nervous, circulatory and endocrine systems. They have also been related to cancer pathways, like renal cell carcinoma, among others. Ssc-miR-126, Ssc-miR-126\*, Ssc-miR-99a and Ssc-let-7f have also been described in some cancer pathways, emphasizing the importance of miRNAs in pathways related with cell cycle, growth and death. No significant related pathways were obtained from targets associated to Bta-miR-193b and Ssc-miR-486.

ID	Sequence	Counts
<b>Ssc-mir-23b</b>	<b>CTCTGGCTGCTTGGGTTCTCTGGCATGCTGATTTGTGACTTAAGATTAAATCACATTGCCAGGGATTACCACGCGAGCCAC</b>	
U0017522	-----ATCACATTGCCAGGGATTACCACGT-----	3
U0002949	-----ATCACATTGCCAGGGATTACCACG-----	7
U0001460	-----ATCACATTGCCAGGGATTACCACT-----	511
U0016038	-----GTATCACATTGCCAGGGATTACCACT-----	3
U0016637	-----ATCACATTGCCAGGGATTACCACCT-----	4
U0001746	-----ATCACATTGCCAGGGATTACCACA-----	28
U0010055	-----ATCACATTGCCAGGGATTACCACC-----	5
U0011971	-----ATCACATTGCCAGGGATTACCAGCT-----	3
U0000076	-----ATCACATTGCCAGGGATTACCAC-----	740
U0009402	-----TCACATTGCCAGGGATTACCACT-----	27
U0020652	-----TCACATTGCCAGGGATTACCACG-----	4
U0005514	-----TCACATTGCCAGGGATTACCACA-----	4
U0010984	-----ATCACATTGCCAGGGATTACTACT-----	3
U0004794	-----ATCACATTGCCAGGGATTACCAT-----	96
U0011426	-----ATCACATTGCCAGGGATTACCATT-----	6
U0014512	-----ATCACATTGCCAGGGATTACCACAT-----	3
U0010324	-----TCACATTGCCAGGGATTACCAC-----	21
U0008048	-----ATCACATTGCCAGGGATTACCACA-----	33
U0021197	-----ATCACATTGCCAGGGATTACCGA-----	5
U0022889	-----ATCACATTGCCAGGGATTACCAGC-----	3
U0000438	-----ATCACATTGCCAGGGATTACCAC-----	2,760*
U0009455	-----GTATCACATTGCCAGGGATTACCA-----	13
U0015675	-----TATCACATTGCCAGGGATTACCA-----	3
U0010120	-----ATCACATTGCCAGGGATTACCAG-----	6
U0020839	-----GTCACATTGCCAGGGATTACCA-----	3
U0006279	-----ATCACATTGCCAGGGATTACCT-----	15
U0016762	-----ATCACATTGCCAGGGATTACCG-----	10
U0006049	-----CACATTGCCAGGGATTACCAC-----	3
U0014676	-----CACATTGCCAGGGATTACCACT-----	3
U0000339	-----TCACATTGCCAGGGATTACCA-----	35
U0000653	-----ATCACATTGCCAGGGATTACC-----	385
U0007998	-----AATCACATTGCCAGGGATTACC-----	5
U0011315	-----ACCACATTGCCAGGGATTACCA-----	4
U0006315	-----ATCACATTGCCAGGGATTACCC-----	34
U0010641	-----ATCACATTGCCAGGGATTACAC-----	3
U0013832	-----ATCACATTGCCAGGGATTACAA-----	12
U0014190	-----ATCACATTGCCAGGGATTACTA-----	4
U0013088	-----ATCACATTGCCAGGGATTATCA-----	5
U0014265	-----TCACATTGCCAGGGATTACC-----	4
U0005203	-----CACATTGCCAGGGATTACCA-----	18
U0001080	-----ATCACATTGCCAGGGATTACT-----	9
U0003671	-----ATCACATTGCCAGGGATTACA-----	46
U0000974	-----ATCACATTGCCAGGGATTAC-----	344
U0011203	-----TCACATTGCCAGGGATTACA-----	4
U0000994	-----CACATTGCCAGGGATTACC-----	5
U0000989	-----ATCACATTGCCAGGGATTAAA-----	24
U0011606	-----ATCACATTGCCAGGGATTAAAA-----	5
U0012900	-----ATCACATTGCCAGGGATTAAAA-----	3
U0019069	-----TCACATTGCCAGGGATTAC-----	7
U0001098	-----ATCACATTGCCAGGGATTAA-----	60
U0007404	-----ATCACATTGCCAGGGATTAT-----	9
U0010188	-----ACATTGCCAGGGATTACCA-----	8
U0001343	-----ATCACATTGCCAGGGATTAA-----	34
U0000184	-----CACATTGCCAGGGATTAC-----	3
U0013118	-----ATCACATTGCCAGGGATAA-----	4
U0017452	-----ATCACATTGCCAGGGATAAA-----	3
U0011023	-----CATTGCCAGGGATTACCA-----	3
U0022053	-----TTGCCAGGGATTACCAC-----	3
U0018250	-----ATCGACATTGCCAGGGATTACCA-----	4

**Figure 1. Sequence variability in Ssc-miR-23b.** The number of total counts for each sequence is indicated. First sequence belongs to the precursor miRNA. Red bases match with precursor miRNA and belong to mature miRNA. Blue bases match also with precursor miRNA but do not belong to mature miRNA sequence. Sequence corresponding to the annotated miRNA in miRBase database is marked with an asterisk. doi:10.1371/journal.pone.0055402.g001

## Discussion

The present study represents the first work that describes the porcine kidney microRNAome in 6 different breeds and the European Ancestor, the Wild Boar, to determine differentially expressed miRNAs related to breed origin. A total of 110 out of 257 known porcine miRNAs in miRBase database have been determined in this study by high throughput sequencing. It is

known that some miRNAs could be tissue specific [64] and this could explain why this study has only described the 43,28% of the porcine miRNAs reported up to date which is in agreement with other studies [22,65]. Interestingly, among all described kidney miRNAs, 119 mammalian orthologous miRNAs were identified. Assuming high miRNA conservation in mammals [16,66], the

**Table 4.** Most expressed miRNAs diverging between the reference miRNA sequence and the miRBase<sup>1</sup> described miRNA sequence.

miRNA name	Total counts	IsomiRs	Most expressed isomiR sequence (5'-3')	Counts	Described miRNA sequence in miRBase database (5'-3')	Counts
Ssc-miR-23a	5,156	51	ATCACATTGCCAGGGATTCCA	3,116	ATCACATTGCCAGGGATTCCA	622
Ssc-miR-99a	3,781	32	AACCCGTAGATCCGATCTTGTGA	1,124	AACCCGTAGATCCGATCTTGTG	1,085
Ssc-miR-30d	1,977	31	TGTAACATCCCCGACTGGAAGC	711	TGTAACATCCCCGACTGGAAGCT	190
Ssc-miR-125a	1,369	29	TCCTGAGACCCTTTAACCTGT	479	TCCTGAGACCCTTTAACCTGTG	326
Bta-miR-193b	473	26	AACTGGCCCAAAAGTCCCGCT	196	AACTGGCCCAAAAGTCCCGCTTT	21
Ssc-miR-30e-5p	461	14	TGTAACATCCTTGACTGGAAGC	268	TGTAACATCCTTGACTGGAAGCT	74
Ssc-miR-139-5p	329	8	TCTACAGTGACAGTGTCTCCAGT	216	TCTACAGTGACAGTGTCTCCAG	85
Ssc-miR-30a-5p	314	10	TGTAACATCCTCGACTGGAAGC	198	TGTAACATCCTCGACTGGAAGCT	15
Ssc-miR-362	303	13	AATCCTTGGAACTAGGTGTGAGT	133	AATCCTTGGAACTAGGTGTGAGT	0
Hsa-miR-29c-5p	260	20	GACCGATTCTCTGGTGTTC	37	TGACCGATTCTCTGGTGTTC	22
Ssc-miR-374a	253	6	TTATAACAACCTGATAAGTGT	117	TTATAACAACCTGATAAGTG	13
Ssc-miR-145	222	12	GTCCAGTTTTCCAGGAATCCCT	91	GTCCAGTTTTCCAGGAATCCCTT	3
Hsa-miR-324-3p	211	22	ACTGCCCCAGGTGCTGCTGGT	45	ACTGCCCCAGGTGCTGCTGG	0
Ssc-miR-21	210	4	TAGCTTATCAGACTGATGTTGAC	105	TAGCTTATCAGACTGATGTTGA	77
Hsa-miR-874	195	8	CTGCCCTGGCCCGAGGGACCGAC	56	CTGCCCTGGCCCGAGGGACCGA	42
Ssc-miR-218b	185	7	TTGTGCTTGATCTAACCATGT	56	TTGTGCTTGATCTAACCATGTG	27
Ssc-miR-191	183	8	CAACGGAATCCCAAAGCAGCT	63	CAACGGAATCCCAAAGCAGCTG	45
Hsa-miR-150-5p	174	12	GTTCTCCCAACCTTGTACCAGT	47	TCTCCCAACCTTGTACCAGTG	14
Ssc-miR-22-3p	106	5	AAGCTGCCAGTTGAAGAAC	43	AAGCTGCCAGTTGAAGAAGTGT	28

<sup>1</sup>: miRBase database (v18, November 2011).

Bta: *Bos taurus*, Hsa: *Homo sapiens*, Ssc: *Sus scrofa*.

Marked in bold the nucleotide variation between pair sequences.

doi:10.1371/journal.pone.0055402.t004

high proportion of orthologous miRNAs (52%) suggests that there are still many miRNAs not described in pigs yet.

miRNA expression profile in kidney revealed that the most expressed miRNAs were Hsa-miR-200b-3p, Ssc-miR-125b and Ssc-miR-23b. Interestingly, Hsa-miR-200b-3p is the most expressed miRNA in all breeds except for Iberian breed, which Ssc-miR-125b (567 reads), Ssc-miR-99a (339 reads) and Hsa-miR-200b-3p (239 reads) are the first, the second and the third most expressed miRNAs, respectively. This variation of the expression in Iberian breed could be explained by many reasons. Firstly, some bias could be assumed because of the low number of reads obtained in this library in comparison with the other libraries (Table 2). However, the total reads of each miRNA in Iberian breed evidence that there is some tendency in its expression. Secondly, this change in the most expressed miRNA could also be explained as Iberian breed was exclusively originated from European Wild Boar and not mixed with any other ancestor neither being in contact with other commercial breeds in Europe.

Importantly, the most expressed miRNA in kidney was Hsa-miR-200b-3p which has not already been described in pig. Its sequence could not be found in the previous pig genome sequence (v9.2) but it could be now annotated in the last version (Sscrofa10, ftp://ftp.ensembl.org/pub/release-67/fasta/sus\_scrofa/dna/) probably due to it was located on a gap in the pig genome sequence that has been solved in the new version. Its expression was significantly higher than other profiled miRNAs, as it had been detected 27,097 times corresponding to the 23.50% of the total reads. This expression pattern in miRNAs profile has been defined in more studies [67], where a miRNA is predominant over the rest meaning that it could have a significant role in the

expressed tissue. It is hypothesised that it could governs or be implicated on the major constitutive functions carried out by this tissue.

The multiple sequences obtained for each specific miRNA were aligned to study the variability in the 3' and 5' miRNA ends. Most of the identified miRNAs (70%) showed nucleotide polymorphisms and variations in sequence length, according to previous studies [31,68]. One example in this study is Ssc-miR-23b, where the length varies from 18 to 25 nucleotides (Figure 1). This variability in the 5' and 3' ends could be favoured by processes prior to the high throughput sequencing sample preparation steps [69], being explained by the cleavage of the same miRNA precursor at different nucleotides during Drosha and Dicer processing. Another reason of this variability could be the nucleotide replacement and additions or deletions in the 5' and 3' miRNA ends. All this post-transcriptional mechanisms, called RNA editing, generate the isomiRs, defined as sequence variations in mature miRNAs and commonly reported in miRNA studies [10,70]. Mechanisms of miRNA sequence diversification are extensively reviewed by Berezikov [16], suggesting a functional meaning to all these processes. Furthermore, different patterns described on isomiR expression suggest different strategies in the regulation of the expression. Assuming the expression data obtained in this study, there are miRNAs where the main isomiR could represent the major regulation function and there are miRNAs which their regulation function could be assumed by more than one isomiR (Table S3).

Considering the proportion of unique sequences obtained per breed in reference to the total sequences, Iberian and Vietnamese breeds obtained the highest proportion (38% and 39%, re-

**Table 5.** Selected miRNA and putative novel miRNAs (Clusters) for further validation through RT-qPCR.

miRNA/Cluster	Total counts	IsomiRs	Sequence (5'-3')
Hsa-miR-200b-3p	27,097	123	TAATACTGCCTGGAATGATGA
Ssc-miR-126	5,274	49	TCGTACCGTGAGTAATAATGCG
Ssc-miR-99a	3,781	32	AACCCGTAGATCCGATCTTGTGA
Hsa-miR-200c-3p	3,478	32	TAATACTGCCGGTAATGATGGA
Ssc-miR-126*	2,796	26	CATTATTACTTTTGGTACGCG
Ssc-miR-92a	797	17	TATTGCACTTGCCCGCCTGT
Ssc-miR-26a	598	17	TTCAAGTAATCCAGGATAGGCT
Bta-miR-193b	493	26	AACTGGCCCAAAAGTCCCGCT
Ssc-miR-532-5p	113	3	CATGCCTTGAGTGTAGGCCGT
Ssc-miR-423-5p	74	10	TGAGGGCAGAGAGCGAGACTTT
Ssc-miR-29c	23	3	TAGCACCATTGAAATCGGTTA
Ssc-miR-486	23	4	TCCTGTACTGAGCTGCCCGAG
Ssc-let-7f	17	1	TGAGGTAGTAGATTGTATAGTT
Cl-2 (Hsa-miR-31-3p) <sup>1</sup>	4	1	CTGCTATGCCAACATATTGCCA
Cl-5 (Hsa-miR-194-5p) <sup>1</sup>	12	2	CTGTAACAGCAACTCCATGTGGAA
Cl-15 (Hsa-miR-551a) <sup>1</sup>	9	2	GCGACCCACTCTTGTTTCCATG
Cl-16	6	1	TTGGTGACCAGGTGCTCAGGGAG
Cl-24	3	2	CTGCATTTCTGGTGCCCTTAATT
Cl-25 (Hsa-miR-138-5p) <sup>1</sup>	7	3	CAGCTGGTGTGTGAATCAGGCCG
Cl-29	3	1	GTTGGTGTACTGGAATAGCT
Cl-38 (Bta-miR-1468) <sup>1</sup>	21	3	TCTCCGTTGCTGTTTGTCTGA

<sup>1</sup>: miRBase (v18) homology (e-value  $\leq 1e-03$ ). Internal mismatches were accepted.

Bta: *Bos taurus*, Hsa: *Homo sapiens*, Ssc: *Sus scrofa*.

doi:10.1371/journal.pone.0055402.t005

spectively) showing that, if the number of miRNAs was not increased in these libraries, what was increased was the number of described isomiRs. The percentage of unique sequences in the other libraries and also in the whole data (all libraries) was between 15 and 19%. Thus, the variability in miRNA sequences seemed to be higher in European breeds like Iberian breed and in Asian breeds such as Vietnamese breed. In contrast, when European commercial breeds were created through the European and Asian breeds introgression, their variability could be reduced due to their strong selection in production traits, as it is indicated if we consider the number of isomiRs in each European commercial breed studied.

Looking at the most expressed isomiR by miRNA in each breed we can determine that largely it was the same isomiR in all studied breeds. However, in 76 miRNAs, like Ssc-miR-99a, the most expressed isomiR varies depending on the breed, being more expressed one isomiR (U0009355: 5'-AACCCGTAGATCCGATCTTGTG-3') in European and Asian breeds, and other isomiR (U0009353: 5'-AACCCGTAGATCCGATCTTGTGA-3') in European commercial breeds. In other cases, there is an agreement in all breeds for the most expressed isomiR except in one breed, like in Ssc-miR-21 and Ssc-miR-100 in Iberian breed, Ssc-miR-152 in Piétrain breed, Hsa-miR-200c-3p in Vietnamese breed or Hsa-let-7d-5p in Landrace breed. There are also few cases where the most expressed isomiR is different in almost each breed, such as in Ssc-miR-199a\* and Ssc-miR-423-5p. All these different situations suggest that the genetic variability degree between breeds is high and it could play an important role in the post-transcriptional regulation mechanisms leading to perform different functions.

Another variation factor was that there were some differences between the most expressed isomiR in each miRNA and its concordance with the described sequence in miRBase database, diverging them in 72 miRNAs found in this study. As example, the most representative sequence for Ssc-miR-23a in this study was 5'-ATCACATTGCCAGGGATTTCCA-3', found 3,116 times and it corresponded to Bta-miR-23a. The sequence described in miRBase database for Ssc-miR-23a (5'-ATCACATTGCCAGGATTTCC-3') was found only 622 times (Table 4). It also suggests that there is not a fixed predominant isomiR, but there is variability in isomiR expression given not only by many factors like age, tissue or disease [71], but also by other factors like breed [50].

Despite of all miRNA homologies found, this study was also focused on describing novel miRNAs. The strict methodology used proposed 8 novel miRNAs to be validated through RT-qPCR and finally 5 miRNAs were successfully validated. These confirmed novel porcine miRNAs were mapped once in the pig genome (Table S7) and their pre-miRNA folding were successfully predicted (Figure 2). Moreover, three out of the thirteen selected miRNAs from the kidney miRNAome to be amplified through RT-qPCR were orthologous (Hsa-miR-200b-3p, Hsa-miR-200c-3p and Bta-miR-193b, Table 5) and, therefore, they were also confirmed as new porcine miRNAs. Significant differential expression regarding breed groups was obtained by RT-qPCR in eight miRNAs: Hsa-miR-200b-3p, Hsa-miR-200c-3p, Ssc-miR-126, Ssc-miR-126\*, Ssc-miR-99a, Bta-miR-193b, Ssc-miR-486 and Ssc-let-7f. However, expression results from RT-qPCR were not in agreement with the differential expression study from high throughput sequencing data in some cases. These differences could be explained by many factors, like some possible bias, the fact that

**Table 6.** Fold Change (FC) comparison from HTS differential expression study and RT-qPCR data.

miRNA/Cluster	HTS				RT-qPCR			
	EU <sup>2</sup> vs. EA <sup>3</sup>	EU vs. AS <sup>4</sup>	EA vs. AS	ANOVA group factor <i>p</i> -value	EU vs. EA	EU vs. AS	EA vs. AS	ANOVA group factor <i>p</i> -value
Hsa-miR-200b-3p	-1.45	-1.54	-1.06	0.794	1.29	1.50*	1.16	0.012
Ssc-miR-126	-1.07	1.24	1.32	0.576	1.09	1.60***	1.47***	<.001
Ssc-miR-99a	1.63	3.05	1.87	0.237	1.10	1.29**	1.18 <sup>§</sup>	0.005
Hsa-miR-200c-3p	-2.31	-1.73	1.34	0.266	1.39	-1.15	-1.61**	0.009
Ssc-miR-126*	-1.36	-1.18	1.15	0.352	1.08	1.35 <sup>§</sup>	1.25	0.044
Ssc-miR-92a	-1.87	-1.68	1.11	0.168	1.20	1.02	-1.18	0.614
Ssc-miR-26a	-2.15	-1.29	1.67	0.235	1.13	1.14	1.01	0.602
Bta-miR-193b	2.23	5.46	2.45	0.285	1.25	1.55**	1.25	0.003
Ssc-miR-532-5p	-1.16	-2.49	-2.14	0.082	1.05	-1.06	-1.12	0.685
Ssc-miR-423-5p	5.28	3.76	-1.40	0.110	1.09	1.14	1.05	0.416
Ssc-miR-29c	-3.78	1.45	5.47	0.028	1.01	1.20	1.19	0.105
Ssc-miR-486	-5.66	-1.55	3.64	0.179	-1.48*	-1.28	1.16	0.025
Ssc-let-7f	-2.76	-11.60	-4.21	0.344	1.22	1.33*	1.09	0.017
Cl-2 (Hsa-miR-31-3p) <sup>1</sup>	-	-	-	-	1.05	1.12	1.07	0.631
Cl-5 (Hsa-miR-194-5p) <sup>1</sup>	-	-	-	-	1.02	1.24	1.22 <sup>§</sup>	0.049
Cl-25 (Hsa-miR-138-5p) <sup>1</sup>	-	-	-	-	1.35**	1.36**	1.00	0.003
Cl-29	-	-	-	-	-1.03	1.21	1.24*	0.009
Cl-38 (Bta-miR-1468) <sup>1</sup>	-	-	-	-	1.03	1.60***	1.55***	<.001

<sup>1</sup>: miRBase (v18) homology (e-value $\leq$ 1e-03). Internal mismatches were accepted.

<sup>2</sup>: EU: European breeds. <sup>3</sup>: EA: European commercial breeds. <sup>4</sup>: AS: Asian breeds.

For HTS data, fold changes from sequence counting between breed groups were calculated from normalised data in counts per thousand for each library and averaged per groups. Positive and negative signs indicate that the level of gene expression is higher for the first or the second group of the test, respectively. Analysis of variance (ANOVA) including breed as fixed factor was performed. Significance was set at  $P < 0.05$ . Fold changes for clusters could not be calculated due to their low total counts. For qPCR data expression study, the quantity obtained of each miRNA in each sample was normalised by the Normalization Factor and corrected in relation to the lowest normalised value. Analysis of variance (ANOVA) including the RT and breed as fixed factors was performed and fold change from Least Squares Means (LSM) between breed groups were calculated. Significance was set at  $P < 0.05$ . Scheffe test determined whether there was significant differential expression between breed groups (<sup>§</sup>: suggestive  $p$ -value $<0.1$ , \* : significant with  $p$ -value $<0.05$ , \*\* : significant with  $p$ -value $<0.01$ , \*\*\* : significant with  $p$ -value $<0.001$ ).

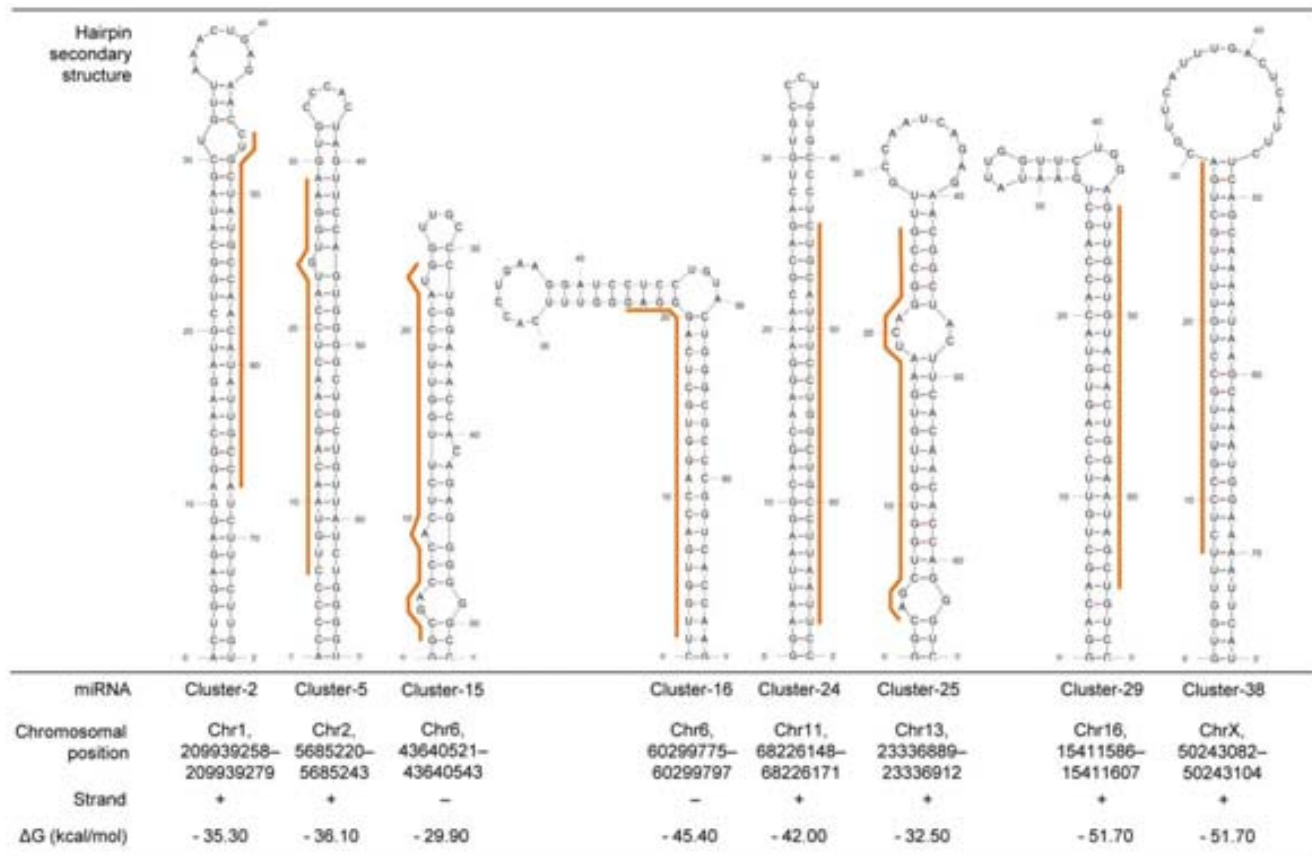
doi:10.1371/journal.pone.0055402.t006

not each miRNA is amplified in the same proportion with high throughput sequencing [72,73] and the presence of isomiRs, which make difficult to perform an accurate filtering of sequences. Other factors to be considered are some diversity and complexity in the miRNAs nomenclature between species listed in the miRNAs databases [15], which make difficult to perform a proper sequence classification. In this sense, cluster analysis methodology for novel miRNA detection was chosen to avoid problems with miRNA nomenclature. It is also important to consider the specificity in qPCR amplification, where primers designed in each miRNA were exclusive for the most expressed isomiR in each miRNA, amplifying only that one. Overall, the relevant topic of next generation sequencing about if it can be treated as expression data taking into account its large amount of reads, is still discussed [74]. The slightly different expression results between high throughput sequencing and RT-qPCR data reveal a change of expression tendency in some miRNAs evidencing the “semi-quantitative” nature of high throughput sequencing methods. Low correlations (from 0.52 to 0.003) were determined between high throughput sequencing by 454 (Roche, Germany) and RT-qPCR data. In this sense, the bias in this study could be explained by the lower number of sequences obtained. For instance, Hsa-miR-200b-3p appears in a higher expression in Asian breeds in sequencing data while it is more expressed in European breeds according to real time RT-qPCR data. Furthermore, Ssc-miR-126\* was more represented in European commercial breeds,

whereas in RT-qPCR it was up regulated in European breeds. The same case happened with Ssc-miR-126, a conclusive fact taking into account that both miRNAs are transcribed together. There are similar cases with low expressed miRNAs (according to sequence count) such as Ssc-let-7f, which was more expressed in Asian breeds but appears up regulated in European breeds in real time RT-qPCR data.

Analysing the target pathways of differentially expressed miRNAs between breed groups, it was found that Hsa-mir-200b-3p, the most expressed miRNA in this study and up regulated in European breeds, was related in several kidney diseases like tubulointerstitial fibrosis [75] or hypertensive glomerulosclerosis [12], elucidating the importance of this miRNA in kidney physiology. Targets of Hsa-miR-200b-3p were involved in several pathways like renal cell carcinoma, associated to some oncogenes such as MET, or tumors suppressors like VHL, FH and BHD. Interestingly, miRNA targets were related with the vascular smooth muscle contraction and calcium signaling pathways. Both pathways are important in muscular growth and development in which Hsa-miR-200b-3p, being more expressed in European pig breeds, could contribute to a greater efficiency. Furthermore, European commercial breeds are precisely selected for production traits and strongly characterized for animal production. Kidney and muscle are strongly related because kidney is in charge of secreting erythropoietin, the hormone responsible to activate erythropoiesis and also to transport oxygen to the muscles and





**Figure 2. Predicted miRNA folding for cluster candidates to novel miRNAs using MFold software<sup>1</sup>.** <sup>1</sup>: [54]. Orange line points out the miRNA position into the pre-miRNA structure. doi:10.1371/journal.pone.0055402.g002

other tissues, and thus, favoring muscular growth. Hsa-miR-200c-3p, which belongs to the same family that Hsa-miR-200b-3p, is also differentially expressed being up regulated in Asian breeds. Some studies have related it to some mechanisms of cancer development in many organs, such as pancreas, bladder, ovaries, prostate or breast. [76–80]. Functional analysis related targets of Hsa-miR-200c-3p with neoplasias, specifically in renal cell carcinoma. Furthermore, Hsa-miR-200c-3p is also involved in some pathways related to reproduction, like oocyte meiosis or progesterone-mediated oocyte maturation. Breeds with an Asian origin have better phenotypes for reproduction traits and, interestingly, Asian breeds have Hsa-miR-200c-3p up regulated suggesting its involvement in the regulation of the reproduction traits. However, Ssc-miR-99a and Hsa-miR-200b-3p targets are also related to these reproduction pathways, both up regulated miRNAs in European breeds which could play an antagonistic role in European breeds. Some targets of Ssc-miR-126, a down regulated miRNA in Asian breeds, participate in cellular processes like focal adhesion and regulation of actin cytoskeleton, while Ssc-miR-126\*, also down regulated in Asian breeds, regulate targets implicated with cellular processes like adherens, gap and tight junction. Moreover, both miRNAs also have a strong relation with many cancer pathways, which were related in all miRNAs. Ssc-let-7f is an ubiquitous miRNA from let-7 family, related to many pathways involved in wide range of physiological functions (cellular, metabolic and environmental information processes) and also to many diseases, like renal cell carcinoma [81,82].

This study opens a wide field about the miRNA specificity, which is not only by tissue, age or species level, but also breed factor is crucial to manage the phenotypic changes and, therefore, the genetic variability. The role of miRNAs in gene expression regulation is much more complex than it was thought initially, being necessary to study in depth which roles isomiRs play and to uncover all factors that make the post-transcriptional regulation through miRNAs so complex.

## Supporting Information

**Table S1 Adaptors used for the construction of each library.** <sup>1</sup>: Forward adaptor, <sup>2</sup>: Reverse adaptor. The 5 nt code used for each library is in bold in the adaptor sequence. Reverse adaptor was used in all breeds libraries. (DOC)

**Table S2 Complete miRNA profile in healthy kidney of 6 porcine breeds and the European ancestor from 454 GS FLX run data.** IB: Iberian breed, WB: European Wild Boar, LD: Landrace breed, LW: Large White breed, PT: Pietrain breed, ME: Meishan breed, VT: Vietnamese breed. miRNA name represents the most expressed sequence in the cluster. Bta: *Bos taurus*, Dre: *Danio rerio*, Eca: *Equus caballus*, Hsa: *Homo sapiens*, Mdo: *Monodelphis domestica*, Mmu: *Mus musculus*, Rno: *Ratus norvegicus*, Sha: *Sarcophilus harrisii*, Ssc: *Sus scrofa*. (DOC)

**Table S3 Summary of the isomiR distribution in the most expressed miRNAs (CN>1,000).** miRNA name represents the most expressed sequence in the cluster. Hsa: *Homo sapiens*, Ssc: *Sus scrofa*. (DOC)

**Table S4 Described miRNAs from HTS presenting variation between the reference miRNA sequence and the miRBase described miRNA sequence.** Bta: *Bos taurus*, Eca: *Equus caballus*, Hsa: *Homo sapiens*, Mdo: *Monodelphis domestica*, Mmu: *Mus musculus*, Rno: *Rattus norvegicus*, Sha: *Sarcophilus harrisii*, Ssc: *Sus scrofa*. Marked in bold the nucleotide variation between pair sequences. (DOC)

**Table S5 List of differentially expressed miRNAs (Fold Change >1.5 times) between porcine breed groups.** EU: European breeds; EA: European commercial breeds; AS: Asian breeds. Bta: *Bos taurus*, Dre: *Danio rerio*, Eca: *Equus caballus*, Hsa: *Homo sapiens*, Mdo: *Monodelphis domestica*, Mmu: *Mus musculus*, Rno: *Rattus norvegicus*, Sha: *Sarcophilus harrisii*, Ssc: *Sus scrofa*. Positive and negative signs indicate that the level of gene expression is higher for the first or the second group of the test, respectively. (DOC)

**Table S6 Normalised 454 GS FLX run data by library in counts per thousand.** IB: Iberian breed, WB: European Wild

## References

- Ambros V (2008) The evolution of our thinking about microRNAs. *Nature Medicine* 14: 1036–1040.
- Ruvkun G (2008) The perfect storm of tiny RNAs. *Nature Medicine* 14: 1041–1045.
- Bartel DP (2009) MicroRNAs: Target recognition and regulatory functions. *Cell* 136: 215–233.
- Krol J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews. Genetics* 11: 597–610.
- Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116: 281–297.
- Guo H, Ingolia NT, Weissman JS, Bartel DP (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466: 835–840.
- Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nature Reviews. Genetics* 12: 99–110.
- Ambros V (2004) The functions of animal microRNAs. *Nature Medicine* 431: 350–355.
- Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120: 15–20.
- Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, et al. (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Research* 18: 610–621.
- Saal S, Harvey SJ (2009) MicroRNAs and the kidney: Coming of age. *Current Opinion in Nephrology and Hypertension* 18: 317–323.
- Li JY, Yong TY, Michael MZ, Gleadle JM (2010) Review: The role of microRNAs in kidney disease. *Nephrology (Carlton, Vic.)* 15: 599–608.
- Kozomara A, Griffiths-Jones S (2011) miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research* 39: D152–7.
- Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: Tools for microRNA genomics. *Nucleic Acids Research* 36: D154–8.
- Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Research* 34: D140–4.
- Berezikov E (2011) Evolution of microRNA diversity and regulation in animals. *Nature Reviews. Genetics* 12: 846–860.
- Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP (2003) Vertebrate microRNA genes. *Science (New York, N.Y.)* 299: 1540.
- Tarver JE, Donoghue PC, Peterson KJ (2012) Do miRNAs have a deep evolutionary history? *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 34: 857–866.
- Lunney JK (2007) Advances in swine biomedical model genomics. *International Journal of Biological Sciences* 3: 179–184.
- Glazov EA, Cottee PA, Barris WC, Moore RJ, Dalrymple BP, et al. (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Research* 18: 957–964.
- Huang TH, Zhu MJ, Li XY, Zhao SH (2008) Discovery of porcine microRNAs and profiling from skeletal muscle tissues during development. *PLoS One* 3: e3225.
- Ribeiro-dos-Santos A, Khayat AS, Silva A, Alencar DO, Lobato J, et al. (2010) Ultra-deep sequencing reveals the microRNA expression pattern of the human stomach. *PLoS One* 5: e13205.
- Shao NY, Hu HY, Yan Z, Xu Y, Hu H, et al. (2010) Comprehensive survey of human brain microRNA by deep sequencing. *BMC Genomics* 11: 409.
- Xie SS, Li XY, Liu T, Cao JH, Zhong Q, et al. (2011) Discovery of porcine microRNAs in multiple tissues by a solexa deep sequencing approach. *PLoS One* 6: e16235.
- Chen X, Ba Y, Ma L, Cai X, Yin Y, et al. (2008) Characterization of microRNAs in serum: A novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Research* 18: 997–1006.
- Chow TF, Youssef YM, Lianidou E, Romaschin AD, Honey RJ, et al. (2010) Differential expression profiling of microRNAs and their potential involvement in renal cell carcinoma pathogenesis. *Clinical Biochemistry* 43: 150–158.
- Godwin JG, Ge X, Stephan K, Jurisch A, Tullius SG, et al. (2010) Identification of a microRNA signature of renal ischemia reperfusion injury. *Proceedings of the National Academy of Sciences of the United States of America* 107: 14339–14344.
- Patel N, Sauter ER (2011) Body fluid micro(mi)RNAs as biomarkers for human cancer. *Journal of Nucleic Acids Investigation* 2: e1.
- Fendler A, Stephan C, Yousef GM, Jung K (2011) MicroRNAs as regulators of signal transduction in urological tumors. *Clinical Chemistry* 57: 954–968.
- Catto JW, Alcaraz A, Bjartell AS, De Vere White R, Evans CP, et al. (2011) MicroRNA in prostate, bladder, and kidney cancer: A systematic review. *European Urology* 59: 671–681.
- Nielsen M, Hansen JH, Hedegaard J, Nielsen RO, Panitz F, et al. (2010) MicroRNA identity and abundance in porcine skeletal muscles determined by deep sequencing. *Animal Genetics* 41: 159–168.
- Reddy AM, Zheng Y, Jagadeeswaran G, Macmill SL, Graham WB, et al. (2009) Cloning, characterization and expression analysis of porcine microRNAs. *BMC Genomics* 10: 65.
- Cho IS, Kim J, Seo HY, Lim do H, Hong JS, et al. (2010) Cloning and characterization of microRNAs from porcine skeletal muscle and adipose tissue. *Molecular Biology Reports* 37: 3567–3574.
- Sharbati S, Friedlander MR, Sharbati J, Hoeke L, Chen W, et al. (2010) Deciphering the porcine intestinal microRNA transcriptome. *BMC Genomics* 11: 275.
- Lian C, Sun B, Niu S, Yang R, Liu B, et al. (2012) A comparative profile of the microRNA transcriptome in immature and mature porcine testes using solexa deep sequencing. *The FEBS Journal*.
- Wu YQ, Chen DJ, He HB, Chen DS, Chen LL, et al. (2012) Pseudorabies virus infected porcine epithelial cell line generates a diverse set of host microRNAs and a special cluster of viral microRNAs. *PLoS One* 7: e30988.

37. Anselmo A, Flori L, Jaffrezic F, Rutigliano T, Cecere M, et al. (2011) Co-expression of host and viral microRNAs in porcine dendritic cells infected by the pseudorabies virus. *PLoS One* 6: e17374.
38. Li S, Ran XQ, Xu L, Wang JF (2011) microRNA and mRNA expression profiling analysis of dichlorvos cytotoxicity in porcine kidney epithelial PK15 cells. *DNA and Cell Biology* 30: 1073–1083.
39. Larson G, Dobney K, Albarella U, Fang M, Matisoo-Smith E, et al. (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science (New York, N.Y.)* 307: 1618–1621.
40. Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, et al. (2007) Ancient DNA, pig domestication, and the spread of the neolithic into Europe. *Proceedings of the National Academy of Sciences of the United States of America* 104: 15276–15281.
41. Larson G, Liu R, Zhao X, Yuan J, Fuller D, et al. (2010) Patterns of east Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America* 107: 7686–7691.
42. Amaral AJ, Ferretti L, Megens HJ, Crooijmans RP, Nie H, et al. (2011) Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. *PLoS One* 6: e14782.
43. Ojeda A, Ramos-Onsins SE, Marletta D, Huang LS, Folch JM, et al. (2011) Evolutionary study of a potential selection target region in the pig. *Heredity* 106: 330–338.
44. Giuffra E, Kijas JM, Amarger V, Carlborg O, Jeon JT, et al. (2000) The origin of the domestic pig: Independent domestication and subsequent introgression. *Genetics* 154: 1785–1791.
45. Gao Y, Zhang YH, Jiang H, Xiao SQ, Wang S, et al. (2011) Detection of differentially expressed genes in the longissimus dorsi of northeastern indigenous and large white pigs. *Genetics and Molecular Research: GMR* 10: 779–791.
46. Fang M, Andersson L (2006) Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication. *Proceedings Biological Sciences/the Royal Society* 273: 1803–1810.
47. Clop A, Amills M, Noguera JL, Fernandez A, Capote J, et al. (2004) Estimating the frequency of Asian cytochrome B haplotypes in standard European and local Spanish pig breeds. *Genetics, Selection, Evolution: GSE* 36: 97–104.
48. Alves E, Ovilo C, Rodriguez MC, Sillio L (2003) Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. *Animal Genetics* 34: 319–324.
49. Kim KI, Lee JH, Li K, Zhang YP, Lee SS, et al. (2002) Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. *Animal Genetics* 33: 19–25.
50. Timoneda O, Balcells I, Cordoba S, Castello A, Sanchez A (2012) Determination of reference microRNAs for relative quantification in porcine tissues. *PLoS One* 7: e44413.
51. Peltier HJ, Latham GJ (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: Identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA (New York, N.Y.)* 14: 844–852.
52. Balcells I, Cirera S, Busk PK (2011) Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC Biotechnology* 11: 70.
53. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3: RESEARCH0034.
54. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* 31: 3406–3415.
55. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, et al. (2003) A uniform system for microRNA annotation. *RNA (New York, N.Y.)* 9: 277–279.
56. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, et al. (2009) Accurate microRNA target prediction correlates with protein repression levels. *BMC Bioinformatics* 10: 295.
57. Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, et al. (2009) DIANA-microT web server: Elucidating microRNA functions through target prediction. *Nucleic Acids Research* 37: W273–6.
58. Zhang B, Kirov S, Snoddy J (2005) WebGestalt: An integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Research* 33: W741–8.
59. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: Tool for the unification of biology. the gene ontology consortium. *Nature Genetics* 25: 25–29.
60. Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28: 27–30.
61. Ogata H, Goto S, Sato K, Fugibuchi W, Bono H, et al. (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 27: 29–34.
62. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate – a practical and powerful approach to multiple testing. *Biometrika* 83: 301–313.
63. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* 55: 611–622.
64. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, et al. (2002) Identification of tissue-specific microRNAs from mouse. *Current Biology: CB* 12: 735–739.
65. Li HY, Xi QY, Xiong YY, Liu XL, Cheng X, et al. (2012) Identification and comparison of microRNAs from skeletal muscle and adipose tissues from two porcine breeds. *Animal Genetics*.
66. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science (New York, N.Y.)* 294: 858–862.
67. Zhou Y, Tang X, Song Q, Ji Y, Wang H, et al. (2012) Identification and characterization of pig embryo microRNAs by Solexa sequencing. *Reproduction in Domestic Animals = Zuchthygiene*.
68. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, et al. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129: 1401–1414.
69. Lee LW, Zhang S, Etheridge A, Ma L, Martin D, et al. (2010) Complexity of the microRNA repertoire revealed by next-generation sequencing. *RNA (New York, N.Y.)* 16: 2170–2180.
70. Fernandez-Valverde SL, Taft RJ, Mattick JS (2010) Dynamic isomiR regulation in *Drosophila* development. *RNA (New York, N.Y.)* 16: 1881–1888.
71. Li SC, Liao YL, Ho MR, Tsai KW, Lai CH, et al. (2012) miRNA arm selection and isomiR distribution in gastric cancer. *BMC Genomics* 13 Suppl 1: S13.
72. Dohm JC, Lottaz C, Borodina T, Himmelbauer H (2008) Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research* 36: e105.
73. Taub MA, Corrada Bravo H, Irizarry RA (2010) Overcoming bias and systematic errors in next generation sequencing data. *Genome Medicine* 2: 87.
74. Malone JH, Oliver B (2011) Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biology* 9: 34.
75. Oba S, Kumano S, Suzuki E, Nishimatsu H, Takahashi M, et al. (2010) miR-200b precursor can ameliorate renal tubulointerstitial fibrosis. *PLoS One* 5: e13614.
76. Li Y, VandenBoom TG, 2nd, Kong D, Wang Z, Ali S, et al. (2009) Up-regulation of miR-200 and let-7 by natural agents leads to the reversal of epithelial-to-mesenchymal transition in gemcitabine-resistant pancreatic cancer cells. *Cancer Research* 69: 6704–6712.
77. Han Y, Chen J, Zhao X, Liang C, Wang Y, et al. (2011) MicroRNA expression signatures of bladder cancer revealed by deep sequencing. *PLoS One* 6: e18286.
78. Marchini S, Cavalieri D, Fruscio R, Calura E, Garavaglia D, et al. (2011) Association between miR-200c and the survival of patients with stage I epithelial ovarian cancer: A retrospective study of two independent tumour tissue collections. *The Lancet Oncology* 12: 273–285.
79. Tang X, Tang X, Gal J, Kyprianou N, Zhu H, et al. (2011) Detection of microRNAs in prostate cancer cells by microRNA array. *Methods in Molecular Biology (Clifton, N.J.)* 732: 69–88.
80. Bockmeyer CL, Christgen M, Muller M, Fischer S, Ahrens P, et al. (2011) MicroRNA profiles of healthy basal and luminal mammary epithelial cells are distinct and reflected in different breast cancer subtypes. *Breast Cancer Research and Treatment*.
81. Gottardo F, Liu CG, Ferracin M, Calin GA, Fassin M, et al. (2007) Micro-RNA profiling in kidney and bladder cancers. *Urologic Oncology* 25: 387–392.
82. White NM, Bui A, Mejia-Guerrero S, Chao J, Soosaipillai A, et al. (2010) Dysregulation of kallikrein-related peptidases in renal cell carcinoma: Potential targets of miRNAs. *Biological Chemistry* 391: 411–423.



**3.2. Estudi de les funcions dels microRNAs virals i de l'hoste durant el procés d'infecció amb el virus de la malaltia d'Aujeszky**

**The Role of Viral and Host microRNAs in the Aujeszky's Disease Virus  
During the Infection Process**

Timoneda O, Núñez F, Balcells I, Muñoz M, Castelló A, Vera G, Pérez LJ, Egea R, Mir G,  
Córdoba S, Rosell R, Segalés J, Tomàs A, Sánchez A, Núñez JI  
(manuscrit en preparació per a ser presentat a PLoS ONE)



## The role of viral and host microRNAs in the Aujeszky's disease virus during the infection process

Oriol Timoneda<sup>1,2,\*</sup>, Fernando Núñez<sup>3</sup>, Ingrid Balcells<sup>1,2</sup>, Marta Muñoz<sup>3</sup>, Anna Castelló<sup>1,2</sup>, Gonzalo Vera<sup>2</sup>, Lester J. Pérez<sup>4</sup>, Raquel Egea<sup>1</sup>, Gisela Mir<sup>2</sup>, Sarai Córdoba<sup>2</sup>, Rosa Rosell<sup>3,5</sup>, Joaquim Segalés<sup>3,6</sup>, Anna Tomàs<sup>7</sup>, Armand Sánchez<sup>1,2</sup>, José I. Núñez<sup>3</sup>.

<sup>1</sup> *Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain.*

<sup>2</sup> *Departament de Genètica Animal, Centre de Recerca en AgriGenòmica (CRAG), CSIC-IRTA-UAB-UB, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.*

<sup>3</sup> *Centre de Recerca en Sanitat Animal (CRESA), UAB-IRTA, Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.*

<sup>4</sup> *Centro Nacional de Sanidad Agropecuaria (CENSA), La Habana, Cuba.*

<sup>5</sup> *Departament d'Agricultura, Ramaderia, Pesca, Alimentació i Medi Natural de la Generalitat de Catalunya (DAR), Spain.*

<sup>6</sup> *Departament de Sanitat i Anatomia Animals, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.*

<sup>7</sup> *Program Infection and Immunity, FISIB, 07110 Bunyola, Spain.*

\* Corresponding author

## ABSTRACT

Porcine production is a primary market in the world economy. Controlling swine diseases in the farm is essential in order to achieve the sector necessities. Aujeszky disease is a viral condition affecting pigs and is endemic in many countries of the world, causing important economic losses in the swine industry. microRNAs (miRNAs) are non-coding RNAs which modulates gene expression in animals, plants and viruses. With the aim of understanding miRNA roles during the Aujeszky's disease virus [ADV] (also known as suid herpesvirus type 1 [SHV-1]) infection, the expression profiles of host and viral miRNAs were determined through deep sequencing in SHV-1 infected porcine cell line (PK-15) and in an animal experimental SHV-1 infection with virulent (NIA-3) and attenuated (Begonia) strains. In the *in vivo* approach, let-7a and miR-378 presented an over expression in NIA-3 infected samples, whereas miR-206 and miR-133a presented differential expression between virus strains infection. In the *in vitro* approach, most miRNAs were down-regulated in infected groups. miR-92a and miR-92b-3p were up-regulated in Begonia infected samples. Functional analysis of all this over expressed miRNAs during the infection revealed their association in pathways related to viral infection processes and immune response. Furthermore, 8 viral miRNAs were confirmed by stem loop RT-qPCR in both *in vitro* and *in vivo* approaches, presenting a gene regulatory network affecting 59 viral genes. Most described viral miRNAs were related to Large Latency Transcript (LLT) and to viral transcription activators *EPO* and *IE180*, and also to regulatory genes regarding their important roles in the host-pathogen interaction during viral infection.



## INTRODUCTION

Domestic pig (*Sus scrofa domesticus*) can be found worldwide because of their extraordinary importance in the production of red meat, lard and cured products [1]. Pork represents the 40% of the world meat production, with more than 110M tones produced every year and with an increasing trend [2]. Genetic improvement of the pig has evolved over the years. Selection for productive characters such as backfat and growth rate was the main objective during the 70s and 80s. In the last decade of the 20<sup>th</sup> century, the interest was focused on reproductive traits, relating good fecundation with good productive life of sows and improving the litter size in dam lines [3]. More recently, product quality and new non-economic factors have been imposed under society pressure and demand, like animal health and welfare or ecologic consequences resulting from intensive porcine production [4]. In this sense, safeguard farm sanitary conditions and animal maintenance are of great interest to provide a better quality of animal life and also a better pork quality and farm performance.

Aujeszky's disease (AD), also known as pseudorabies, is a disease of great economic importance, especially for intensive production systems that concentrate large farms located relatively near from each other. The etiological agent of the disease is the AD virus (ADV), also known as suid herpesvirus type 1 (SHV-1), from the *Alphaherpesvirinae* subfamily, genus *Varicellovirus*. The disease causes significant economic losses in pig farms, mainly by reducing the litter size, abortions and slow growth of the animals, as well as restrictions on movements related to disease control. Young animals develop symptoms of central nervous infection while growing pigs and adults display respiratory signs. Pig is the primary host and the main disseminator of the disease. Moreover, in pigs surviving the acute infection, virus can develop a latency stage, and, subsequently, act as a reservoir [5]. It is a notifiable disease by the World Organisation for Animal Health (OIE).

SHV-1 has a linear double-stranded DNA genome of about 142 Kb. Its genome has been fully sequenced [6] and comprises two unique sequences, the 5' long and the 3' short, the last one flanked by two inverted repeated sequences. The genome encodes

more than 70 proteins [7], and two spliced transcripts are described: US1 and the Large Latency Transcript (LLT), while a third (UL15) has a putative splicing [6]. The virus has the possibility to induce latency in nervous system after an acute infection circularizing its genome and persisting like an episome with limited viral gene expression [8]. The molecular bases of latency in herpes viruses are still unknown. The control of the disease has been based on vaccination campaigns with the use of marker vaccines obtained by deletion of determined genes like TK and gE.

During last years, microRNAs (miRNAs) have been described and characterised as small non-coding RNAs involved in post transcriptional regulation of gene expression in animals, plants and some DNA viruses. They participate in a wide range of biological processes acting mainly through down-regulation of target messenger RNAs (mRNAs) by blocking protein translation or inducing mRNA degradation [9-15]. In addition, miRNA expression has also been associated with many different pathological processes, such as cancer, neurological disorders, inflammatory pathologies, cardiovascular diseases and infectious diseases [16-21].

Recently, several studies have confirmed the existence of viral miRNAs [22-24]. They are miRNAs codified by the viral genome, and mainly, they are described in the herpes virus family [24-26]. These viral miRNAs work supporting the development of viral infection by using host miRNA biogenesis system to control the expression of their own and host genes [27,28]. They can take advantage of a conserved gene regulatory mechanism within the host cell and establish a cellular environment conducive to viral replication [29]. Viral miRNAs act by regulating fundamental cellular processes in immunity, apoptosis and key steps in the transition from latent to lytic infection [8,30-32]. In contrast to viral proteins, miRNAs can regulate host and own gene expression avoiding protein factors exposed to host antigenic immune response and, moreover, they also have relative little coding capacity [26,29,33]. Most studied herpes virus express miRNAs during latency, even being a restrictive gene expression stage. In SHV-1, for instance, the 13 miRNAs described up to date in miRBase database (v19, August 2012, URL: <http://www.mirbase.org/>, [34-36]) are encoded in the large latency transcript (LLT), the unique transcript expressed during the latency stage which

generates the Latency Associated Transcripts (LAT) [37,38]. However, little is known about the key roles that viral miRNAs can develop during the virus infection process. The virus is transmitted primarily through physical contact between pigs by secretions. First replication takes part in epithelial tissues and entries directly into the sensory nerve endings in the nasopharynx. The normal AD incubation period is from 2 to 6 days. There are several porcine cells used in many scientific studies that allow the laboratory culture of SHV-1, such as the cell line derived from pig kidney PK-15 [5,37] or dendritic cells [38], but there are no studies involving pig tissues infected with SHV-1.

In order to unravel the role of viral miRNAs during the ADV infection process, a SHV-1 experimental infection using Landrace pigs was performed using two viral strains (NIA-3 and Begonia). Furthermore, *in vitro* infections with the same viral strains were also conducted. Characterisation and functional study of viral miRNAs are crucial to understand the molecular bases of herpes virus pathogeny and, consequently, to develop mechanisms to fight against the disease and improve pork production.

## **MATERIALS AND METHODS**

### **Biological material**

Two strains of SHV-1 were used in the experiment: the NIA-3 virulent strain and the Begonia attenuated strain. Begonia strain is derived from NIA-3 strain and is used as live attenuated vaccine.

A total of 20 4-week-old Landrace pigs were used in the experimental infection with both SHV-1 strains. All animal procedures were performed in CReSA biosafety level 3 (BSL3) facilities (Centre de Recerca en Sanitat Animal, Universitat Autònoma de Barcelona, Bellaterra, Spain) and were carried out according to Spanish and European animal experimentation ethics law and approved by the institutional animal ethics committee of Universitat Autònoma de Barcelona. Porcine Kidney (PK-15) cell lines were used for *in vitro* infection also with NIA-3 and Begonia SHV-1 strains.

### **Cell culture and animal infections**

PK-15 cell line was used for viral stocks preparation and *in vitro* infections. Cells were grown at 37°C and 5% CO<sub>2</sub> and maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 5% FCS, 100 µg/ml streptomycin and 100 IU/ml penicillin. A viral stock with a titre of 10<sup>7,57</sup> TCID<sub>50</sub>/mL was prepared in PK-15 cells for NIA-3 virulent strain and 10<sup>8,49</sup> TCID<sub>50</sub>/mL for Begonia attenuated strain.

PK-15 cell cultures at 65% of confluence were infected with a MOI of 0.05 with NIA-3 or Begonia strains. PK-15 cells inoculated with DMEM were maintained as non-infected cells. Samples were recovered at 12, 24 and 30 hours post infection (hpi). Cytopathic effect for both viruses was observed at 24 and 30 hours affecting 20-30 and 90% of the cells, respectively.

In the *in vivo* experimental infection, 9 animals were intranasally inoculated with 10<sup>3</sup> TCID<sub>50</sub> NIA-3 virulent strain, 6 with the same dose of Begonia attenuated strain and 5 control animals with PBS. One animal per group was euthanized at 4, 5 and 6 days post infection (dpi). The remaining animals were slaughtered at 7 dpi. Olfactory bulb (OB) and trigeminal ganglia (TG) samples for each animal were collected, immediately snap-frozen in liquid nitrogen and stored at -80°C until use. All samples were taken from CReSA BSL3 facilities (Bellaterra, Spain) under veterinary supervision.

### **RNA isolation**

Total RNA was isolated using TRIzol<sup>®</sup> reagent (Invitrogen, Carlsbad, USA) following the manufacturer's recommendations, quantified using ND 1000 Nanodrop<sup>®</sup> Spectrophotometer (Thermo Scientific, Wilmington, USA) and its quality was assessed on an Agilent 2100 Bioanalyzer using the RNA 6000 Nano kit (Agilent Technologies, Santa Clara, USA).

### **Small RNA library construction and high throughput sequencing**

A total of 21 libraries were performed in order to high throughput sequence them (Table 1). 9 libraries were from cell cultures infected with NIA-3, Begonia or mock-infected (control) cells at 12, 24 and 30 hpi. The remaining 12 libraries belonged to

animal infections, 6 for each tissue (OB and TG). Selected animals were 3 NIA-3 infected animals euthanized at 4, 5 and 7 dpi, 2 Begonia infected animals sacrificed at 4 and 7 dpi and 1 control animal (mock-infected) necropsied at 5 dpi.

Small RNA fraction from each sample was excised and isolated from denaturing 12.5% polyacrilamide gel electrophoresis (PAGE) using miSpike<sup>™</sup> (IDT<sup>®</sup>, Coralville, USA) as internal size marker. 50 µg of total RNA for each sample were loaded on separate gels to avoid cross-contamination. Gels were stained with GelStar<sup>®</sup> Acid Nucleic Gel Stain (Lonza, Basel, Switzerland) for UV visualization. Excised small RNA fraction were purified using Performa<sup>®</sup> DTR gel filtration cartridges (EdgeBio, Gaithersburg, USA). Briefly, 3' and 5' linkers from miRCat<sup>™</sup> kit (IDT, Coralville, USA) were ligated at both ends of the small RNAs in two separated reactions using a T4 RNA ligase without ATP (Fermentas, Germany) and T4 RNA ligase with ATP (Ambion, Austin, USA), respectively. Between 3' and 5' primer ligations, the 60 nt RNAs were purified by PAGE to eliminate unligated products. Then, linked products were used to perform a reverse transcription reaction using the SuperScript<sup>™</sup> III Reverse Transcriptase kit (Invitrogen<sup>™</sup>, Carlsbad, USA) and the cDNA obtained was amplified with the Expand High Fidelity System (Roche, Germany). PCRs were done to amplify the cDNA with primers complementary to 3' and 5' linkers and, in addition, they included multiplex identifiers at the 5' end (a five nucleotide sequence tag) to allow differentiation between libraries. The number of PCR cycles was optimized for each sample, in order to minimize/avoid saturation, ranging from 21 to 30. Purification was carried out by using QIAquick PCR Purification Kit (Qiagen<sup>®</sup>, Germany). Libraries were quantified with Qubit<sup>™</sup> fluorometer, Quant-IT<sup>™</sup> (Invitrogen<sup>™</sup>, Carlsbad, USA), prepared to a 10<sup>11</sup> DNA molecules/µL and equimolecular pooled according to their indexes. Ion Torrent adapters were ligated to 30 ng of pooled DNA and libraries were then amplified with Ion Torrent primers for 8 cycles, size selected (2% E-Gel Size Select, Invitrogen), and sequenced in four 314 chips in Ion PGM<sup>™</sup> sequencer (Life Technologies, Carlsbad, USA) following the manufacturer's protocol at DNA sequencing facilities at CRAG (Bellaterra, Spain). Software version for base calling was Torrent-Suite v2.0.1 (Life Technologies, Carlsbad, USA).

### **Sequence processing scheme**

Primers sequences were trimmed and only those insert sequences between 15 and 29 nucleotides and with total number of sequences  $\geq 3$  were kept for further analysis. For porcine miRNA profiling, sequences were compared to all available miRNA sequences (miRBase v19) using local Blast. Parameters were set to 100% identity and up to 4 mismatches allowed at the end of the sequences to assume variability on 3' and 5' ends [39].

For viral miRNA discovery, sequences were blasted to SHV-1 genome (NCBI Reference Sequence: NC\_006151.1) considering 100% of alignment and identity (perfect match). Sequences positioned at annotated regions were discarded. Remaining sequences were clustered taking into account only the position in the genome. Hence, sequences positioned in the same region were grouped and the sequence with higher copy number (CN) was selected as the reference sequence for each cluster. A total of 14 clusters (viRs) were described and considered as putative viral miRNAs (Table S1). viRs were blasted to viral miRNAs described in miRBase v19 and compared to the described SHV-1 miRNAs up to date [37,38]. Flanking regions (50 nt) of the selected reference sequences for each cluster were used to predict pre-miRNA folding structure using MFold software [40] following the guidelines reported by Ambros et al. [41] for animal miRNAs (Figure S1). At the end, 8 viRs were selected for RT-qPCR validation (Table 2).

Differences in host and viral miRNA expression were assessed. Total number of sequences obtained for each porcine miRNA or viR was normalised by library size (in counts per thousand) and, then, averaged by group. Fold changes (FC) between groups were calculated using normalised data.

### **RT-qPCR validation**

For RT-qPCR validations, additional samples from animal infected tissues were added (OB tissue: 5 NIA-3, 4 Begonia and 2 mock-infected; TG tissue: 5 NIA-3, 3 Begonia and 2 mock-infected, reaching a total of 40 samples. Two samples (37TG and 41TG from NIA-3 group) were not available for RT-qPCR validation due to lack of amount of cDNA. See Table 1). RT reactions were performed in duplicate using total RNA as previously

described by Balcells et al. [42]. Briefly, 1 µg of total RNA in a final volume of 20 µL including 2 µL of 10x poly(A) polymerase buffer, 0.1 mM of ATP, 0.1 mM of each dNTP, 1 µM of RT-primer, 200 U of M-MuLV Reverse Transcriptase (New England Biolabs, USA) and 2,5 U of poly(A) polimerase (New England Biolabs, USA) was incubated at 42°C for 1 hour and at 95°C for 5 minutes for enzyme inactivation. Non template controls (NTC), minus RT and minus poly(A) polymerase controls for each sample were included.

DNA primers for each viR were designed following the methodology suggested by Balcells et al. [42] (Table S2). qPCR reactions were performed in duplicate in 20 µL final volume including 10 µL SYBR® Select Master Mix (Life Technologies, Carlsbad, USA), 300 nM of each primer and 5 µL of a 1:20 dilution of the cDNA cell cultures or 1:15 dilution of the cDNA animal infected tissues on an 7900HT Sequence Detection System (Applied Biosystems, Warrington, UK). Standard curves were generated by 5 fold serial dilutions of a pool of NIA-3 and Begonia infected cell cultures cDNAs in order to calculate the qPCR efficiency. Thermal profile was set as follows: 50°C for 2 min, 95°C for 10 min and 40 cycles at 95°C for 15 sec and 60°C for 60 sec. NTC and minus poly(A) polymerase controls were included. Melting curve analysis was included at the end of the qPCR to detect unspecific amplifications. Hsa-miR-93, Hsa-miR-25, Ssc-miR-106a, Ssc-miR-17-5p, Ssc-miR-26a were used as reference miRNAs [43,44].

Quantities from each sample were obtained from the calibration (standard) curve added in each RT-qPCR reaction, and only those samples classified as quantifiable were used for statistical analyses. GeNorm v.3.5 software [45] was used to examine the stability of the reference miRNAs ( $M < 1.5$ ) and to obtain a normalization factor (NF). The quantity obtained from each miRNA was normalised by the NF and FCs were calculated in relation to the lowest normalised value. Finally, FCs were  $\log_2$  transformed in order to perform the statistical analyses with the General Linear Models procedure of the Statistical Package for the Social Scientists (IBM® SPSS® Statistics 19; IBM Corporation, Armonk, USA). *In vitro* expression data were analysed to study the differences between the infection groups (NIA-3, Begonia and mock-infected) and time groups (12/24/30 hours) by a two-way analysis of variance

(ANOVA). Significance threshold was set at  $\alpha < 0.001$  due to the unequal sample size and variances. Estimated marginal means were also calculated using the least significance difference (LSD) as confidence interval adjustment. The same strategy was followed in NIA-3 *in vivo* expression data for the tissue group (OB, TG) and time groups (4/5/6/7 days).

### **Target prediction and functional analysis**

DIANA - microT v3.0 web server [46,47] was used to identify *in silico* potential mRNA targets for the most abundant and the differentially expressed porcine miRNAs. Porcine genes are not included in the current version of DIANA - microT v3.0 and predictions were based on the human mRNA:miRNA interactions assuming sequence conservation. *In silico* functional annotation of putative mRNA target genes for each miRNA were analyzed with WEB-based Gene Set Analysis Toolkit (WebGestalt, [48]). Predicted miRNA targets were functionally annotated through the biological process information supported by Gene Ontology (GO, [49]) and the pathways in which they were involved were described by using the Kyoto Encyclopedia of Genes and Genomes database (KEGG, [50,51]). Over or under represented functional categories were identified with hypergeometric test corrected by the multiple test adjustment proposed by Benjamini & Hochberg [52]. Significant threshold was set at  $\alpha < 0.05$ .

The miRanda algorithm [53] was utilised to predict putative targets for viral miRNAs (viRs), using the following parameters: -sc 140 -en 20. Strict alignments were required in the seed region (G:U wobble is not allowed). Cytoscape 2.8.2 software [54,55] was used to build the gene regulatory network formed by viral miRNAs their target genes from SHV-1 genome. Most abundant DE porcine miRNAs in infected samples were also added to gene regulatory network.



## RESULTS

### Sequence miRNA annotation

A total of 21 small RNA libraries (9 from PK-15 cell line cultures and 12 from *in vivo* animal infection) were sequenced in a Ion PGM™ sequencer (Life Technologies). After trimming the adaptors sequences, inserts ranging from 15 to 29 nt (corresponding to miRNA size) and found more than two times (copy number (CN)>2) were aligned to miRBase database (v19). 435,434 counts (4,029 unique sequences) could be aligned to miRBase database, representing a 37% of total counts used in this study (Table 3). In PK-15 cell line libraries, 229 miRNAs were described, 109 of them had been already described in pig, 113 were orthologous and 7 were viral miRNAs (6 from SHV-1 and 1 from Rhesus Rhadinovirus – RRV). In *in vivo* animal infection libraries, a total of 302 miRNAs were described, 150 that had been previously described in *Sus scrofa*, 151 were orthologous and only 1 viral miRNA from SHV-1 was found. No viral miRNAs were identified in mock-infected PK-15 cell cultures and in mock-infected animals.

### Differential expression analysis

#### *In vitro*

Looking for miRNA abundance differences among groups (NIA-3 infected, Begonia infected and mock-infected), 138 miRNAs (60%) were more expressed in mock-infected group and 91 miRNAs (40%) were more expressed in infected groups. Furthermore, 35 miRNAs were only expressed in infected groups, and two miRNAs were only expressed in mock-infected group.

miRNAs were considered differentially expressed (DE) when fold change (FC) difference between groups was greater than 5 or when a miRNA was not expressed in both infected groups or in mock-infected group. Out of the 229 miRNAs described in the *in vitro* profile, 111 (48%) miRNAs were DE; of which, 69 (30%) were overexpressed in mock-infected group and 42 (18%) over-expressed in infected groups (Table S4). Looking at those most abundant miRNAs (CN>100), we observed a clear predominance of those miRNAs overexpressed in mock-infected group (Table 4), such as miR-125b-5p, miR-99b-5p and miR-100. It was relevant that the only miRNA overexpressed in

infected groups with CN>100 was the viral miRNA miR-LLT1 (CN=3,280). There were two miRNAs presenting an up-regulation in Begonia infected group and a down-regulation in NIA-3 infected group, being, consequently, DE between both infected groups: miR-92a and miR-92b-3p. Interestingly, there was a miRNA, miR-2887, which presented high FC differences between mock-infected and NIA-3 infected groups (204 fold). Although it was over-expressed in mock-infected group, it also was DE between NIA-3 and Begonia infected groups, like miR-4286 and let-7d-3p.

Focusing on the described viral miRNAs, miR-LLT9 (CN=3) was only described in NIA-3 infected group. miR-LLT6, miR-LLT8 and miR-LLT11a were more expressed in NIA-3 infected group and miR-LLT1 and miR-LLT2 presented more expression in Begonia infected group. Only miR-LLT2 was DE between infected groups, although its abundance was only 13 counts. The viral miRNA from RRV (miR-rR1-5, CN=415) was more expressed in mock-infected group.

### ***In vivo***

In contrast to *in vitro* approach, miRNA abundance differences revealed a turnaround in which 211 (70%) miRNAs were more expressed in infected NIA-3 and Begonia groups, and the remaining 91 (30%) miRNAs were more expressed in mock-infected group. Moreover, 75 miRNAs presented only expression in infected groups, and only two miRNAs were only expressed in mock-infected group.

Differential miRNA expression analysis showed a total of 133 (44%) DE miRNAs out of the 302 described miRNAs, of which 99 (33%) were overexpressed in infected groups and 34 (11%) overexpressed in mock-infected group (Table S5). Looking at those miRNAs which were more abundant (CN>100), most of them were more expressed in infected groups, like let-7a, miR-378 and miR-129b (Table 4). If we focus on those more expressed miRNAs in infected groups and DE between them, we observed that miR-378 was more expressed in NIA-3 group. Moreover, miR-206 was a very high DE miRNA (FC=648) between both infected groups and not expressed in mock-infected group, presenting more expression in NIA-3 infected group. Curiously, miR-133a and miR-133b were up-regulated in NIA-3 infected group and down-regulated in Begonia

infected group, being also DE between infected groups. On the other hand, miR-451 and miR-1249 were the most expressed miRNAs presenting an up-regulation in mock-infected group. The only viral miRNA detected (miR-LLT1, CN=9) were only expressed in NIA-3 infected group, diverging from its expression in the *in vitro* approach, where it was more expressed in Begonia infected group.

Out of 361 total described miRNAs taking into account both approaches, 170 miRNAs were present in both *in vitro* and *in vivo* profiles and 191 miRNAs were present only in one profile (59 miRNAs from *in vitro* approach and 122 miRNAs from *in vivo* approach, see Table S3). There was not much correlation in the expression of profile-shared miRNAs. Only 27 miRNAs followed the same expression profile in both approaches, and 11 of them were DE in both groups: 4 miRNAs (miR-26b-5p, miR-29b-2-5p, miR-450b-5p and miR-450c-5p) were only expressed in infected groups and the remaining 7 miRNAs (let-7b-3p, miR-193b, miR-345-5p, miR-1306-5p, miR-2779, miR-2898 and miR-4286) were more expressed in mock-infected group, although none of them presented a CN>100 in both profiles. There were 35 miRNAs expressed with a minimum of 100 counts which showed a different pattern of expression between *in vitro* and *in vivo* approaches, and only two presented differential expression also in both approaches: let-7i-5p and miR-30d-5p (Table 4). In both cases, they were more expressed in mock-infected group in the *in vitro* approach, and their expression changed in the *in vivo* approach, having a major expression in NIA-3 infected group. Among the non-profile-shared miRNAs, there were 6 high expressed miRNAs (CN>1,000 counts): miR-99a-5p (11,178 counts), miR-10a-5p (3,570 counts), miR-133a (1,990 counts), miR-218b (1,887 counts), miR-9-3p (1,620 counts) and miR-129a (1,566 counts). All these 6 miRNAs were expressed in the *in vivo* profile, except miR-10a-5p which was expressed in the *in vitro* profile.

### **Target prediction and functional analysis of host miRNAs**

*In silico* target prediction were performed for those most abundant and DE porcine miRNAs in NIA-3 and Begonia infected groups in the *in vivo* approach: miR-206, let-7a and miR-378, and also for those miRNAs presenting differential expression between NIA-3 and Begonia infected groups: miR-92a and miR-92b-3p from the *in vitro*

approach and miR-133a from the *in vivo* approach (Table 4). A total of 2,051 target genes were identified (Table S6) and functionally analysed through KEGG pathways database. Significant related pathways to target genes were found for all miRNAs except for miR-378. Interestingly, many pathways related to viral infection process and immune response resulted significant, such as RIG-I-like receptor signalling pathway, B and T cell receptor signalling pathways, Fc gamma R-mediated phagocytosis and chemokine signalling pathway. Furthermore, there also were described pathways involved in more general biological processes, such as cell cycle, apoptosis, endocytosis, focal adhesion and RNA transport and degradation. Interestingly, pathways focused on nervous system were also described, like axon guidance, neurotrophin signalling pathway and long-term potentiation.

#### **Viral miRNA discovery and RT-qPCR validation**

Trimmed and non-empty reads ranging from 15 to 29 nt were also aligned to SHV-1 genome (NC\_006151.1) considering only 100% of alignment and identity. A total of 3,979 counts (50 unique sequences) yielded a positive match, 3,948 counts (47 unique sequences) from PK-15 cell line libraries and 31 counts (5 unique sequences) from *in vivo* animal infection libraries (Table 3). All sequences which matched at an annotated region were removed and the remaining sequences were clustered by the position in the SHV-1 genome resulting in 14 putative viral miRNAs. All of them were identified in cell cultures and only 2 of them could also be described in animal infection samples: viR02 (mir-LLT1 in miRBase v19) and viR09 (Table 2). None of them were found in mock-infected cell cultures or in mock-infected animals. From these 14 putative viral miRNAs, viR03, viR07, viR10, viR12 and viR13 were already described at miRBase as miR-LLT2, miR-LLT6, miR-LLT8, miR-LLT9 and miR-LLT11a, respectively, and, consequently, were discarded for RT-qPCR validation. viR02 was also described at miRBase but it was maintained to be the most expressed viral miRNA in the study. viR01, the only putative viR which was not located at LLT intronic region, did not succeed at performing their pre-miRNA structure, and was also removed. Finally, 8 viRs remained to confirm their expression (Table 2) and a validation protocol through RT-qPCR was designed for them, which were successfully amplified with high RT-qPCR

efficiencies, ranging from 90% to 110%, and standard curves correlations were at least of 0.98.

Out of these 8 validated viRs, 5 of them were already described in previous studies [37,38] and remaining 3 viRs were new described viRs (Table 2). viR09 was described as miRNA offset RNA (moRNA) in the Wu study [37] and our study also confirmed that viR09 was the moRNA originated from miR-LLT8 (Figure S1). As moRNAs functions remain unknown [56,57], the described moRNA in this study (viR09) was kept in the functional analysis, as it was found in a major relative abundance than its contiguous mature miRNA and it was favourably detected through RT-qPCR.

All detected and validated viRs were located in the intronic region of the LLT transcript, such as all previously described viRs. They were detected in the NIA-3 and Begonia groups from cell cultures and only presented expression in some samples from NIA-3 group from *in vivo* infection (Figure 1). Begonia group from *in vivo* infection as well as all mock-infected samples from both *in vitro* and *in vivo* infection resulted without expression for all studied viRs.

Statistical analyses from RT-qPCR *in vitro* data revealed no significant expression differences between NIA-3 and Begonia groups. Referring at time group, there were significant expression differences ( $p$ -value <.001), being viRs more expressed in those samples taken at 24 or 30 hpi than those taken at 12 hpi ( $p$ -value <.001), excepting for viR05, which showed more expression in 12 hpi than in 30 hpi, although without significant differences in any time group.

Regarding NIA-3 *in vivo* group RT-qPCR data, statistical analyses only could be performed for viR02, viR06 and viR09. Results showed differential expression ( $p$ -value <.001) for tissue group and time group in all three viRs, being more expressed in olfactory bulb than in trigeminal ganglia. In time group, viRs presented differential expression in 4, 5, 6 and 7 dpi, presenting directional increasing from dpi 4 to dpi 6 and an interesting decrease of their expression at dpi 7.

### Gene regulatory network viral miRNAs

miRanda algorithm was used to form the gene interaction network between the 8 described viRs and the 70 annotated genes in SHV-1 genome, including large latency transcript (LLT). It found 110 significant interactions between 7 viRs and 59 SHV-1 genes (Figure 2). No interactions could be retrieved for viR11. The most interacting viRs were viR04, viR14 and viR06 reaching at 48, 20 and 15 gene interactions, respectively. LLT was the gene associated to more viRs, up to 5: viR02, viR04, viR05, viR06 and viR14.

Most abundant DE porcine miRNAs were also added to the gene interaction network. Thus, miR-92a, miR-92b-3p, miR-133a, miR-378, let-7a and miR-206 generated a total of 63 significant interactions with 39 SHV-1 genes (Figure 2). miR-133a had 33 gene interactions, including LLT and the regulatory genes *EPO*, *IE180*, *UL41* and *UL48*. miR-92a, miR-92b-3p and miR-378 also interacted with LLT and with the regulatory genes *IE180*, *UL41* and *UL48*. Let-7a was associated to few viral genes, most of them related to structural role, like glycoproteins and capsid proteins. miR-206 was not associated to any viral gene.

### DISCUSSION

This study is the first work that describes the host and pathogen miRNA profile in an *in vitro* and an *in vivo* SHV-1 infection through high throughput sequencing. Regarding the host miRNAome, 193 porcine miRNAs out of a total of 306 annotated miRNAs in miRBase v.19 were described. Approximately half of all described miRNAs were orthologous, suggesting that there are still many porcine miRNAs to be discovered in order to complete the current annotation of porcine miRNAs. First conclusion observing both profiles was that there was low correlation among described miRNAs and also among their expression pattern. Looking at those most abundant miRNAs in each profile (CN>100), only 33% of miRNAs were shared. It is known that there are usually discrepancies between *in vitro* and *in vivo* results, and, moreover, our study has been developed with material of different nature: cell lines derived from kidney and

nervous tissues, helping to perform different miRNA profiles and, assuming different expression patterns. Just comparing the PK-15 profile with the kidney profile from Timoneda et al., previous study [58], the shared miRNAs were up to 42%, more than in the *in vivo* profile, showing some analogy between PK-15 cells and kidney. However, PK-15 cell lines were chosen because they were a good substrate for the laboratory culture of SHV-1, and, on the other hand, OB and TG tissues were chosen to be the first tissues where the virus replicates.

Sequenced libraries revealed a different porcine miRNA profile when the SHV-1 infection was present, and presented some homology between the two viral strains used in this study, NIA-3 (virulent strain) and Begonia (attenuated strain). The change of expression of some porcine miRNAs between infected and mock-infected samples reflects that miRNAs can play key roles during the viral infection process, where virus can affect cellular miRNA expression profile on their own benefit. In this sense, many porcine miRNAs were described to be down-regulated in the infected samples, particularly in the *in vitro* infection, such as miR-125b-5p, miR-99b-5p, miR-100 and miR-2887, suggesting that viral mechanisms can affect host miRNA expression. For instance, miR-100 has already been described to be down-regulated in human cytomegalovirus infection [59].

On the other hand, there were miRNAs up-regulated during SHV-1 infection, mainly in the *in vivo* approach, like let-7a, miR-378 and miR-206, the most DE expressed miRNAs in infected samples. Interestingly, they presented target genes related to pathways involved in immune response, such as RIG-I-like receptor signalling pathway, responsible for detecting viral pathogens, or also B and T cell receptor signalling pathways, which are key components for the activation of adaptive immunity and T lymphocytes, respectively, and ensure an efficient response of the immune system. They also were related to Fc gamma R-mediated phagocytosis, which plays an important role in host-defence mechanisms through the uptake and destruction of infectious pathogens, and chemokine signalling pathway, which works on the inflammatory immune response. All these 3 miRNAs were particularly more expressed in NIA-3 infected group, so they could work activating those pathways related to the

response against the viral infection. Moreover, differential expression between NIA-3 and Begonia infected groups were identified in miR-206 (FC=648) and miR-378 (FC=5). We can also add miR-133a (FC=108) in this group of DE miRNAs between virus strains (Table 4). This change of behaviour could be explained by the differences in both virus strains, as Begonia is an attenuated strain which has been genetically modified, removing glycoprotein E gene (gE) and thymidine kinase gene (tK), in order to be less effective at virus replication. In this sense, these miRNAs increase their expression when the virulent strain NIA-3 is present in order to react against infection.

Another miRNA which changed its expression pattern between NIA-3 and Begonia infections was miR-92a. While in the *in vitro* samples miR-92a was DE presenting an up-regulation in Begonia strain infection, in the *in vivo* samples its expression was higher in NIA-3 strain infection. A wide range of significant pathways were associated to miR-92a, from pathways related to neuronal functions such as axon guidance, neuro active-ligand receptor activation and neurotrophin signalling pathway, to pathways related to more general cellular functions like endocytosis, RNA degradation and focal adhesion. We could argue that in the *in vitro* approach, the Begonia strain virus does not receive a strong defensive response from the host and it could replicate in a similar level as NIA-3 begonia strain, according to RT-qPCR results from viral miRNAs expression. In this sense, the host response generated in the Begonia *in vivo* infection could cause the miR-92a fall of expression. Regarding to miR-92b-3p, it maintained its expression in both approaches, being more expressed in Begonia strain infection than in NIA-3 strain infection. As miR-92a, miR-92b-3p target genes were associated to a wide range of biological processes as well as nervous system pathways. As its expression has shown invariable in different tissues, despite of the viral infection, it could play a more general cell function.

As mentioned above, the differences observed in miRNA profiles between *in vitro* and *in vivo* approaches in both infected and mock-infected samples could be explained by the employed biological approach: cell cultures or animal tissues. We observed that the development of the viral infection did not happen in the same way in cell cultures than in tissues from experimentally infected pigs. The ability to react and fight against



a viral attack was not the same due, among others, to the lack of immune response in cell culture. The miRNA expression variability can be determined by the frequency of most expressed miRNAs in each approach. In the *in vitro* infection, miR-23a-3p was the most expressed miRNA representing the 50% of annotated reads, while miR-125b-5p was the most expressed miRNA in the *in vivo* infection, in both olfactory bulb and trigeminal ganglia, and only represented the 25% of the annotated reads. Thus, there was a major variability in the animal infection model. In comparison to previous studies, Anselmo et al. [38] described miR-21 as the most expressed miRNA in their study by using dendritic cells as approach, representing almost 91% of all small RNA sequence tags, while Wu et al. [37] found miR-7f as the most expressed miRNA in a PK-15 cell line culture, being the 17% of total small RNA reads. As miRNAs are spatial and temporal specific, the disparity of results must be taken in consideration.

Moreover, miRNA differences in *in vitro* and *in vivo* profiles became more evident by using two different strains of SHV-1. In this sense, the two strains were capable of infecting cell cultures producing cytopathic effect, while in the animals, the attenuated strain was unable to produce clinical signs showing a less effective replication, and therefore, we were not able to detect viral miRNAs in these samples through RT-qPCR. Focusing on viral miRNAs (viRs), we could detect differences in sequenced data expression between *in vitro* and *in vivo* groups through miRBase homology. In cell cultures, 7 viRs could be detected, being miR-LLT1 (CN=3,280) the most expressed viral miRNA. In the animal infection, only one viR could be detected in sequenced data, which was also miR-LLT1 (CN=9). Even so, viral miRNA discovery based on genome sequence homology approach was able to detect and confirm through RT-qPCR the expression of 8 viRs. It is of great interest because this study confirms the expression of viral miRNAs originated from SHV-1 LLT transcript during lytic infection, assuming their role in the early stage infection process.

RT-qPCR expression data confirmed that there were no significant differences in viral miRNAs expression between infected groups (NIA-3 virulent strain and Begonia attenuated strain) in the *in vitro* infection, although no expression were detected in Begonia infected group in the *in vivo* infection. As expected, samples taken at 24 and

30 hpi in the *in vitro* infection presented more viR expression than samples taken at 12 hpi. In the *in vivo* approach, significant differential expression were described between all 4 times where samples were taken (4, 5, 6 and 7 dpi), and presented an increasing expression from day 4 to day 6 and revealing a sudden decrease at day 7. This result could be explained because the expressed miRNAs in this initial phase of the acute infection in our animal experiment could play a role in the establishment of the infection and can be different from miRNAs expressed in a latent phase, which can be the same than whose are found in cell culture, as has been found for other herpesvirus [60].

Viral gene network analysis deciphered the complex interaction between the described viRs and SHV-1 genes (Figure 2). Interestingly, viRs were associated with almost all described SHV-1 genes [7], as described in a previous study [37]. Large Latency Transcript (LLT) was the most interacted transcript by viRs, confirming the important role of them in the latency stage development. As all previously described viral miRNAs in SHV-1, the new viRs described in this study were also encoded in the LLT intronic region, confirming it as the primary miRNA precursor region. In addition, most DE host miRNAs were also associated to viral regulatory mechanisms, relating the interaction of host miRNAs with the expression of viral genes. The gene regulatory network deciphered the important role of viR04 presenting up to 24 gene interactions related to structural role, from proteins related to the virion envelope (glycoproteins) to tegument and capsid proteins. Thus, viR04 may play an important structural role, although it also presented numerous interactions to genes involved in processes like DNA repair and recombination, DNA cleavage, encapsidation and packaging. viR02 and viR14 were mainly associated to structural genes related to virion envelope glycoproteins and tegument proteins, and could play important roles during viral entry and virion morphogenesis. In contrast, viR05 was associated to non-structural genes related to DNA cleavage and encapsidation and DNA replication and packaging. Curiously, viR06 presented many interactions with genes related to viral egress (*UL3.5*, *UL31*, *UL37* and *UL53*), which could play an interesting role in the viral egress process. viR02, viR04, viR06 and viR14 were also linked to genes associated to regulatory functions (*EPO*, *IE180*, *UL41*, *UL48* and *UL54*). viR09, the miRNA described in this

study, was associated to few viral genes, which were associated to DNA cleavage, packaging and replication functions, as well as viral egress and structural role regarding virion envelope proteins. Regarding the involvement of host miRNAs in the regulatory network, miR-133a, miR-92a, miR-92b-3p and miR-378 were related to LLT and the regulatory genes *EPO*, *IE180*, *UL41* and *UL48*. It could mean that they could play an active role in fighting against the viral infection process. Moreover, miR-133a was related with many structural and non structural genes, confirming the importance of miR-133a in the ADV infection, as it was previously determined to be related with host-defence pathways. As let-7a is an ubiquitous miRNA [43,44] and it was related to viral genes with structural roles, it could be developing functions related to viral packaging recognition.

The regulation roles developed by miRNAs not only in the own host gene machinery but also in the viral infection mechanisms were evidenced in this study, using *in vitro* and *in vivo* approaches. In addition, this study increases the knowledge about miRNAs functional roles in a herpesvirus infection and their host-pathogen interactions, supported by an *in vivo* approaching using nervous tissue.

## **AKNOWLEDGEMENTS**

This work was supported by the projects AGL2007-66371-C02 and AGL2010-22358-C02-01 from Ministerio de Ciencia e Innovación. OT is recipient of PhD fellowship (Programa de Formación al Personal Investigador – FPI) from the Spanish government. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors have declared that no competing interests exist.

## REFERENCES

1. Amills M, Clop A, Ramírez O, Perez-Enciso M. (2010) Origin and genetic diversity of pig breeds. *eLS* .
2. FAOSTAT. (2011) [Http://faostat3.fao.org/home/index.html#DOWNLOAD](http://faostat3.fao.org/home/index.html#DOWNLOAD).
3. Distl O. (2007) Mechanisms of regulation of litter size in pigs on the genome level. *Reprod Domest Anim* 42 Suppl 2: 10-16.
4. Kanis E, De Greef KH, Hiemstra A, van Arendonk JA. (2005) Breeding for societally important traits in pigs. *J Anim Sci* 83: 948-957.
5. Pejsak Z, Truszczynski M. (2006) Aujeszky's disease (pseudorabies). In: Straw B, Zimmerman J, D'Allaire S, Taylor D, editors. *Diseases of Swine*, Ninth Edition. Oxford, UK: Blackwell Science. pp. 419-433.
6. Klupp BG, Hengartner CJ, Mettenleiter TC, Enquist LW. (2004) Complete, annotated sequence of the pseudorabies virus genome. *J Virol* 78: 424-440.
7. Pomeranz LE, Reynolds AE, Hengartner CJ. (2005) Molecular biology of pseudorabies virus: Impact on neurovirology and veterinary medicine. *Microbiol Mol Biol Rev* 69: 462-500.
8. Boss IW, Plaisance KB, Renne R. (2009) Role of virus-encoded microRNAs in herpesvirus biology. *Trends Microbiol* 17: 544-553.
9. Huntzinger E, Izaurralde E. (2011) Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nat Rev Genet* 12: 99-110.
10. Krol J, Loedige I, Filipowicz W. (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* 11: 597-610.
11. Guo H, Ingolia NT, Weissman JS, Bartel DP. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466: 835-840.
12. Bartel DP. (2009) MicroRNAs: Target recognition and regulatory functions. *Cell* 136: 215-233.
13. Ambros V. (2008) The evolution of our thinking about microRNAs. *Nat Med* 14: 1036-1040.
14. Ruvkun G. (2008) The perfect storm of tiny RNAs. *Nat Med* 14: 1041-1045.
15. Bartel DP. (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.
16. Esteller M. (2011) Non-coding RNAs in human disease. *Nat Rev Genet* 12: 861-874.

17. Akkina S, Becker BN. (2011) MicroRNAs in kidney function and disease. *Transl Res* 157: 236-240.
18. Sayed D, Abdellatif M. (2011) MicroRNAs in development and disease. *Physiol Rev* 91: 827-887.
19. Hata A. (2013) Functions of microRNAs in cardiovascular biology and disease. *Annu Rev Physiol* 75: 69-93.
20. Ma L, Qu L. (2013) The function of MicroRNAs in renal development and pathophysiology. *J Genet Genomics* 40: 143-152.
21. Rederstorff M, Hüttenhofer A. (2010) Small non-coding RNAs in disease development and host-pathogen interactions. *Current Opinion in Molecular Therapeutics, Curr Opin Mol Ther* 6: 684-694.
22. Grundhoff A, Sullivan CS, Ganem D. (2006) A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. *RNA* 12: 733-750.
23. Sullivan CS, Grundhoff AT, Tevethia S, Pipas JM, Ganem D. (2005) SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* 435: 682-686.
24. Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, et al. (2004) Identification of virus-encoded microRNAs. *Science* 304: 734-736.
25. Kincaid RP, Sullivan CS. (2012) Virus-encoded microRNAs: An overview and a look to the future. *PLoS Pathog* 8: e1003018.
26. Grundhoff A, Sullivan CS. (2011) Virus-encoded microRNAs. *Virology* 411: 325-343.
27. tenOever BR. (2013) RNA viruses and the host microRNA machinery. *Nat Rev Microbiol* 11: 169-180.
28. Cullen BR. (2013) MicroRNAs as mediators of viral evasion of the immune system. *Nat Immunol* 14: 205-210.
29. Skalsky RL, Cullen BR. (2010) Viruses, microRNAs, and host interactions. *Annu Rev Microbiol* 64: 123-141.
30. Ouellet DL, Plante I, Landry P, Barat C, Janelle ME, et al. (2008) Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. *Nucleic Acids Res* 36: 2353-2365.
31. Cullen BR. (2009) Viral and cellular messenger RNA targets of viral microRNAs. *Nature* 457: 421-425.

32. Besecker MI, Harden ME, Li G, Wang XJ, Griffiths A. (2009) Discovery of herpes B virus-encoded microRNAs. *J Virol* 83: 3413-3416.
33. Jurak I, Griffiths A, Coen DM. (2011) Mammalian alphaherpesvirus miRNAs. *Biochim Biophys Acta* 1809: 641-653.
34. Kozomara A, Griffiths-Jones S. (2011) miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* 39: D152-7.
35. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. (2008) miRBase: Tools for microRNA genomics. *Nucleic Acids Res* 36: D154-8.
36. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. (2006) miRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34: D140-4.
37. Wu YQ, Chen DJ, He HB, Chen DS, Chen LL, et al. (2012) Pseudorabies virus infected porcine epithelial cell line generates a diverse set of host microRNAs and a special cluster of viral microRNAs. *PLoS One* 7: e30988.
38. Anselmo A, Flori L, Jaffrezic F, Rutigliano T, Cecere M, et al. (2011) Co-expression of host and viral microRNAs in porcine dendritic cells infected by the pseudorabies virus. *PLoS One* 6: e17374.
39. Nielsen M, Hansen JH, Hedegaard J, Nielsen RO, Panitz F, et al. (2010) MicroRNA identity and abundance in porcine skeletal muscles determined by deep sequencing. *Anim Genet* 41: 159-168.
40. Zuker M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31: 3406-3415.
41. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, et al. (2003) A uniform system for microRNA annotation. *RNA* 9: 277-279.
42. Balcells I, Cirera S, Busk PK. (2011) Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC Biotechnol* 11: 70.
43. Timoneda O, Balcells I, Cordoba S, Castello A, Sanchez A. (2012) Determination of reference microRNAs for relative quantification in porcine tissues. *PLoS One* 7: e44413.
44. Peltier HJ, Latham GJ. (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: Identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* 14: 844-852.
45. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3: RESEARCH0034.

46. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, et al. (2009) Accurate microRNA target prediction correlates with protein repression levels. *BMC Bioinformatics* 10: 295.
47. Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, et al. (2009) DIANA-microT web server: Elucidating microRNA functions through target prediction. *Nucleic Acids Res* 37: W273-6.
48. Zhang B, Kirov S, Snoddy J. (2005) WebGestalt: An integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 33: W741-8.
49. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: Tool for the unification of biology. the gene ontology consortium. *Nat Genet* 25: 25-29.
50. Kanehisa M, Goto S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28: 27-30.
51. Ogata H, Goto S, Sato K, Fugibuchi W, Bono H, et al. (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 27: 29-34.
52. Benjamini Y, Hochberg Y. (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 57: 283-300.
53. Enright AJ, John B, Gaul U, Tuschl T, Sander C, et al. (2003) MicroRNA targets in drosophila. *Genome Biol* 5: R1.
54. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498-2504.
55. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. (2011) Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* 27: 431-432.
56. Shi W, Hendrix D, Levine M, Haley B. (2009) A distinct class of small RNAs arises from pre-miRNA-proximal regions in a simple chordate. *Nature Structural & Molecular Biology, Nat Struct Mol Biol* 2: 183-189.
57. Bortoluzzi S, Biasiolo M, Bisognin A. (2011) MicroRNA-offset RNAs (moRNAs): By-product spectators or functional players? *Trends Mol Med* 17: 473-474.
58. Timoneda O, Balcells I, Nunez JI, Egea R, Vera G, et al. (2013) miRNA expression profile analysis in kidney of different porcine breeds. *PLoS One* 8: e55402.
59. Wang Y, Liang Y, Lu Q. (2008) MicroRNA epigenetic alterations: Predicting biomarkers and therapeutic targets in human diseases. *Clin Genet* 74: 307-315.

60. Burnside J, Bernberg E, Anderson A, Lu C, Meyers BC, et al. (2006) Marek's disease virus encodes MicroRNAs that map to meq and the latency-associated transcript. *J Virol* 80: 8778-8786.

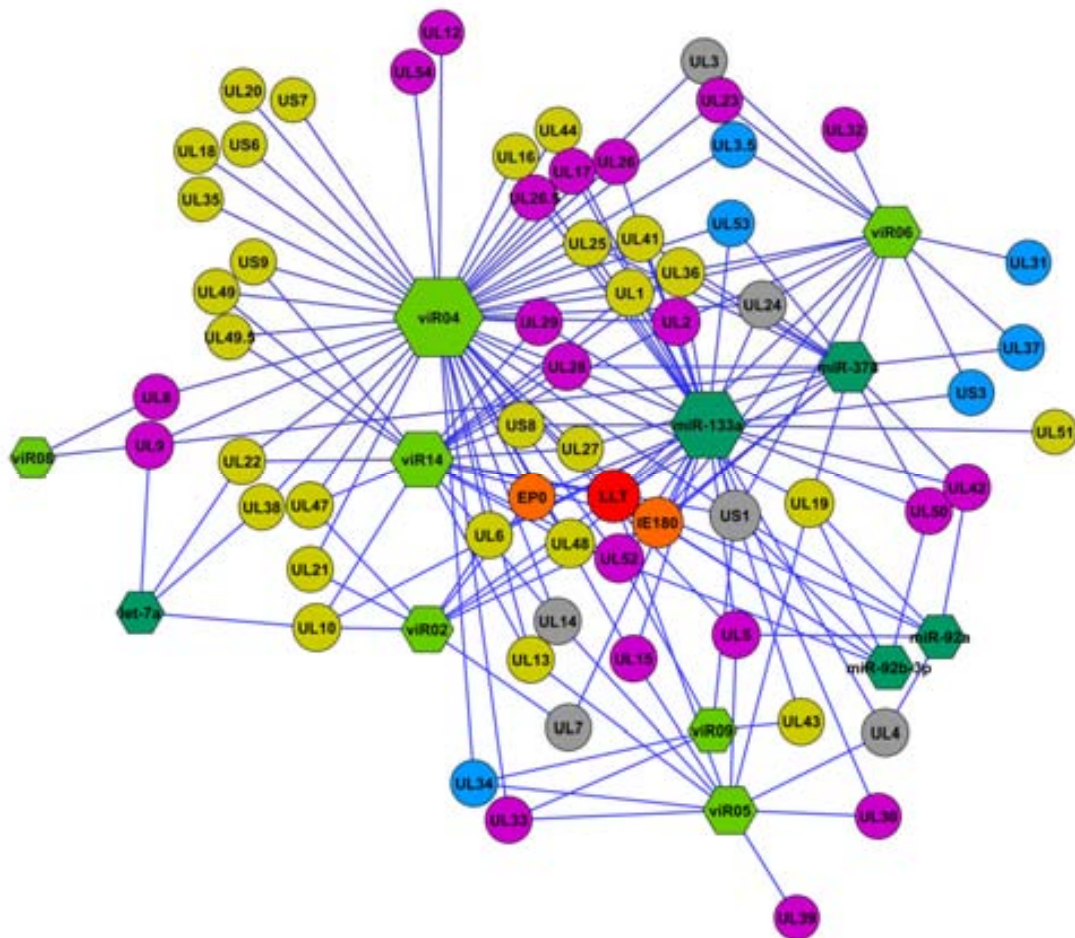


## FIGURES

**Figure 1. Results in viR validation through RT-qPCR for NIA-3 group from *in vivo* animal infection.** Each line represents one sample and each column describes each studied viR. OB: Olfactory Bulb; TG: Trigeminal Ganglia. Colour determines either the viR could be detected or not in the sample. Green reflects the detection of the viR in the sample in quantifiable parameters. Yellow defines the detection of the viR without quantifiable parameters. Red determines that the viR was not detected in the sample. Parameters were considered quantifiable when: (1) viR detection is achieved in all reactions done per sample (4 RT-qPCR reactions, corresponding to 2 different RT per duplicate) and (2) a minimum of 3 out of the 4 RT-qPCR reactions have a cycle threshold (Ct) under 35.

	viR02	viR04	viR05	viR06	viR08	viR09	viR11	viR14
31OB	Green	Red	Red	Yellow	Red	Green	Red	Red
34OB	Green	Red	Red	Red	Red	Red	Red	Red
35OB	Red	Red	Red	Red	Red	Red	Red	Red
37OB	Green	Red	Red	Red	Red	Yellow	Red	Red
38OB	Green	Green	Green	Green	Green	Green	Green	Green
39OB	Green	Red	Red	Yellow	Red	Yellow	Red	Red
41OB	Green	Green	Green	Green	Green	Green	Green	Green
45OB	Green	Red	Red	Red	Red	Red	Red	Red
31TG	Green	Red	Red	Green	Red	Green	Red	Red
34TG	Yellow	Red	Red	Yellow	Red	Red	Red	Red
35TG	Green	Red	Red	Yellow	Red	Yellow	Red	Red
38TG	Green	Yellow	Red	Green	Green	Green	Green	Green
39TG	Green	Red	Red	Yellow	Yellow	Green	Red	Red
45TG	Red	Red	Red	Red	Red	Red	Red	Red

**Figure 2. Regulatory gene network between SHV-1 genes and described viRs and most abundant DE host miRNAs in infected samples.** Node shape represents either is a viral miRNA (hexagon, bright green), a host miRNA (hexagon, dark green) or a viral gene (circular). Node size simulates the number of interactions, which is directly proportional. Viral gene function is represented by colour: yellow means structural function (virion envelope, tegument or capsid proteins); purple means regulatory function; blue represents viral egress function and grey means unknown function. LLT transcript is marked in RED and *EPO* and *IE180* viral activators are marked in orange.



## TABLES

Table 1. Summary of employed samples in the study.

Sample	Infection	Group	Approach	PI <sup>1</sup> sample taken time	Sequenced by Ion Torrent	RT-qPCR validation
CC12	<i>In vitro</i>	Mock-infected	Cell culture	12 hours	Yes	Yes
CC24	<i>In vitro</i>	Mock-infected	Cell culture	24 hours	Yes	Yes
CC30	<i>In vitro</i>	Mock-infected	Cell culture	30 hours	Yes	Yes
NIA12	<i>In vitro</i>	NIA-3 infected	Cell culture	12 hours	Yes	Yes
NIA24	<i>In vitro</i>	NIA-3 infected	Cell culture	24 hours	Yes	Yes
NIA30	<i>In vitro</i>	NIA-3 infected	Cell culture	30 hours	Yes	Yes
BEG12	<i>In vitro</i>	Begonia infected	Cell culture	12 hours	Yes	Yes
BEG24	<i>In vitro</i>	Begonia infected	Cell culture	24 hours	Yes	Yes
BEG30	<i>In vitro</i>	Begonia infected	Cell culture	30 hours	Yes	Yes
32BO	<i>In vivo</i>	Mock-infected	OB	5 days	Yes	Yes
37BO	<i>In vivo</i>	NIA-3 infected	OB	4 days	Yes	Yes
41BO	<i>In vivo</i>	NIA-3 infected	OB	6 days	Yes	Yes
43BO	<i>In vivo</i>	Begonia infected	OB	7 days	Yes	Yes
45BO	<i>In vivo</i>	NIA-3 infected	OB	7 days	Yes	Yes
48BO	<i>In vivo</i>	Begonia infected	OB	7 days	Yes	Yes
32TG	<i>In vivo</i>	Mock-infected	TG	5 days	Yes	Yes
37TG	<i>In vivo</i>	NIA-3 infected	TG	4 days	Yes	No
41TG	<i>In vivo</i>	NIA-3 infected	TG	6 days	Yes	No
43TG	<i>In vivo</i>	Begonia infected	TG	7 days	Yes	Yes
45TG	<i>In vivo</i>	NIA-3 infected	TG	7 days	Yes	Yes
48TG	<i>In vivo</i>	Begonia infected	TG	7 days	Yes	Yes
31BO	<i>In vivo</i>	NIA-3 infected	OB	7 days	No	Yes
34BO	<i>In vivo</i>	NIA-3 infected	OB	4 days	No	Yes
35BO	<i>In vivo</i>	NIA-3 infected	OB	7 days	No	Yes
38BO	<i>In vivo</i>	NIA-3 infected	OB	5 days	No	Yes
39BO	<i>In vivo</i>	NIA-3 infected	OB	7 days	No	Yes
42BO	<i>In vivo</i>	Begonia infected	OB	7 days	No	Yes
44BO	<i>In vivo</i>	Begonia infected	OB	7 days	No	Yes
46BO	<i>In vivo</i>	Begonia infected	OB	5 days	No	Yes
47BO	<i>In vivo</i>	Begonia infected	OB	4 days	No	Yes
49BO	<i>In vivo</i>	Mock-infected	OB	7 days	No	Yes
50BO	<i>In vivo</i>	Mock-infected	OB	7 days	No	Yes
31TG	<i>In vivo</i>	NIA-3 infected	TG	7 days	No	Yes
34TG	<i>In vivo</i>	NIA-3 infected	TG	4 days	No	Yes
35TG	<i>In vivo</i>	NIA-3 infected	TG	7 days	No	Yes
36TG	<i>In vivo</i>	Mock-infected	TG	4 days	No	Yes
38TG	<i>In vivo</i>	NIA-3 infected	TG	5 days	No	Yes
39TG	<i>In vivo</i>	NIA-3 infected	TG	7 days	No	Yes
44TG	<i>In vivo</i>	Begonia infected	TG	7 days	No	Yes
46TG	<i>In vivo</i>	Begonia infected	TG	5 days	No	Yes
47TG	<i>In vivo</i>	Begonia infected	TG	4 days	No	Yes
49TG	<i>In vivo</i>	Mock-infected	TG	7 days	No	Yes

<sup>1</sup>: Post Infection. OB: Olfactory bulb, TG: Trigeminal ganglia.

Two last columns indicate which samples were used to create the libraries for sequencing by Ion PGM™ sequencer (n=21) and which samples were added later in order to validate viRs through RT-qPCR (n=40).

**Table 2. Putative viral miRNAs (viRs) selected for RT-qPCR validation.**

viR	information	Genome position <sup>1</sup>	Sequence (5'-3')	Length	Copy number	Approach <sup>2</sup>	RT-qPCR validation
viR02	prv-miR-LLT1 <sup>a</sup>	97929-97949	TCTCACCCCTGGGTCCGTCGC	21	2,299	CC+IT	+
viR04	prv-miR-7-5' <sup>b</sup>	99282-99301	CCGCCCCCGGGGGTTGATG	20	27	CC	+
viR05	new viR	99302-99322	GGGATGGGCGCTCGGGGGTGA	21	7	CC	+
viR06	prv-miR-7-3' <sup>b</sup>	99342-99363	ACCACCGTCCCCCTGTCCCTCA	22	6	CC	+
viR08	new viR	99843-99862	TCAAACCTCCTCGTGCCCC	20	57	CC	+
viR09	prv-miR-4 <sup>c</sup> , moR8 <sup>b</sup>	100203-100220	CGGAACCGGGTGAGGCG	18	872	CC+IT	+
viR11	prv-miR-8-3' <sup>b</sup>	100267-100287	CAACCCTTCTGGAGCCCTACC	21	569	CC	+
viR14	new viR	102016-102040	TTCCGCCCGCTCTCCACCGCCTTT	25	4	CC	+

<sup>a</sup>: homology in miRBase v19 (*p-value* <0.001).

<sup>b</sup>: described at (Wu 2012).

<sup>c</sup>: described at (Anselmo 2011).

<sup>1</sup>: Genome Position: start-end.

<sup>2</sup>: CC: Cell Culture (*In vitro* approach); IT: Infected Tissue (*In vivo* approach).

**Table 3. Summary of sequence processing scheme.**

	PK-15 cell line cultures	Animal infection	General
Raw reads obtained	705,846	855,701	1,561,547
Trimmed and non empty reads	500,870	694,944	1,195,814
Counts ranging from 15 to 29 nt	490,848	674,538	1,165,386
Counts aligned to miRBase (unique sequences)	212,519 (2,151)	222,915 (2,841)	435,434 (4,029)
miRNA profile	229	302	361
<i>Sus scrofa</i> miRNAs	109	150	193
Orthologous miRNAs	113	151	161
Viral miRNAs	7	1	7
Counts aligned to SHV-1 genome (unique sequences)	3,948 (47)	31 (5)	3,979 (50)
Putative viral miRNAs (viRs)	14	2	14

**Table 4. The most abundant (CN>100) DE miRNAs between infected groups (NIA and BEG) and mock-infected group (CC) in the *in vivo* and *in vitro* approaches.**

miRNA	Counts	BEG vs. CC	NIA vs. CC	NIA vs. BEG
<b><i>in vitro</i></b>				
miR-92a	5,495	3.37	-1.54	-5.20
miR-125b-5p	4,896	-7.37	-8.36	-1.13
miR-LLT1	3,280	-	-	-4.04
miR-99b-5p	2,220	-3.48	-6.90	-1.99
miR-100	1,891	-5.47	-7.90	-1.44
miR-92b-3p	704	2.18	-2.62	-5.70
miR-2779	512	-1.38	-5.41	-3.91
miR-2887	479	-20.24	-203.79	-10.07
miR-2904	408	-13.18	-21.41	-1.62
miR-125a-5p	346	-8.13	-13.31	-1.64
miR-5109	294	-9.38	-18.18	-1.94
miR-4286	243	-2.60	-16.78	-6.45
miR-30a-5p	226	-7.87	-9.07	-1.15
let-7b-5p	209	-3.64	-5.76	-1.58
miR-26a-5p	206	-5.33	-3.84	1.39
miR-339-5p	176	-4.97	-22.89	-4.60
let-7d-3p	171	-1.36	-8.66	-6.38
miR-19b	144	-6.97	-4.30	1.62
miR-23a-5p	126	-11.83	-27.63	-2.34
let-7i-5p	122	-5.40	-6.29	-1.17
miR-505-5p	117	-15.46	-25.93	-1.68
miR-4454	104	-3.32	-7.51	-2.26
<b><i>in vivo</i></b>				
miR-133a	1,990	-18.70	5.76	107.79
let-7a	953	2.07	9.57	4.62
miR-378	923	1.02	5.18	5.07
miR-451	867	-4.36	-5.38	-1.23
miR-206	512	-	-	648.10
miR-129b	449	9.65	5.31	-1.82
let-7c	432	7.78	8.16	1.05
miR-21-5p	252	12.33	14.07	1.14
miR-1249	247	-4.36	-22.82	-5.23
miR-133b	228	-60.94	1.44	87.89
let-7i-5p	218	3.66	8.28	2.26
miR-137	189	3.25	-1.96	-6.39
miR-193a-5p	179	6.56	12.13	1.85
miR-10b	170	5.11	4.33	-1.18
let-7b-3p	151	-2.09	-7.36	-3.53
miR-24-3p	145	5.10	2.85	-1.79
miR-30a-5p	130	4.27	6.38	1.49
miR-345-5p	119	-4.83	-5.21	-1.08
let-7f	106	5.12	21.29	4.16

**SUPPORTING INFORMATION**

**Table S1. Described 14 viral clusters from high throughput sequencing.**

viR	Sequence (5'-3')	Length	Genome position <sup>1</sup>	isomiRs	Copy number	Approach <sup>2</sup>
viR01	ACAACCCGGAGCGCCGCCGTC	21	63745-63765	1	4	CC
viR02	TCTCACCCCTGGGTCCGTCGC	21	97929-97949	5	2,299	CC+IT
viR03	CTCATCCCGTCAGACCTGCGCC	22	98386-98407	2	12	CC
viR04	CCGCCCCCGGGGGGTTGATG	20	99282-99301	2	27	CC
viR05	GGGATGGGCGCTCGGGGGTGA	21	99302-99322	1	7	CC
viR06	ACCACCGTCCCCCTGTCCCTCA	22	99342-99363	1	6	CC
viR07	CGTACCGACCCGCCTACCAGGC	22	99550-99571	2	13	CC
viR08	TCAAACCTCCTCGTGTCCCC	20	99843-99862	5	57	CC
viR09	CGGAACCGGGTGCAGGCG	18	100203-100220	5	872	CC+IT
viR10	GTGGGGGCGAAGATTGGGTTGGG	23	100221-100243	1	10	CC
viR11	CAACCCTTCTGGAGCCCTACC	21	100267-100287	6	569	CC
viR12	CGAGGAGATGTGGAGGGGTG	20	100377-100396	1	3	CC
viR13	AGGCTGGGAGTGGGGACGGAAGA	23	101974-101996	2	38	CC
viR14	TTCCGCCCGCTCTCCACCGCCTTT	25	102016-102040	1	4	CC

<sup>1</sup>: Genome Position: start-end.

<sup>2</sup>: Approaches where viRs were found. CC: Cell Culture (*In vitro* infection); IT: Infected Tissue (*In vivo* infection).

**Table S2. Primers and viR/miRNA sequences used for the RT-qPCR design.**

viR/miRNA	Sequence (5'-3')	Forward primer (5'-3')	Reverse primer (5'-3')
viR02	TCTCACCCCTGGGTCCGTCGC	AGTCTCACCCCTGGGTC	TCCAGTTTTTTTTTTTTTTTTCGAC
viR04	CCGCCCCCGGGGGGTTGATG	CGCCCCCGGGGGGTT	AGGTCCAGTTTTTTTTTTTTTTTCATC
viR05	GGGATGGGCGCTCGGGGGTGA	AGGGGATGGGCGCTCG	TCCAGTTTTTTTTTTTTTTTCACCC
viR06	ACCACCGTCCCCCTGTCCCTCA	CACCACCGTCCCCCTG	TCCAGTTTTTTTTTTTTTTTGTAGGGA
viR08	TCAAACCTCCTCGTGTCCCC	GCAGTCAAACCTCCTCGTG	TCCAGTTTTTTTTTTTTTTTGGGGA
viR09	CGGAACCGGGTGCAGGCG	GCAGCGGAACCGGGTG	CCAGTTTTTTTTTTTTTTTCGCCTG
viR11	CAACCCTTCTGGAGCCCTACC	GCAGCAACCCTTCTGGAG	GTCCAGTTTTTTTTTTTTTTTGGTAG
viR14	TTCCGCCCGCTCTCCACCGCCTTT	CAGTTCCGCCCGCTCTC	GGTCCAGTTTTTTTTTTTTTTTAAAGG
miR-25	CATTGCACTGTCTCGGTCTGA	CATTGCACTGTCTCGGT	GGTCCAGTTTTTTTTTTTTTTTCAGA
miR-93	CAAAGTGCTGTTCTGTCAGGTAG	GCAAAGTGCTGTTCTGTG	TCCAGTTTTTTTTTTTTTTTCTACCT
miR-106a	AAAAGTGCTTACAGTGCAGGTAGC	GAAAAGTGCTTACAGTGCAG	TCCAGTTTTTTTTTTTTTTTGTCTAC
miR-26a	TTCAAGTAATCCAGGATAGGCT	GCAGTTCAAGTAATCCAGGA	TCCAGTTTTTTTTTTTTTTTAGCCT
miR-17-5p	CAAAGTGCTTACAGTGCAGGTAG	CAAAGTGCTTACAGTGCAG	GGTCCAGTTTTTTTTTTTTTTTCTAC

**Table S3. Described miRNAs in only one approach: cell culture (CC) or infected tissue (IT).**

<b>miRNA</b>	<b>Copy number</b>	<b>isomiRs</b>	<b>approach</b>
miR-99a-5p	11,178	76	IT
miR-124a-3p	3,810	63	IT
miR-10a-5p	3,570	52	CC
miR-133a	1,990	35	IT
miR-218b	1,887	25	IT
miR-9-3p	1,620	32	IT
miR-129a	1,566	35	IT
miR-150	917	15	IT
miR-451	867	11	IT
miR-145-5p	773	28	IT
miR-874-3p	768	23	IT
miR-486	637	15	IT
miR-3065-3p	633	24	IT
miR-409-3p	629	29	IT
miR-9-5p	548	15	IT
miR-204	534	17	IT
miR-206	512	8	IT
miR-181a-5p	500	12	IT
miR-129b	449	7	IT
let-7c	432	12	IT
miR-rR1-5	415	10	CC
miR-409-5p	395	8	IT
miR-329	364	11	IT
miR-551b-3p	356	9	IT
miR-187	321	18	IT
miR-181c-5p	318	11	IT
miR-485-5p	261	10	IT
miR-676-3p	234	8	IT
miR-133b	228	10	IT
miR-1224-5p	226	16	IT
miR-193b-3p	198	11	IT
miR-137	189	8	IT
miR-127	154	19	IT
miR-181a-3p	151	4	IT
miR-1247-5p	120	13	IT
miR-20b	120	5	IT
miR-184	109	5	CC
miR-363	108	7	IT
miR-99a-3p	97	5	IT
miR-299-5p	96	5	IT
miR-154b	91	3	IT
miR-483-3p	89	5	IT
miR-342	84	4	IT
miR-338-3p	82	4	IT
miR-182	81	6	IT
miR-1468	80	3	IT
miR-6240	80	11	CC
miR-15b-3p	79	9	CC
miR-2889	78	9	CC

miR-495-3p	78	5	IT
miR-214-5p	75	5	IT
miR-181c-3p	73	5	IT
miR-92c	65	7	CC
miR-199a-5p	63	5	IT
miR-199a-3p	62	8	IT
miR-325	62	3	IT
miR-758	60	6	IT
miR-670-5p	58	3	IT
miR-2483	56	5	IT
miR-3065-5p	55	2	IT
miR-377-5p	54	5	IT
miR-199b-3p	51	6	IT
miR-1-3p	49	3	IT
miR-504	48	3	IT
miR-502-3p	42	5	IT
miR-105-5p	41	6	IT
miR-502b	40	2	CC
miR-543	39	5	IT
miR-LLT11a	39	2	CC
miR-155-5p	36	6	CC
miR-219-5p	36	4	IT
miR-329a	34	7	IT
miR-1343	33	4	IT
miR-17-3p	28	3	CC
miR-497	28	4	IT
miR-132-3p	24	3	IT
miR-542-5p	24	3	IT
miR-330-5p	23	2	IT
miR-615	23	2	CC
miR-369-3p	21	1	IT
miR-4324	21	4	IT
miR-LLT6	21	3	CC
miR-135a-5p	20	2	IT
miR-143-3p	20	3	IT
miR-132-5p	19	3	IT
miR-211	19	3	IT
miR-25-5p	19	3	CC
miR-326	18	1	IT
miR-381	18	3	IT
miR-138-3p	17	2	IT
miR-592	14	1	IT
miR-124a-5p	13	2	IT
miR-1386	13	2	CC
miR-154c	13	3	IT
miR-LLT2	13	2	CC
miR-196b-5p	12	2	CC
miR-19a	12	1	CC
miR-3195	12	3	CC
miR-3956-3p	12	3	IT
miR-141	11	3	IT
miR-382-5p	11	2	IT



miR-412-3p	11	3	IT
miR-5097	11	3	CC
miR-5105	11	2	CC
miR-376c-3p	10	1	IT
miR-LLT8	10	1	CC
miR-192-3p	9	2	CC
miR-210	9	3	CC
miR-26a-1-3p	9	1	CC
miR-466i-5p	9	3	CC
miR-96-5p	9	2	IT
miR-1298	8	2	IT
miR-215	8	1	CC
miR-31-5p	8	1	CC
miR-885-5p	8	1	IT
miR-122	7	1	IT
miR-181d-5p	7	2	IT
miR-316-3p	7	1	CC
miR-335-5p	7	1	CC
miR-346	7	1	IT
miR-455-3p	7	1	CC
miR-5100	7	2	IT
miR-1458	6	2	CC
miR-146a-5p	6	1	IT
miR-193b-5p	6	1	IT
miR-296-5p	6	2	CC
miR-3074-5p	6	1	CC
miR-33b-3p	6	1	CC
miR-345-3p	6	1	IT
miR-3576	6	1	IT
miR-3957-3p	6	2	IT
miR-7-3p	6	1	IT
miR-935	6	1	IT
miR-101b-5p	5	1	CC
miR-146b	5	1	IT
miR-2310	5	1	CC
miR-26b-3p	5	1	IT
miR-30e-3p	5	1	IT
miR-375	5	1	IT
miR-432-5p	5	1	IT
miR-488-3p	5	1	IT
miR-107-3p	4	1	CC
miR-107-5p	4	1	CC
miR-125b-1-3p	4	1	IT
miR-148a-3p	4	1	IT
miR-219-1-3p	4	1	CC
miR-219-3p	4	1	IT
miR-23c	4	1	CC
miR-301b	4	1	CC
miR-3596	4	1	IT
miR-450a	4	1	IT
miR-489	4	1	IT
miR-656	4	1	IT

miR-93-3p	4	1	IT
miR-103a-5p	3	1	CC
miR-10a-3p	3	1	CC
miR-1246	3	1	CC
miR-1271-5p	3	1	IT
miR-148a-5p	3	1	IT
miR-1-5p	3	1	IT
miR-1912	3	1	IT
miR-194-5p	3	1	CC
miR-222-5p	3	1	CC
miR-22-5p	3	1	CC
miR-2411	3	1	CC
miR-24-5p	3	1	CC
miR-2478	3	1	CC
miR-26a-2-3p	3	1	CC
miR-29b-1-5p	3	1	CC
miR-323	3	1	IT
miR-3587	3	1	IT
miR-369-5p	3	1	IT
miR-370	3	1	IT
miR-380	3	1	IT
miR-384	3	1	IT
miR-410	3	1	IT
miR-411-5p	3	1	IT
miR-4485	3	1	CC
miR-454	3	1	CC
miR-485-3p	3	1	IT
miR-487b	3	1	IT
miR-545-5p	3	1	IT
miR-6243	3	1	CC
miR-652-5p	3	1	CC
miR-660-5p	3	1	CC
miR-664-3p	3	1	IT
miR-668-3p	3	1	IT
miR-874-5p	3	1	IT
miR-885-3p	3	1	IT
miR-98	3	1	IT
miR-LLT9	3	1	CC

---

**Table S4. Complete *in vitro* miRNA profile from sequencing and differential expression analysis through fold change between NIA-3 strain infected group (NIA), Begonia strain infected group (BEG) and mock-infected group (CC).**

miRNA	Copy number	IsomiRs	Fold Change	Fold Change	Fold Change
			BEG vs. CC	NIA vs. CC	NIA vs. BEG
miR-23a-3p	105,318	245	1.26	1.43	1.14
miR-221-3p	18,709	97	2.04	2.74	1.34
miR-200b-3p	16,910	94	1.68	3.13	1.86
miR-23b-3p	9,797	88	1.74	1.65	-1.05
miR-92a	5,495	49	3.37	-1.54	-5.20
miR-125b-5p	4,896	35	-7.37	-8.36	-1.13
miR-378	4,733	56	-1.16	1.27	1.48
miR-10a-5p	3,570	52	-2.04	-1.38	1.49
miR-LLT1	3,280	30	-	-	-4.04
miR-183-5p	2,622	40	-1.06	1.46	1.55
miR-99b-5p	2,220	20	-3.48	-6.90	-1.99
miR-365-3p	1,925	27	-1.27	-2.71	-2.13
miR-100	1,891	19	-5.47	-7.90	-1.44
miR-192-5p	1,679	31	-2.08	-1.33	1.57
miR-21-5p	1,664	25	1.53	1.67	1.09
miR-222-3p	1,520	40	-1.11	-1.04	1.07
miR-29a-5p	1,250	12	-2.44	-1.66	1.47
miR-1983	1,247	41	1.85	3.01	1.63
miR-429	1,179	17	1.42	1.67	1.18
miR-30d	970	27	-3.64	-2.93	1.24
miR-25-3p	920	10	-1.47	-1.98	-1.34
miR-10b	918	33	-2.73	-1.21	2.26
miR-139-5p	852	25	-1.63	-1.79	-1.10
miR-193a-5p	777	15	-2.97	-1.78	1.67
miR-200a-3p	755	10	-1.21	1.00	1.22
miR-92b-3p	704	33	2.18	-2.62	-5.70
miR-186	689	16	-2.90	-2.15	1.35
miR-324-3p	663	34	-1.52	-1.46	1.04
miR-2779	512	27	-1.38	-5.41	-3.91
miR-191-5p	502	18	-2.87	-2.68	1.07
miR-2887	479	13	-20.24	-203.79	-10.07
miR-15b-5p	438	9	-2.95	-1.87	1.58
miR-18a-5p	437	15	-1.95	-2.08	-1.07
miR-34a	425	14	1.32	2.17	1.64
miR-rR1-5	415	10	-4.09	-2.74	1.49
miR-2904	408	15	-13.18	-21.41	-1.62
miR-551a	370	11	-3.28	-4.13	-1.26
miR-20a-5p	366	11	-1.24	1.35	1.68
miR-16	365	13	1.07	1.29	1.21
miR-125a-5p	346	15	-8.13	-13.31	-1.64
miR-27a	345	17	-2.04	-1.81	1.13
miR-5109	294	20	-9.38	-18.18	-1.94
let-7a	277	11	-1.80	-2.60	-1.45
miR-374b-5p	261	4	-1.79	1.09	1.95
miR-29b-3p	257	10	-4.11	-4.21	-1.02

miR-29c-5p	245	18	-2.35	-2.24	1.05
miR-4286	243	7	-2.60	-16.78	-6.45
miR-30a-5p	226	12	-7.87	-9.07	-1.15
miR-24-3p	221	12	-2.50	-1.68	1.48
let-7b-5p	209	16	-3.64	-5.76	-1.58
miR-574-5p	209	11	-3.09	-2.27	1.36
miR-26a-5p	206	11	-5.33	-3.84	1.39
miR-93-5p	203	6	-3.02	-1.62	1.86
miR-22-3p	202	6	-4.05	-3.92	1.03
miR-130b-3p	182	12	-2.98	-2.17	1.37
miR-505-3p	181	10	-2.00	-2.84	-1.42
miR-500a-5p	178	6	3.84	1.28	-3.01
miR-339-5p	176	9	-4.97	-22.89	-4.60
miR-135b-5p	173	8	-1.23	1.11	1.36
miR-151-3p	172	10	1.24	-1.06	-1.32
let-7d-3p	171	9	-1.36	-8.66	-6.38
miR-320	171	16	-2.40	-2.93	-1.22
miR-30e-5p	164	9	-3.74	-3.40	1.10
miR-532-5p	164	8	1.10	-1.88	-2.06
miR-29a-3p	160	5	-1.36	-1.53	-1.13
miR-374a-5p	155	5	-3.00	-1.38	2.17
miR-18a-3p	147	11	-1.28	-2.00	-1.56
miR-19b	144	4	-6.97	-4.30	1.62
miR-532-3p	144	6	-1.49	-3.13	-2.09
let-7f	134	6	1.33	1.09	-1.21
miR-23a-5p	126	6	-11.83	-27.63	-2.34
let-7i-5p	122	11	-5.40	-6.29	-1.17
miR-652-3p	122	6	1.54	1.42	-1.09
miR-505-5p	117	9	-15.46	-25.93	-1.68
miR-17-5p	110	6	-1.14	-1.05	1.09
miR-184	109	5	-4.25	-1.86	2.28
let-7e	105	6	-1.74	-1.36	1.28
miR-4454	104	11	-3.32	-7.51	-2.26
miR-361-5p	102	9	2.83	3.16	1.12
miR-2898	92	6	-3.32	-18.69	-5.63
miR-193a-3p	89	5	-3.06	-14.68	-4.80
miR-30b-5p	85	3	9.19	4.61	-1.99
miR-4497	82	9	-3.63	-13.83	-3.81
miR-6240	80	11	4.02	-7.90	-31.73
miR-15b-3p	79	9	1.48	2.41	1.63
miR-2889	78	9	-8.55	-64.68	-7.57
miR-769-5p	71	6	-1.24	1.06	1.31
miR-339-3p	68	6	-3.26	-3.02	1.08
miR-423-3p	68	10	-2.80	-1.87	1.50
miR-423-5p	67	10	-3.49	-7.91	-2.27
miR-6529	66	3	-4.41	-2.34	1.89
miR-92c	65	7	-1.31	-4.37	-3.32
miR-128	62	6	-2.35	-1.59	1.48
miR-324-5p	59	4	-7.55	-6.65	1.14
miR-335-3p	59	6	2.80	2.05	-1.36
miR-500a-3p	59	8	-1.47	1.62	2.38
miR-125a-3p	57	11	-5.74	-5.19	1.10

miR-99b-3p	56	5	-5.52	-4.69	1.18
miR-185	54	3	-25.20	-23.46	1.07
let-7d-5p	53	3	-2.25	-2.36	-1.05
miR-503	49	6	-1.19	-1.01	1.17
miR-345-5p	48	1	-1.75	-16.10	-9.23
miR-130b-5p	47	5	-3.04	-1.99	1.53
miR-27b-3p	46	2	-1.17	-1.20	-1.03
let-7b-3p	45	6	-3.59	-5.54	-1.54
miR-103	44	6	-4.66	-8.21	-1.76
miR-664-5p	44	3	1.46	3.06	2.09
miR-148b-3p	41	3	-17.32	-14.09	1.23
miR-191-3p	40	5	-1.86	-1.62	1.15
miR-502b	40	2	-1.87	-1.38	1.35
miR-LLT11a	39	2	-	-	1.47
miR-155-5p	36	6	-1.62	1.47	2.38
miR-301a-3p	35	3	-3.21	-1.70	1.89
miR-484	35	4	-10.32	-40.50	-3.92
miR-152	34	4	-13.19	-3.91	3.37
miR-106b-3p	32	4	-3.36	-3.80	-1.13
miR-126-5p	30	1	1.01	1.97	1.95
miR-200a-5p	29	6	1.15	-2.90	-3.34
miR-362-3p	29	2	6.71	6.47	-1.04
miR-425-3p	29	6	-5.92	-11.31	-1.91
miR-1260	28	3	-1.24	-31.05	-24.97
miR-17-3p	28	3	-2.35	-4.09	-1.74
miR-30c-5p	28	2	-5.69	-4.58	1.24
miR-716b	28	5	-14.78	-8.65	1.71
miR-140-3p	27	3	1.65	1.40	-1.18
miR-378a-5p	27	2	-	-	-4.02
miR-28-3p	26	3	-4.68	-7.31	-1.56
miR-126-3p	23	2	4.60	4.96	1.08
miR-362-5p	23	4	-5.90	-1.50	3.93
miR-615	23	2	-4.89	-6.36	-1.30
miR-200c-3p	21	2	-	-	-1.05
miR-LLT6	21	3	-	-	1.09
miR-1306-5p	20	3	-4.41	-7.01	-1.59
miR-25-5p	19	3	-31.70	-15.48	2.05
miR-140-5p	17	2	-11.17	-4.14	2.70
miR-30a-3p	17	3	-2.12	1.01	2.15
miR-6651-5p	17	3	-1.77	4.18	7.39
let-7g	16	3	-22.01	-5.37	4.10
miR-340-5p	16	3	-1.18	-1.53	-1.30
miR-744	16	2	-	-	-1.73
miR-23b-5p	15	3	-11.25	-13.73	-1.22
miR-29c-3p	15	1	-16.76	-4.98	3.37
miR-106b-5p	14	2	-3.92	-1.64	2.39
miR-130a	14	2	-9.91	-2.29	4.34
miR-421-3p	14	1	-8.21	-2.62	3.13
miR-106a	13	2	1.34	-1.06	-1.41
miR-1386	13	2	-11.43	-12.60	-1.10
miR-3184-5p	13	2	1.55	1.14	-1.36
miR-LLT2	13	2	-	-	-6.75

miR-1285	12	2	-	-	-2.78
miR-196b-5p	12	2	-24.73	-6.04	4.10
miR-19a	12	1	-2.93	1.08	3.16
miR-3195	12	3	-40.58	-30.11	1.35
miR-361-3p	12	2	-1.18	1.15	1.36
miR-455-5p	12	2	-2.71	-3.25	-1.20
miR-5097	11	3	1.02	-2.20	-2.25
miR-5105	11	2	-2.95	-	-
miR-101	10	2	-	-1.32	-
miR-2320-3p	10	2	-	-	2.71
miR-LLT8	10	1	-	-	3.09
miR-1249	9	2	4.18	-2.68	-11.20
miR-125b-2-3p	9	2	-1.78	-2.62	-1.47
miR-192-3p	9	2	2.02	-2.63	-5.31
miR-210	9	3	-1.42	2.70	3.84
miR-26a-1-3p	9	1	-1.01	2.32	2.35
miR-331-5p	9	1	-	-	-2.84
miR-466i-5p	9	3	-22.92	-8.54	2.68
miR-1271-3p	8	2	1.70	-2.62	-4.45
miR-151-5p	8	1	-4.25	-5.21	-1.23
miR-183-3p	8	2	-11.14	-	-
miR-200b-5p	8	1	-3.53	-1.09	3.25
miR-215	8	1	-	-	-3.53
miR-26b-5p	8	2	-	-	2.04
miR-31-3p	8	1	1.69	1.54	-1.10
miR-31-5p	8	1	-2.05	-7.83	-3.82
miR-328	8	2	1.67	-2.62	-4.37
miR-499-5p	8	1	-	-	1.36
miR-316-3p	7	1	-	-	-3.57
miR-335-5p	7	1	1.27	1.14	-1.11
miR-425-5p	7	2	-2.03	-1.75	1.16
miR-455-3p	7	1	-8.78	-1.32	6.66
miR-1458	6	2	-2.36	-5.18	-2.20
miR-15a	6	1	-1.01	1.14	1.16
miR-181b	6	1	-	-	2.02
miR-195	6	1	-	-	-
miR-296-5p	6	2	-4.27	-3.97	1.07
miR-3074-5p	6	1	-	-	-2.08
miR-32	6	1	-	-	1.17
miR-33b-3p	6	1	-6.16	-4.51	1.37
miR-490	6	1	-3.53	-2.62	1.35
miR-101b-5p	5	1	-	-10.35	-
miR-2310	5	1	-6.16	-9.02	-1.47
miR-338-5p	5	1	-	-	2.00
miR-3586-3p	5	1	1.13	-	-
miR-450b-5p	5	1	-	-	1.18
miR-7-5p	5	1	-1.28	-	-
miR-760	5	1	-	-	-1.10
miR-107-3p	4	1	-	-	-
miR-107-5p	4	1	-	-	-
miR-219-1-3p	4	1	-3.53	-	-
miR-23c	4	1	-	-	1.62

---

miR-301b	4	1	-3.53	-2.59	1.37
miR-450c-5p	4	1	-	-	-3.64
miR-103a-5p	3	1	1.17	-	-
miR-10a-3p	3	1	-	-	1.32
miR-1246	3	1	-	-2.59	-
miR-1296-5p	3	1	-	-1.32	-
miR-138-5p	3	1	-	-	-
miR-149	3	1	-7.07	-	-
miR-194-5p	3	1	-	-5.18	-
miR-222-5p	3	1	-	-5.18	-
miR-22-5p	3	1	-3.53	-2.68	1.32
miR-2411	3	1	-	-	2.73
miR-24-5p	3	1	1.41	-	-
miR-2478	3	1	1.89	-	-
miR-26a-2-3p	3	1	-	-	-1.47
miR-29b-1-5p	3	1	-	-2.59	-
miR-29b-2-5p	3	1	-	-	-2.55
miR-330-3p	3	1	-	-5.18	-
miR-3613-3p	3	1	-8.78	-6.44	1.37
miR-4485	3	1	-	-	-
miR-454	3	1	-	-	-1.47
miR-539-5p	3	1	-	-	1.35
miR-545-3p	3	1	-	-3.27	-
miR-574-3p	3	1	-1.01	-	-
miR-6243	3	1	-	-	-
miR-652-5p	3	1	-	-	-1.82
miR-660-5p	3	1	-	-5.18	-
miR-LLT9	3	1	-	-	-

---

**Table S5. Complete *in vivo* miRNA profile from sequencing data and differential expression analysis through fold change between NIA-3 strain infected group (NIA), Begonia strain infected group (BEG) and mock-infected group (CC).**

miRNA	Copy number	IsomiRs	Fold Change BEG vs. CC	Fold Change NIA vs. CC	Fold Change NIA vs. BEG
miR-125b-5p	57,400	171	1.13	-1.43	-1.62
miR-23b-3p	21,832	114	-1.35	-1.38	-1.02
miR-200b-3p	17,544	92	-1.48	2.51	3.71
miR-99a-5p	11,178	76	-1.05	-1.45	-1.38
miR-23a-3p	10,690	74	-1.06	1.23	1.30
miR-92a	6,689	62	1.98	4.00	2.02
miR-125a-5p	5,560	84	-1.19	-3.05	-2.57
miR-100	5,044	34	1.37	1.06	-1.30
miR-338-5p	4,766	28	1.45	1.91	1.32
miR-183-5p	4,598	44	-1.41	-1.67	-1.18
miR-200c-3p	4,354	32	-1.38	2.25	3.11
miR-126-5p	3,945	26	1.35	2.75	2.03
miR-124a-3p	3,810	63	-2.83	-1.39	2.03
miR-126-3p	3,251	37	2.09	1.68	-1.25
miR-138-5p	3,017	44	2.51	-1.46	-3.66
miR-139-5p	3,002	27	3.51	2.12	-1.66
miR-539-5p	2,991	34	2.48	-1.56	-3.87
miR-99b-5p	2,900	30	1.77	1.41	-1.25
miR-30d	2,324	40	2.27	2.83	1.25
miR-133a	1,990	35	-18.70	5.76	107.79
miR-218b	1,887	25	3.78	1.49	-2.54
miR-9-3p	1,620	32	1.34	-1.20	-1.61
miR-129a	1,566	35	-3.89	-4.85	-1.25
miR-429	1,475	25	-2.04	2.28	4.64
miR-128	1,432	45	-1.49	-2.83	-1.90
miR-26a-5p	1,374	21	-1.02	2.74	2.79
let-7b-5p	1,271	26	1.66	3.03	1.82
miR-200a-3p	1,182	23	-3.44	-1.60	2.15
miR-92b-3p	1,173	24	1.34	-1.20	-1.60
miR-29c-5p	1,007	23	1.55	1.37	-1.13
miR-25-3p	987	13	1.40	2.46	1.76
let-7a	953	15	2.07	9.57	4.62
miR-324-3p	947	39	1.69	1.36	-1.24
miR-378	923	25	1.02	5.18	5.07
miR-150	917	15	2.39	2.80	1.17
miR-451	867	11	-4.36	-5.38	-1.23
miR-191-5p	795	15	1.63	-1.45	-2.37
miR-145-5p	773	28	-2.70	-2.29	1.18
miR-324-5p	772	9	1.31	-2.56	-3.35
miR-874-3p	768	23	1.60	1.10	-1.46
let-7d-3p	668	17	-1.16	-1.84	-1.59
miR-486	637	15	1.23	1.09	-1.13
miR-3065-3p	633	24	-1.87	-2.15	-1.15
miR-409-3p	629	29	2.42	1.64	-1.47
miR-500a-5p	578	15	2.07	1.17	-1.76



miR-30e-5p	564	18	2.03	2.62	1.29
miR-9-5p	548	15	1.59	3.03	1.91
miR-365-3p	541	14	-4.84	-2.69	1.80
miR-204	534	17	-1.19	-1.40	-1.18
miR-206	512	8	-	-	648.10
miR-181a-5p	500	12	1.38	1.10	-1.26
miR-129b	449	7	9.65	5.31	-1.82
let-7c	432	12	7.78	8.16	1.05
miR-29a-5p	419	12	-1.50	-2.80	-1.86
miR-221-3p	417	10	1.33	-2.32	-3.09
miR-532-3p	412	9	-1.23	-2.84	-2.30
miR-409-5p	395	8	1.29	1.33	1.03
miR-329	364	11	-1.71	-1.50	1.14
miR-551b-3p	356	9	-1.97	-2.22	-1.13
miR-187	321	18	3.11	2.58	-1.21
miR-181c-5p	318	11	-1.46	-1.93	-1.32
miR-149	314	13	-1.71	-1.96	-1.15
miR-27b-3p	286	6	1.07	-1.13	-1.21
miR-186	280	9	-1.01	-1.06	-1.05
let-7d-5p	278	6	1.41	2.08	1.48
miR-320	261	18	1.53	-2.14	-3.26
miR-485-5p	261	10	2.62	1.50	-1.75
miR-21-5p	252	6	12.33	14.07	1.14
miR-1249	247	8	-4.36	-22.82	-5.23
miR-374a-5p	238	5	1.20	1.33	1.11
miR-676-3p	234	8	1.54	2.25	1.46
miR-133b	228	10	-60.94	1.44	87.89
miR-1224-5p	226	16	1.93	-1.24	-2.40
miR-328	219	11	-1.33	-2.35	-1.76
let-7i-5p	218	8	3.66	8.28	2.26
miR-193b-3p	198	11	-3.27	-4.38	-1.34
miR-532-5p	198	7	4.33	2.37	-1.83
miR-137	189	8	3.25	-1.96	-6.39
miR-15b-5p	189	3	-1.03	2.00	2.05
miR-151-3p	185	11	2.03	2.91	1.43
miR-29b-3p	183	6	1.34	1.64	1.22
miR-16	182	8	1.64	1.09	-1.50
miR-664-5p	181	6	3.38	1.19	-2.84
miR-193a-5p	179	5	6.56	12.13	1.85
miR-1983	177	13	1.27	1.06	-1.20
miR-10b	170	12	5.11	4.33	-1.18
miR-127	154	19	-1.05	1.61	1.69
let-7b-3p	151	7	-2.09	-7.36	-3.53
miR-181a-3p	151	4	1.92	3.53	1.84
miR-24-3p	145	9	5.10	2.85	-1.79
miR-423-5p	144	15	4.17	3.36	-1.24
miR-551a	136	7	1.07	-2.46	-2.64
miR-181b	132	9	1.40	-1.47	-2.07
miR-30a-5p	130	6	4.27	6.38	1.49
miR-1247-5p	120	13	4.51	-1.00	-4.52
miR-20b	120	5	1.18	-2.29	-2.69
miR-345-5p	119	4	-4.83	-5.21	-1.08

miR-652-3p	111	4	2.53	3.07	1.21
miR-361-5p	109	7	-2.12	2.03	4.29
miR-363	108	7	-1.08	-2.34	-2.17
let-7f	106	5	5.12	21.29	4.16
miR-34a	97	6	-2.34	-6.56	-2.80
miR-99a-3p	97	5	-5.39	-24.53	-4.55
miR-299-5p	96	5	-4.27	-1.56	2.73
miR-93-5p	92	7	1.46	-2.27	-3.31
miR-154b	91	3	4.86	1.62	-3.01
miR-125b-2-3p	90	13	-2.36	-3.44	-1.45
miR-483-3p	89	5	1.13	-5.00	-5.64
miR-152	84	6	-1.99	1.00	2.00
miR-342	84	4	1.63	3.83	2.35
let-7e	83	5	5.04	13.42	2.66
miR-338-3p	82	4	-1.24	-6.24	-5.04
miR-182	81	6	-	-	3.61
miR-1468	80	3	-1.10	1.20	1.32
miR-22-3p	80	6	-1.02	1.60	1.64
miR-29a-3p	79	4	1.17	1.29	1.10
miR-339-5p	79	7	-3.59	1.23	4.40
miR-769-5p	79	6	1.26	-1.23	-1.54
miR-495-3p	78	5	4.49	4.62	1.03
miR-214-5p	75	5	1.96	-1.15	-2.25
miR-200a-5p	73	5	-4.37	2.15	9.38
miR-181c-3p	73	5	-1.25	1.07	1.33
miR-20a-5p	72	5	1.72	1.49	-1.16
miR-374b-5p	72	4	1.04	-1.19	-1.23
miR-484	68	7	-3.59	-3.48	1.03
miR-4454	67	5	-3.76	-3.85	-1.03
miR-362-5p	64	5	2.48	4.21	1.70
miR-19b	64	6	-	-	1.52
miR-140-5p	63	7	5.15	1.71	-3.01
miR-199a-5p	63	5	-1.25	-2.56	-2.05
miR-199a-3p	62	8	-2.04	-2.53	-1.24
miR-325	62	3	4.18	3.18	-1.31
miR-183-3p	60	8	-1.07	1.15	1.23
miR-758	60	6	2.21	-1.03	-2.28
miR-670-5p	58	3	1.20	1.88	1.56
miR-29c-3p	57	3	-1.09	1.16	1.26
miR-2483	56	5	-1.02	-1.09	-1.07
miR-28-3p	56	6	1.61	1.20	-1.34
miR-3065-5p	55	2	-2.06	-13.24	-6.42
miR-6529	55	3	-1.13	-1.94	-1.71
miR-377-5p	54	5	2.06	2.11	1.02
miR-505-5p	53	3	-	-	-1.71
miR-199b-3p	51	6	6.90	6.55	-1.05
miR-574-3p	51	2	-3.16	-19.75	-6.25
miR-99b-3p	50	5	1.34	-2.70	-3.63
miR-1-3p	49	3	-	-	57.10
miR-4286	49	8	-3.35	-9.19	-2.74
miR-504	48	3	-3.90	-3.97	-1.02
miR-505-3p	47	7	1.10	3.62	3.29

miR-425-3p	46	5	-1.28	-1.81	-1.41
miR-30c-5p	43	4	-1.40	1.45	2.03
miR-502-3p	42	5	6.33	3.38	-1.87
miR-195	42	4	-1.15	2.76	3.16
miR-2898	42	5	-5.03	-31.66	-6.29
miR-105-5p	41	6	1.09	-1.70	-1.85
miR-574-5p	41	6	-	-	-4.13
miR-543	39	5	1.08	-1.66	-1.79
miR-1306-5p	37	4	-4.76	-6.20	-1.30
miR-15a	36	2	3.00	4.80	1.60
miR-219-5p	36	4	-1.55	-15.43	-9.98
miR-106b-3p	35	2	-1.65	-7.36	-4.45
miR-130a	35	5	-2.88	-1.25	2.30
miR-18a-5p	35	5	-1.26	-1.05	1.20
miR-331-5p	35	5	1.29	1.53	1.19
miR-30b-5p	34	1	2.37	7.07	2.98
miR-130b-3p	34	4	3.91	4.69	1.20
miR-140-3p	34	4	-2.25	-2.17	1.04
miR-329a	34	7	-4.70	-2.86	1.64
miR-455-5p	34	3	-1.60	-7.20	-4.50
miR-1343	33	4	-1.19	-5.03	-4.23
miR-135b-5p	32	2	2.43	4.47	1.84
miR-125a-3p	31	3	3.97	3.04	-1.31
miR-101	30	7	2.38	3.48	1.47
miR-27a	30	3	-1.14	1.12	1.28
miR-222-3p	28	4	1.26	-1.64	-2.07
miR-3613-3p	28	3	-4.72	-10.51	-2.23
miR-497	28	4	-7.02	-17.94	-2.56
miR-500a-3p	27	5	4.48	3.05	-1.47
miR-503	27	5	-1.19	-2.54	-2.13
miR-499-5p	26	2	-	-	-
miR-193a-3p	25	2	-15.15	-26.45	-1.75
miR-5109	25	5	-1.44	-3.50	-2.42
miR-151-5p	24	1	1.80	3.23	1.80
miR-378a-5p	24	1	-4.72	-2.19	2.16
miR-542-5p	24	3	-3.88	-15.19	-3.92
miR-132-3p	24	3	-	-	-1.04
miR-2779	23	4	-12.76	-87.64	-6.87
miR-330-5p	23	2	2.03	3.43	1.69
miR-103	22	3	1.62	1.22	-1.32
let-7g	22	2	-	-	2.25
miR-18a-3p	21	4	2.29	-2.19	-5.02
miR-4324	21	4	-	-	-7.81
miR-330-3p	21	3	-1.10	3.90	4.31
miR-490	21	3	-2.40	-3.21	-1.33
miR-760	21	5	-1.07	-1.92	-1.79
miR-26b-5p	21	3	-	-	1.24
miR-369-3p	21	1	-	-	4.52
miR-361-3p	20	4	2.03	-6.46	-13.12
miR-135a-5p	20	2	1.09	-1.06	-1.15
miR-545-3p	20	3	-8.55	-4.41	1.94
miR-143-3p	20	3	-	-	3.33

miR-132-5p	19	3	-6.11	1.03	6.29
miR-211	19	3	3.38	1.49	-2.27
miR-421-3p	19	2	2.06	1.85	-1.11
miR-326	18	1	-4.26	-14.96	-3.51
miR-381	18	3	1.98	1.58	-1.26
miR-138-3p	17	2	-2.43	1.32	3.22
miR-23b-5p	17	4	2.47	1.49	-1.66
miR-106b-5p	17	2	-	-	2.42
miR-185	16	1	-	-	1.30
miR-423-3p	16	3	-	-	1.58
miR-148b-3p	15	2	-1.57	1.77	2.79
miR-4497	15	1	2.19	1.50	-1.46
miR-29b-2-5p	14	3	-	-	-1.09
miR-32	14	1	-	-	-1.00
miR-335-3p	14	3	-	-	3.66
miR-592	14	1	-	-	3.92
miR-124a-5p	13	2	1.29	1.65	1.27
miR-154c	13	3	-1.12	-1.80	-1.61
miR-3956-3p	12	3	-3.46	-1.45	2.39
miR-716b	11	3	-	-	7.38
miR-130b-5p	11	2	-6.44	-2.36	2.73
miR-141	11	3	-	-	2.71
miR-382-5p	11	2	-	-	2.91
miR-412-3p	11	3	-	-	-1.83
miR-376c-3p	10	1	-3.87	-3.49	1.11
miR-425-5p	10	2	-	1.88	-
miR-192-5p	9	2	-	-	1.30
miR-96-5p	9	2	-	-	1.39
miR-LLT1	9	1	-	-	-
miR-1298	8	2	-2.70	-2.54	1.06
miR-301a-3p	8	2	-1.90	-3.13	-1.64
miR-17-5p	8	2	-	-	2.48
miR-3586-3p	8	1	-	-	4.69
miR-744	8	2	-1.47	-13.46	-9.16
miR-885-5p	8	1	-4.20	-1.04	4.04
miR-450b-5p	8	1	-	-	1.07
miR-122	7	1	-	-	1.91
miR-2904	7	2	-1.08	-1.11	-1.03
miR-181d-5p	7	2	-	-	-1.05
miR-5100	7	2	-5.00	-6.31	-1.26
miR-346	7	1	-	-	-3.25
miR-339-3p	6	2	-	-	10.17
miR-3576	6	1	-	-	5.03
miR-7-5p	6	1	-	-	-7.75
miR-146a-5p	6	1	-	-	-2.21
miR-191-3p	6	2	-2.74	-3.13	-1.14
miR-193b-5p	6	1	1.02	-1.13	-1.15
miR-200b-5p	6	2	-	-32.54	-
miR-3184-5p	6	2	-1.42	-	-
miR-345-3p	6	1	-	-	-1.42
miR-935	6	1	-	-	-2.13
miR-3957-3p	6	2	-11.88	-20.42	-1.72

miR-1296-5p	6	2	-	-	-
miR-7-3p	6	1	-	-	-
miR-1271-3p	5	1	-	-	-1.67
miR-30e-3p	5	1	-	-	2.11
miR-432-5p	5	1	-	-	-2.85
miR-450c-5p	5	1	-	-	-1.71
miR-488-3p	5	1	-	-	-2.41
miR-26b-3p	5	1	1.18	-	-
miR-146b	5	1	-	-	-
miR-375	5	1	-	-	-
miR-3596	4	1	-	-	1.24
miR-450a	4	1	-	-	1.64
miR-489	4	1	-	-	-1.08
miR-656	4	1	-	-	-4.45
miR-1260	4	1	1.13	-	-
miR-148a-3p	4	1	-10.31	-13.91	-1.35
miR-219-3p	4	1	-2.10	-	-
miR-30a-3p	4	1	-2.41	-6.51	-2.70
miR-362-3p	4	1	-3.65	-	-
miR-93-3p	4	1	-	-9.69	-
miR-125b-1-3p	4	1	-	-	-
miR-106a	3	1	-	-	-1.25
miR-148a-5p	3	1	-	-	-1.51
miR-2887	3	1	-	-	1.77
miR-340-5p	3	1	-	-	1.84
miR-545-5p	3	1	-	-	2.25
miR-6651-5p	3	1	-	-	1.08
miR-874-5p	3	1	-	-	-3.16
miR-1-5p	3	1	-	-	-
miR-1912	3	1	-	-2.14	-
miR-1271-5p	3	1	-	-	-
miR-1285	3	1	-	-	-
miR-2320-3p	3	1	-	-	-
miR-23a-5p	3	1	-	-	-
miR-31-3p	3	1	-	-	-
miR-370	3	1	-1.89	-	-
miR-323	3	1	-	-	-
miR-384	3	1	-	-6.51	-
miR-3587	3	1	-	-	-
miR-369-5p	3	1	-	-	-
miR-485-3p	3	1	-	-2.14	-
miR-380	3	1	-	-	-
miR-410	3	1	-	-	-
miR-411-5p	3	1	-	-	-
miR-487b	3	1	-	-	-
miR-668-3p	3	1	-	-	-
miR-664-3p	3	1	-	-	-
miR-885-3p	3	1	-	-	-
miR-98	3	1	-	-	-



	<p>ATXN1, ATXN3, ATXN3L, AURKA, AXL, AZIN1, B3GALT2, BACH1, BAI3, BAZ2A, BAZ2B, BCAT2, BCL11A, BCL11B, BCL2L11, BCL9, BICC1, BMPR2, BRMS1L, BSDC1, BSN, BTBD12, BTBD14A, BTG2, C10orf118, C10orf22, C11orf24, C11orf56, C12orf5, C14orf28, C17orf39, C18orf1, C18orf25, C1orf21, C1orf80, C20orf111, C20orf39, C21orf66, C21orf91, C4orf16, C4orf40, C5orf24, C5orf28, C5orf30, C5orf5, C6orf103, C6orf107, C6orf62, C6orf98, CACNA1C, CACNA1H, CACNA1I, CADM1, CALM1, CALM2, CALM3, CALN1, CAMK2A, CAMTA1, CAND1, CASD1, CASK, CASKIN1, CBFA2T3, CBLN4, CCDC113, CCDC131, CCNE2, CCNJ, CCNJL, CCT6A, CD2AP, CD69, CDC27, CDC42, CDC42BPA, CDCA7L, CDH10, CDK5R1, CDK6, CDKN1C, CEP350, CFL2, CHD9, CHGA, CHKA, CHMP7, CHR2, CHR5, CHST1, CHST7, CIC, CLCN5, CLDN11, CLEC16A, CLGN, CLIP4, CLK3, CLTA, CNIH, CNNM4, CNOT2, CNR1, CNTN4, COL11A1, COL12A1, COL19A1, COL1A2, COL27A1, COL5A1, COPS2, CPEB1, CPEB2, CPEB3, CPEB4, CPLX2, CPNE8, CPSF6, CRB1, CREB1, CREB3L2, CREM, CSMD3, CTNNBIP1, CTTNBP2, CUL3, CXADR, CXCL5, CXXC4, CXXC5, DAAM1, DAB2IP, DACT1, DAG1, DBT, DCLK2, DCP2, DCX, DDIT4, DDX3X, DDX3Y, DENND4B, DIAPH2, DIRAS1, DISC1, DKK3, DLGAP2, DMXL1, DNAJB12, DNAJB9, DNMT3A, DOCK5, DOCK9, DPP10, DRGX, DSC2, DSCAML1, DST, DTX2, DUS2L, DUSP1, DUSP10, DUSP5, DUSP5P, DUSP6, DYNLT3, DYRK2, E2F3, EBAG9, EDEM1, EDEM3, EDG1, EDNRB, EGR2, EIF1, EIF2C4, EIF4G2, ELOVL4, ELOVL6, EN2, EOMES, EPC2, EPHA8, EPS8, ERBB2IP, ERC2, ERGIC2, ESCO1, ESRRG, EVI5, EXOC5, EXOSC6, EZH2, FAM110B, FAM120A, FAM126B, FAM135A, FAM46A, FAM70A, FAM76B, FAM91A1, FARP1, FASLG, FBIN1, FBIN2, FBXL11, FBXO28, FBXO33, FBXW7, FCHO2, FHL2, FHL3, FIGN, FKBP1C, FLI1, FMN2, FMR1, FNBP4, FNDC3B, FNIP1, FOXG1, FOXN2, FOXN3, FOXP2, FREM1, FRS2, FRY, FRYL, FRZB, FSTL1, FUNDC1, FXR1, FZD10, FZD6, G3BP2, GAN, GAP43, GATA2, GATA6, GATAD2B, GCLM, GDF11, GFPT2, GIT2, GLCE, GNPDA2, GNS, GOLGA1, GOLGA3, GOLGA4, GOLGA7, GOLGA8A, GOLGA8B, GOLGA8E, GOLGA8F, GOLGA8G, GPATCH8, GPBP1L1, GPM6A, GPR124, GPR137C, GPR158, GPR162, GPR180, GPR85, GRAMD3, GRHL1, GRIA1, GRIA4, GRID2, GRIP2, GRK5, GRM7, HAND1, HAND2, HAS3, HCN2, HDAC9, HECTD2, HECW1, HERC2, HERC3, HERC4, HERPUD2, HHIP, HIPK1, HIPK3, HIVEP1, HMBOX1, HNF1B, HNRNPA0, HNRNPA1, HNRNPU, HNRPH1, HOXA9, HOXB8, HOXC8, HOXD10, ICK, IDH1, IKZF2, IKZF4, IL6ST, ING2, INSIG1, INSL5, IQGAP2, IQWD1, IRS2, ITGA5, ITGA6, ITGAV, ITM2B, ITPR1, JAKMIP1, JARID2, JAZF1, JOSD1, JUN, KBTBD8, KCNA1, KCNC4, KCND2, KCNJ3, KCNK10, KIAA0240, KIAA0329, KIAA0831, KIAA1109, KIAA1128, KIAA1279, KIAA1370, KIAA1409, KIAA1432, KIAA1600, KIAA1632, KIAA1712, KIF1B, KIF3B, KIF5B, KLF12, KLF2, KLF4, KLF5, KLHL14, KLHL29, KLHL3, KPNA5, KRT1, LATS2, LBX1, LCOR, LEPREL2, LHFPL2, LIN28, LIN28B, LIN54, LMBR1L, LMO2, LRP1B, LRRC1, LRRC20, LRRC4, LRRC8B, LUZP1, LYST, MACF1, MACROD2, MAN2A1, MAP2K4, MAP4K5, MAPK8, MARCH4, MARK1, MBOAT2, MCOLN2, MDN1, ME1, MEF2D, MEGF10, MFHAS1, MIA3, MIER3, MITF, MKL2, MLL5, MLLT10, MLSTD2, MMD, MMP16, MOBKL2A, MORC3, MPP1, MRPS25, MRS2L, MSRB3, MTDH, MTF1, MTHFD2, MTMR10, MTMR9, MTPN, MYCBP2, MYH3, MYH9, MYLIP, MYLK, MYO18A, MYO1B, MYO5A, MYT1L, NARG1, NAT12, NAT13, NAV3, NCK2, NCOA1, NCOA3, NCOA6, NECAP1, NEDD4L, NEFH, NEFL, NEFM, NEGR1, NELLF, NEUROD1, NF2, NFAT5, NFIA, NFIB, NFIX, NFYB, NFYC, NHLH2, NHSL1, NICN1, NIPBL, NKX2-4, NLK, NOL7, NOTCH1, NOV, NOVA1, NOX4, NPAL1, NPAS3, NPC1, NPNT, NPTN, NR4A3, NRF1, NRIP1, NRK, NSF, NSMAF, NTRK2, NUP43, NUPL1, NUTF2, OAZ3, ODZ1, OTUD3, OTUD4, OXSR1, PAFAH1B1, PAIP1, PALLD, PALM2, PANK3, PAPD5, PAPOLA, PAX3, PAX9, PCAF, PCDH10, PCDH11X, PCDH11Y, PCDH20, PCDH9, PCGF3, PCMTD1, PCOLCE2, PCTK1, PDCD6IP, PDE10A, PDE4D, PDE8A, PDS5A, PDS5B, PDS52, PDXDC1, PDZD2, PER2, PFN2, PHF15, PHF17, PHF3, PHLPL, PHTF2, PIAS4, PIK3AP1, PIK3R1, PIK3R3, PIP4K2C, PIP5K1C, PIP5K3, PIP3NA, PIP3NC1, PIP3NM2, PLCB4, PLEKHA1, PLEKHA6, PLEKHB2, PLEKHG3, PLEKHM1, POLK, POLS, PPCS, PPP1R12A, PPP1R12C, PPP1R9A, PRDM13, PRDM16, PRKAB2, PRKAR2B, PRKCE, PROX1, PSMD14, PTAR1, PTEN, PTENP1, PTF1A, PTGER4, PTGFR, PTPRD, PTPRK, PTPRO, PVRL1, QKI, QSER1, RAB14, RAB23, RAB7A, RAB8B, RAD21, RAG1, RANBP9, RAP1A, RAP1B, RAP2C, RAPGEF6, RASAL2, RASSF2, RBJ, RBL2, RBM24, RBM35A, RBM9, RBPJ, RECK, REV3L, REXO1, RFX1, RFXDC2, RGL1, RGS17, RGS3, RHPN2, RIMBP2, RIPK5, RIT1, RNF11, RNF141, RNF38, RNF4, RNF44, ROBO1, ROBO2, RPL15, RPS6KA4, RPS6KB1, RRPB1, RREB1, RSBN1, RUND2B, RYR3, SAR1B, SATB1, SATB2, SCN3A, SCUBE3, SDC2, SEC24A, SEC24C, SEMA3A, SERTAD2, SERTAD3, SESN3, SETD5, SETD7, SETDB1, SFMBT2, SFRS2IP, SGK, SGK3, SGMS2, SGPP1, SH3PXD2A, SHOX, SIM1, SIM2, SIRPA, SLC12A2, SLC12A5, SLC17A6, SLC1A2, SLC24A3, SLC25A32, SLC2A14, SLC2A3, SLC32A1, SLC38A2, SLC44A1, SLC5A7, SLC6A1, SLC7A11, SLC9A1, SLC9A2, SLC9A3R2, SLITRK5, SMAD6, SMAD7, SMURF1, SNAP91, SNAPC1, SNF1LK, SNN, SNX2, SOBP, SOCS5, SOCS6, SORCS3, SORL1, SOSTDC1, SOX11, SOX4, SP4, SPEN, SPHK2, SPN, SPOCK2, SPRYD4, SPTBN4, SRPK2, SRPR, SSBP3, SSFA2, STAG2, STAT2, STK39, STRN, STRN3, STYX, SUDS3, SUHW4, SUPV3L1, SUV420H1, SYNE1, SYNJ1, SYT1, TACC2, TAF15, TAGAP, TBC1D12, TBX3, TCF5, TEAD1, TEF, TEX2, TFAP2A, TGIF1, TMCC1, TMCC3, TMEM16C, TMEM16D, TMEM16H, TMEM184B, TMEM87A, TMEPAI, TMF1, TNPO1, TNRC6B, TOB2, TOP1, TP63, TPCN1, TRAF3, TRAK2, TRAM2, TRIM33, TRIM34, TRIM36, TRIM6, TRIO, TSC1, TSGA14, TTC28, TTC9, TULP4, TWF1, TWIST1, UBE2G1, UBE2W, UBE2Z, UBR1, UGP2, UNC84A, UPF2, USF2, USP28, USP44, VAMP2, VASH2, VAX1, VCL, VEZF1, VPS13B, VPS4B, VPS54, WASL, WDFY3, WDR32, WDR42A, WDR81, WNT5A, WWC1, WWP2, XRN1, XYLT2, YIPF4, YTHDF3, YWHAZ, ZBTB34, ZBTB40, ZBTB46, ZBTB8, ZCCHC2, ZDHHC21, ZDHHC3, ZDHHC5, ZEB2, ZFXH3, ZFPM2, ZNF238, ZNF287, ZNF295, ZNF512B, ZNF521, ZNF532, ZNF533, ZNF626, ZNF654, ZNF659, ZNF711, ZNF804A.</p>
miR-133a	<p>ACVR2B, ADCY1, ADIPOQ, ANGPTL4, AZIN1, BAZ2A, BCL2L2, BCORL2, BICC1, BRUNOL4, BTBD14A, BTN2A3, C10orf38, C20orf117, CAPN5, CCDC88B, CCDC88C, CDC2L5, CDH23, CKAP5, CLTA, CMPK, CORO1C, CREB5, DDX3X, DLGAP3, DNM2, DOK1, ELAVL1, EPHA7, FAM46A, FBXL19, FGF1, FOSL2, FOXP2, GDNF, GGTLA1, GLI3, GPM6A, GPR162, GSN, IGF1R, IHPK1, ITGA3, KCND3, KCNIP2, KIAA0323, KIAA0652, KIAA1147, KIAA1602, KIAA1853, KIF1A, KIF3C, KPNA6, KREMEN1, LASP1, LASS2, LEPREL2, LHFPL, LIMD2, MDN1, MGAT3, MLL, MLL5, MMP24, NADK, NAT11, NCOA6, NLGN3, NOS1, NUP153, PAN3, PCDHGA11, PCDHGA12, PCDHGA2, PCDHGA3, PCDHGA6, PCDHGA8, PCDHGB7, PCDHGC3, PCDHGC4, PCDHGC5, PER1, PER2, PIP3NM2, POU2F1, PPP1R12C, PPP1R9B, PPP2CA, PPP2R5D, PRRT2, PTBP1, PTBP2,</p>

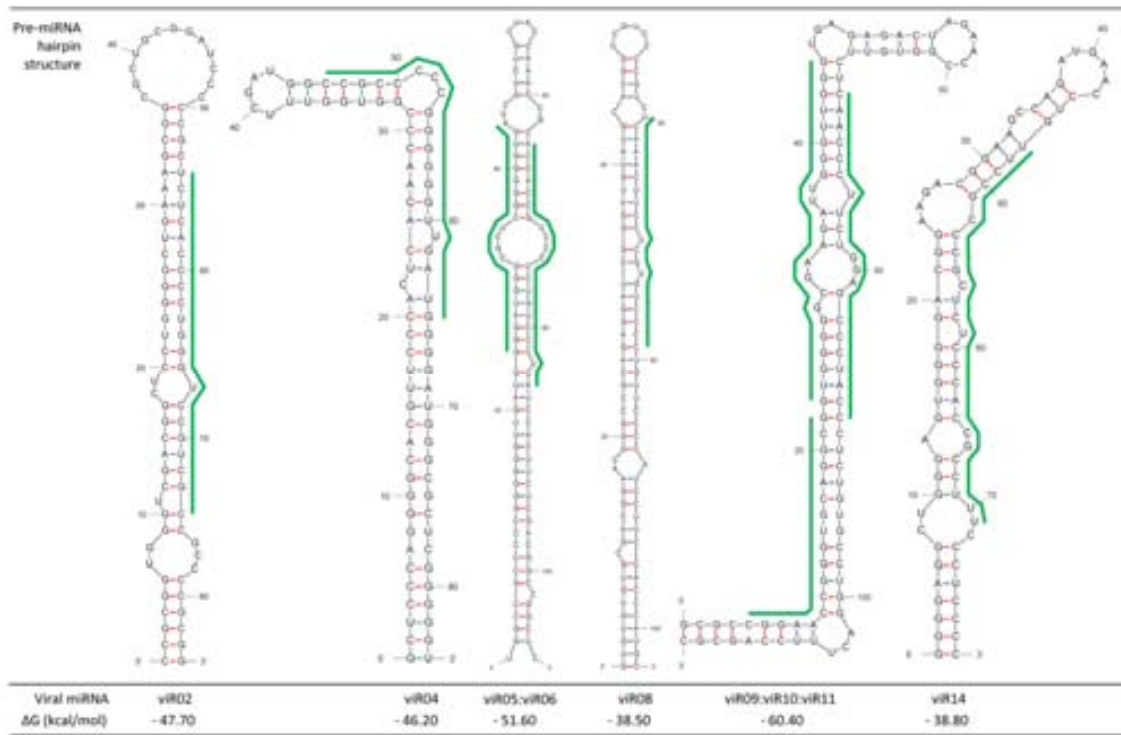
	PTPRD, PTPRK, QKI, RAG1, RIMS4, RNF44, ROD1, SCN2B, SEC14L5, SFXN5, SLC2A12, SLC46A1, SLC6A1, SLC7A8, SMARCD1, SMEK1, SMEK2, SMPD3, SNRK, SNX30, SOBP, SOLH, SREBF2, SRGAP2, STAU1, SVOP, SYDE1, SYT2, TAGLN2, TMC6, TNR, TRAM2, TRO, TTBK1, TTYH3, TUSC5, TXLNA, TXNDC4, UBR4, WIPI2, YPEL2, ZBTB7B, ZHX3, ZNF436.
let-7a	AATK, ABCB9, ABCC10, ABCC13, ABCC5, ABCC8, ABCD4, ABL2, ACP1, ACSL6, ACTR10, ACTR2, ACVR1B, ACVR1C, ACVR2A, ACVR2B, ADAM15, ADAMTS1, ADAMTS14, ADAMTS5, ADAMTS6, ADAMTS8, ADCY9, ADIPOR2, ADRB1, ADRB2, ADRB3, AFF2, AGBL2, AGPAT6, AHCTF1, AHCTF1P, AKAP6, AKT2, ALKBH1, ALPK3, AMMECR1L, AMOT, AMT, ANGPTL2, ANK3, ANKFY1, ANKRA2, ANKRD18A, ANKRD46, ANKRD49, AP1S1, AP4B1, APBB3, APC2, APPBP2, ARG2, ARHGAP12, ARHGAP20, ARHGAP28, ARHGEF15, ARID3A, ARID3B, ARL4D, ARL5A, ARL6IP6, ARMC8, ARRDC4, ASAH3L, ASH1L, ATF7, ATG16L1, ATG9A, ATP2A2, ATP2B1, ATP2B3, ATP2B4, ATP2C1, ATP6V1F, ATP7B, ATPAF1, ATXN1, ATXN2, ATXN7L2, AURKB, AVEN, B3GAT1, B3GNT1, B3GNT7, BACH1, BAH1, BAT3, BCAP29, BCAT1, BCL2L1, BCL7A, BDP1, BEGAIN, BIN3, BMP2, BNC2, BOLL, BPTF, BRD3, BSDC1, BSN, BTBD3, BTF3L4, BZW1, BZW2, C10orf56, C10orf6, C10orf64, C11orf57, C14orf28, C14orf32, C15orf29, C15orf39, C15orf41, C16orf63, C18orf21, C18orf23, C19orf39, C19orf47, C1orf21, C1orf26, C20orf54, C21orf29, C22orf30, C22orf9, C3orf52, C3orf57, C3orf63, C3orf64, C6orf107, C6orf120, C6orf167, C6orf168, C6orf211, C8orf44, C8orf58, C9orf100, C9orf41, C9orf7, CABLES2, CACNA1D, CACNA1E, CACNA1I, CACNG4, CALD1, CALML4, CALU, CAP1, CAPN3, CASKIN1, CASP3, CFA2T3, CBL, CBX2, CCDC100, CCDC113, CCDC4, CCDC76, CCL3, CCL3L1, CCL3L3, CCL7, CCND1, CCND2, CCNF, CCNJ, CCNJL, CCNY, CCR7, CD164, CD200, CD200R1, CD276, CD59, CD86, CDC25A, CDC34, CDC42SE1, CDCA8, CDH1, CDK6, CDKN1A, CDV3, CDYL, CECR6, CEECAM1, CEP110, CEP135, CFL2, CGNL1, CHD4, CHD7, CHD9, CHR, CHRDL2, CHST3, CLASP2, CLCN5, CLDN12, CLP1, CMTM6, CNOT2, CNOT6L, COIL, COL11A1, COL14A1, COL15A1, COL19A1, COL1A1, COL1A2, COL24A1, COL27A1, COL3A1, COL4A1, COL4A2, COL4A3BP, COL4A5, COL4A6, COL5A2, COL9A1, COL9A3, CPA4, CPEB1, CPEB2, CPEB3, CPEB4, CPM, CPSF4, CRB1, CRB2, CRCT1, CREB3L2, CREM, CRTAP, CRY2, CTDSPL2, CTNS, CYB561D1, CYP19A1, CYP2E1, CYP46A1, DAB1, DAGLA, DAPK1, DARS2, DCLRE1B, DCUN1D2, DCUN1D3, DCX, DDEF1, DDI2, DDN, DDO, DDTL, DDX19A, DDX19B, DDX26B, DGAT2L4, DIABLO, DIAPH2, DICER1, DKK3, DLC1, DLGAP1, DLGAP4, DLST, DLSTP, DMD, DNA2L, DNAH2, DNAJC11, DNMBP, DOCK3, DPF2, DPH1, DPYSL3, DST, DTX2, DTX4, DUSP1, DUSP16, DUSP4, DUSP9, DVL3, DYRK1A, DZIP1, E2F2, E2F5, E2F6, EDA, EDEM3, EDN1, EEA1, EEF2K, EFHD2, EGLN2, EGR3, EIF2C4, EIF2S2, EIF4G2, ELF4, ELOVL4, ENOX2, ENPP3, EPB41, EPHA3, EPHA4, EPHB1, ERGIC2, ERO1L, ETNK1, ETNK2, EXOC3L, EZH1, FAM103A1, FAM104A, FAM105B, FAM116A, FAM118A, FAM125B, FAM135A, FAM135B, FAM43A, FAM70A, FAM83G, FAM96A, FARP1, FASLG, FASTK, FBN1, FBXL12, FBXL19, FBXO24, FBXO30, FCHSD1, FGF11, FGF5, FIGN, FLNA, FNDC3A, FNDC3B, FNIP1, FOXN3, FOXP1, FOXP2, FRAS1, FREM2, FRMD4A, FRMD4B, FRS2, FSD1L, FZD4, G3BP1, GAB2, GABBR2, GABPA, GALT, GALE, GALNT1, GALNT2, GALNTL2, GAN, GAS7, GATM, GCNT4, GDAP2, GDF6, GDDP1, GEMIN7, GFOD1, GGA3, GHR, GIPC1, GIYD1, GIYD2, GJA7, GLMN, GLRX, GLT8D3, GNAL, GNAT1, GNG5, GNP1TAB, GNS, GOLGA4, GOLGA7, GOLT1B, GOPC, GPATCH3, GPR137, GPR137C, GPR162, GPR26, GRAMD3, GRB10, GRIK2, GRIN3A, GRPEL2, GSG1L, GTF2I, GYG2, HABP4, HAND1, HAND2, HAS2, HDLBP, HDX, HECTD2, HFM1, HIC2, HIF3A, HK2, HLF, HMGA1, HOOK1, HOXA1, HOXA9, HOXB4, HOXC11, HOXD1, HS2ST1, HSD17B11, HSPA14, HTR1E, HTR4, ICK, ICMT, ICOS, IDE, IDH2, IGF1, IGF1R, IGF2BP1, IGF2BP2, IGF2BP3, IGSF1, IKBKAP, IKBKE, IKZF2, IL10, IL13, IL28RA, IL6, INDOL1, INPP5A, INSR, INTS2, IPO9, IQCB1, IRS2, ISG20L1, ITGA4, ITGB3, ITGB8, ITSN1, JMJD1A, KATNAL1, KCNC4, KCNE3, KCTD10, KCTD17, KCTD21, KIAA0040, KIAA0258, KIAA0319L, KIAA0323, KIAA0329, KIAA0406, KIAA0427, KIAA0515, KIAA0776, KIAA1033, KIAA1147, KIAA1274, KIAA1467, KIAA1539, KIAA2022, KIF21B, KIF2A, KLF9, KLHDC8B, KLHL13, KLHL23, KLHL31, KLHL6, KPNA1, KPNA4, KPNA5, KREMEN1, KTELC1, L2HGDH, LASS2, LBH, LBR, LCOR, LEPREL2, LEPROTL1, LGR4, LIMD1, LIMD2, LIMK2, LIN28, LIN28B, LINGO1, LIPH, LMX1A, LONRF3, LOR, LOXL3, LOXL4, LPGAT1, LRIG2, LRIG3, LRRC17, LRRC20, LRRC59, LRRC8B, LRRFIP1, LSM11, LYPLA3, LYVE1, MACF1, MAN2A2, MAP3K1, MAP3K3, MAP3K7IP2, MAP4K3, MAP4K4, MAPK4, MAPK6, MAPK8, MARCH9, MARS2, MDFI, MECP2, MED28, MED6, MEF2D, MEGF11, MEIS2, MEIS3, MESDC1, MEST, MGA, MGAT3, MGAT4A, MGLL, MIB1, MLL5, MLLT10, MLXIP, MMP11, MNT, MOBKL3, MON2, MRM1, MRS2L, MSN, MTPN, MUSTN1, MUTED, MXD1, MYB, MYCBP, MYCL1, MYCN, MYO1F, MYO5B, MYRIP, NAB1, NAGA, NAP1L1, NARG1, NARG1L, NAT12, NCOA1, NCOA3, NCOA5, NCOA6, NCOR1, NDST2, NDST3, NEFM, NEK3, NEK9, NGFB, NID2, NIPA1, NKD1, NKIRAS2, NLK, NLN, NME4, NME6, NOVA1, NPAL1, NPEPL1, NR4A1, NR6A1, NRAS, NRK, NSMCE2, NT5C2, NUMBL, NXT2, OLFM4, ONECUT2, OPA3, OPRM1, OSBPL3, OSMR, OSTF1, OTOF, OTUD3, OTUD4, P4HA2, PAG1, PAK1, PANX2, PAPP, PARD6B, PAX3, PBX2, PBX3, PCDH19, PCDH20, PCGF3, PCYT1B, PDE1C, PDGFB, PDSS1, PEX11B, PGM2L1, PGRMC1, PHACTR4, PHF8, PI4K2B, PIAS4, PIGA, PIK3IP1, PKIA, PKN2, PLA2G3, PLAGL2, PLCXD3, PLD3, PLD5, PLDN, PLEKHA6, PLEKHG6, PLEKHO1, PLXND1, PNKD, POGZ, POLL, POLR2D, POLR3D, POMT1, POU2F1, POU2F2, PPAPDC2, PPARA, PPARGC1A, PPARGC1B, PPP1R12B, PPP1R15B, PPP1R16B, PPP3CA, PPTC7, PQLC2, PRDM1, PRKAB2, PRMT8, PRPF38B, PRPF39, PRRX1, PRSS22, PRTG, PSCD3, PSD3, PTCH1, PTPN7, PTPRD, PTPRO, PTPRU, PURB, PXDN, PXT1, PYGO2, PYY2, RAB11FIP2, RAB11FIP4, RAB15, RAB22A, RAB3GAP2, RAB40C, RAB8B, RAG1, RAI16, RALB, RALGPS1, RANBP2, RAPGEF6, RASGRP1, RASL10B, RAVER2, RB1, RBM38, RBM9, RBPJ, RCN1, RCS1, RDH10, RDX, REEP1, RFXDC1, RGAG1, RGS16, RIOK3, RNF165, RNF20, RNF216L, RNF38, RNF44, RNF7, RNMT, RORC, RPS6KA3, RRAGD, RRM2, RRP1B, RSPO2, RSPRY1, RTKN, RUFY3, SAMD10, SBK1, SCD, SCN4B, SCN5A, SCRT2, SCUBE3, SCYL3, SDK1, SEC14L1, SEC14L5, SEC16B, SEC24C, SEC31B, SEMA3F, SEMA4C, SEMA4F, SEMA4G, SENP2, SENP5, SERPINB9, SETDB1, SFRS12, SFT2D3, SGCD, SH2B3, SIM2, SIPA1L2, SLC10A2, SLC10A7, SLC12A9, SLC16A10, SLC16A9, SLC17A3, SLC1A4, SLC20A1, SLC25A18, SLC25A24, SLC25A27, SLC25A32, SLC25A4, SLC2A12, SLC30A4, SLC30A7, SLC31A1, SLC31A2, SLC35B1, SLC35D2, SLC37A4, SLC45A4, SLC4A4, SLC4A7, SLC5A6, SLC5A9, SLC6A1, SLC7A6, SLC8A2, SLC9A9, SLC05A1, SMAD2, SMAP1, SMAP1L, SMARCD1, SMARCC1, SMC1A, SMC7, SMUG1, SNAP23, SNX11, SNX16, SNX30, SNX6, SOCS1, SOCS4, SP8, SPATA2, SPEG, SPIRE1, SPOCD1, SPRY4, SPTBN4, SREBF2, SRGAP3, SSH1, ST7L, STAB2, STARD13,



	STARD3NL, STEAP3, STK24, STK40, STRBP, STX3, STXBP5, STYX, SUB1, SUCLG2, SUHW2, SULF1, SULF2, SULT1A3, SULT1A4, SURF4, SYT1, SYT11, SYT14, SYT2, SYT7, TACC3, TADA2L, TAF5, TAF9B, TARBP2, TAT, TBKBP1, TBX5, TEAD3, TEX261, TGDS, TGFBR1, TGFBR3, THBS1, THOC2, THRA, THRSP, TIMM17B, TMC7, TMED5, TMEM110, TMEM135, TMEM143, TMEM164, TMEM2, TMEM65, TMPRSS11F, TMTC3, TNFAIP3, TNFRSF1B, TNFSF9, TOB2, TOR1AIP2, TP53, TPP1, TRABD, TRAK2, TRAM2, TRHDE, TRIB1, TRIB2, TRIM33, TRIM41, TRIM67, TRIM71, TRPM6, TSC1, TSC22D2, TSPAN18, TSPAN2, TTC31, TTC9C, TTL, TLL4, TLL6, TUSC2, TXLNA, TXNDC13, TXNDC5, UBE1, UBE2G2, UBN1, UBXD2, UFM1, UGCGL1, UHRF1, UHRF2, ULK2, UNC5A, USP12, USP21, USP24, USP32, USP38, USP44, USP47, USP49, USP6, UTRN, VANGL2, VASH2, VAV3, VCIPI1, VGLL3, VPS25, VSNL1, WAPAL, WARS2, WASL, WDFY3, WDFY4, WDR26, WDR37, WDR42A, WIPI2, WNT1, XK, XKR8, XRN1, XYLT1, XYLT2, YAF2, YAP1, YIPF1, YOD1, YPEL2, YTHDF3, ZBTB10, ZBTB39, ZBTB5, ZC3H3, ZCCHC3, ZCCHC5, ZFYVE16, ZFYVE26, ZMAT1, ZNF10, ZNF197, ZNF200, ZNF248, ZNF275, ZNF282, ZNF318, ZNF341, ZNF343, ZNF354A, ZNF354B, ZNF362, ZNF462, ZNF473, ZNF512, ZNF512B, ZNF518, ZNF566, ZNF583, ZNF641, ZNF644, ZNF689, ZNF710, ZNF740, ZNF784, ZNF81, ZPLD1, ZRF1, ZSWIM4, ZSWIM5.
miR-378	ADAR, BANP, C1orf132, C20orf91, C9orf5, CHRNA4, CRLS1, E2F2, EIF2C1, FANCA, FBXO40, FMNL3, GNA13, IPO9, KCNAB2, KLK4, MLL, OTUB2, PI4K2A, PIPNA, POU2F1, PTPRT, SLC38A1, TMCC3, UHRF1, VANGL1, ZNF134, ZNF776.
miR-206	ABC7, ABHD2, ADAM12, ADAR, ADCY1, ADPGK, AKAP11, AKAP2, PALM2, AMOT, ANKRD29, ANKRD38, ANXA2, ANXA2P3, ANXA4, AP1G1, API5, ARCN1, ARF3, ARFIP1, ARHGEF18, ARL10, ARSD, ASAH3L, ASH2L, ASXL2, ATF2, ATXN7, AZIN1, BACH2, BAG4, BAIAP2, BCL11A, BCL11B, BCL7A, BDNF, BICD1, BSCL2, BSN, BZRAP1, C10orf26, C10orf68, CCDC7, C11orf61, C17orf62, RNF165, C18orf23, C1orf96, C20orf112, C4orf34, C6orf120, C6orf89, C7orf23, C9orf82, CAND1, CAPRIN1, CBL, CCDC141, CCND2, CD164, CD28, CD2AP, CDC42, CDC42SE1, CDK6, CDON, CEBPZ, CENTB2, CLCN3, CLTC, CNN3, CNTNAP2, COL19A1, COL4A3, COL4A3BP, CORO1C, CPEB1, CPLX2, CREB5, CREBL2, CSNK1G1, CTBP2, CTTNBP2NL, CXorf23, CYP2U1, DCP2, DDX5, DENND2C, DGKH, DICER1, DLG4, DMRT2, E2F5, EDN1, EFN2, EIF2C1, EIF4E, EIF4G3, ELOF1, ELOVL6, EPB41L1, ERC1, ETS1, FAM107B, FAM46C, FAM63B, FAM91A1, FBXO33, FN1, FNBP1L, FNDC3A, FNDC3B, FOSB, FOXP1, FRAS1, FRMD4A, FRS2, FUBP1, G3BP1, G6PD, GAN, GARNL1, GAS2L1, GCH1, GIT1, GJA1, GK5, GLCCI1, GLIS2, GNE, GNPDA2, GOLPH3, GPD2, GPR137C, GPR158, GPR6, GPR85, H3F3A, H3F3B, HACE1, HDAC4, HEYL, HIAT1, HIC2, HIVEP1, HIVEP3, HMBBOX1, HMG2L1, HMGCR, HMG1, HNRNP1, HNRPA3, HOOK1, HOXB4, HS3ST3B1, HSPD1, IGF1, IPO9, ITGB1BP1, JARID2, JOSD1, KCND3, KCNIP3, KCNJ2, KIAA0323, KIAA0329, KIAA0652, KIAA1045, KIAA1462, KIAA2022, KIF2A, KLF13, KLHDC5, KRAS, KTN1, LASP1, LRAT, LRRC8A, LYRM5, MAB21L1, MAL2, MAN1C1, MAP1A, MAP3K1, MAP4K3, MAPK1, MAPKB1, MATR3, MDN1, MED1, MEIS1, MEOX2, MET, MEX3C, MGAT4A, MIER1, MIPOL1, MLL, MLL5, MLLT3, MLLT4, MMD, MMD2, MNT, MOBKL2C, MON2, MPP5, MPP7, MRAS, MXD1, MYLK, MYO15A, NANP, NBEA, NCBP1, NCL, NCOA1, NDRG3, HNRPA3P2, NFAT5, NPAS3, NR3C1, NRP1, NXT2, OAT, OLFML2A, ONECUT2, OSBPL7, OTX2, PAFAH1B1, PAQR5, PAX3, PBEF1, PCDH17, PDCD4, PDGFA, PDIK1L, PFN2, PFTK1, PGD, PHLPL, PLEKHA7, PLEKHQ1, POGK, PPIB, PPP4R2, PRKACB, PRKCE, PRKRIR, PTBP1, PTPLAD1, PTPN1, PTPRC, PTPRG, PTPRK, PTPRT, PVRL3, QKI, RAB43, ISY1, RAB5A, RABEPK, RABGAP1, RASA1, RGS7, RIMS4, RIT2, RNF138, RNF145, RNF44, RSNB1, S100A7A, SDPR, SEC63, SEMA6D, SETBP1, SFRP1, SFRS1, SFRS10, SFRS2IP, SFRS9, SHANK2, SLC10A7, SLC1A2, SLC25A16, SLC25A22, SLC25A25, SLC25A30, SLC29A3, SLC2A13, SLC31A1, SLC35B4, SLC35F1, SLC38A3, SLC39A10, SLC44A1, SLC6A15, SLC7A11, SLC7A2, SLC8A1, SLC8A2, SMAD4, SMARCB1, SMARCC1, SMEK2, SMG7, SNAP25, SNX2, SOS1, SOX5, SOX6, SOX9, SP1, SP2, SPEG, SPHK2, SPRED1, SPTLC3, SRGAP2, STARD7, STC2, STX12, STXBP4, SUHW4, SULF1, SYNJ2, SYT1, TACR1, TAGLN2, TCF7L2, TEX2, TGIF2, THBS1, TIMP3, TLE4, TLOC1, TMCC1, TMEM178, TMEM55B, TMSB4X, TMSL1, TMSL2, TMSL6, TMSL3, TMSL4, TNKS2, TNPO1, TNPO2, TNRC6B, TNS3, TPM3, TPPP, TRAPPC3, TRHDE, TRIM2, TSPYL4, TWF1, UBE2H, UBE4A, UBN1, UBQLN1, UBXD3, UNC50, UST, UTRN, VAMP2, VAMP4, VGLL4, WDR1, WDR48, WEE1, WIPF2, WNK3, WNT3, XPO6, YLPM1, YWHAQ, YWHAZ, ZBTB4, ZC3H7B, ZFP36L2, ZMAT3, ZNF236, ZNF571.

**Figure S1. Predicted viral miRNA folding using MFold software<sup>1</sup>.**

<sup>1</sup>: [40]. Green line points out the miRNA position into the pre-miRNA structure.



**3.3. Estudi dels microRNAs per a ser utilitzats com a gens de referència en estudis de quantificació relativa**

**Determination of Reference microRNAs for Relative Quantification  
in Porcine Tissues**

Timoneda O, Balcells I, Córdoba S, Castelló A, Sánchez A

2012 PLoS ONE 7(9): e44413. doi:10.1371/journal.pone.0044413



# Determination of Reference microRNAs for Relative Quantification in Porcine Tissues

Oriol Timoneda<sup>\*†</sup>, Ingrid Balcells<sup>†</sup>, Sarai Córdoba, Anna Castelló, Armand Sánchez

Departament de Genètica Animal, Centre de Recerca en AgriGenòmica (CRAG), Universitat Autònoma de Barcelona (UAB), Bellaterra, Barcelona, Spain

## Abstract

Relative quantification is the strategy of choice for processing RT-qPCR data in microRNAs (miRNAs) expression studies. Normalisation of relative quantification data is performed by using reference genes. In livestock species, such as pigs, the determination of reference miRNAs and the optimal number of them has not been widely studied. In this study, the stability of ten miRNAs (Ssc-let-7a, Ssc-miR-103, Ssc-miR-17-3p, Hsa-miR-25, Hsa-miR-93, Ssc-miR-106a, Ssc-miR-191, Ssc-miR-16, Ssc-miR-26a and Ssc-miR-17-5p) was investigated by RT-qPCR in different tissues (skeletal muscle, kidney, liver, ovary and uterus) and in different pig breeds (Iberian, Landrace, Large White, Meishan and Vietnamese) as variation factors. Stability values were calculated with geNorm and NormFinder algorithms obtaining high correlation between them ( $r^2 = 0.99$ ). The analyses showed that tissue is an important variability factor in miRNAs expression stability whereas breed is not a determinant factor. All ten miRNAs analysed had good stability values and, therefore, can be used as reference miRNAs. When all tissues were considered, miR-93 was the most stable miRNA. Dividing data set by tissues, let-7a was the most stable in skeletal muscle and ovary, miR-17-5p in kidney, miR-26a in liver and miR-103 in uterus. Moreover, the optimal number of reference miRNAs to be used for proper normalisation data was determined. It is suggested the use of five reference miRNAs (miR-93, miR-25, miR-106a, miR-17-5p and miR-26a) in multi-tissue experimental designs and the use of three reference miRNAs as the optimal number in single tissues studies (let-7a, miR-17-5p and miR-25 in skeletal muscle; miR-17-5p, miR-93 and miR-26a in kidney, miR-26a, miR-103 and let-7a in liver, let-7a, miR-25 and miR-106a in ovary and miR-103, let-7a and miR-93 in uterus). Overall, this study provides valuable information about the porcine reference miRNAs that can be used in order to perform a proper normalisation when relative quantification by RT-qPCR studies is undertaken.

**Citation:** Timoneda O, Balcells I, Córdoba S, Castelló A, Sánchez A (2012) Determination of Reference microRNAs for Relative Quantification in Porcine Tissues. PLoS ONE 7(9): e44413. doi:10.1371/journal.pone.0044413

**Editor:** Atsushi Asakura, University of Minnesota Medical School, United States of America

**Received:** April 16, 2012; **Accepted:** August 2, 2012; **Published:** September 10, 2012

**Copyright:** © 2012 Timoneda et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the projects AGL2007-66371-C02-01 and AGL2010-22358-C02-01 and by the Consolider-Ingenio 2010 program (CSD2007-00036) from Ministerio de Ciencia e Innovación. OT is recipient of PhD fellowship (Programa de Formación al Personal Investigador - FPI) from the Spanish government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: oriol.timoneda@uab.cat

† These authors contributed equally to this work.

## Introduction

MicroRNAs (miRNAs) are small non-coding RNAs involved in gene expression regulation at the post-transcriptional level in animals, plants and viruses [1–3]. They participate in a wide range of biological processes where they play important roles. Largely, their role involves blocking protein translation and/or inducing mRNA degradation [4]. Moreover, miRNA expression has been associated with different pathological processes, such as cancer, neurological disorders, inflammatory pathologies and cardiovascular diseases [5–8]. In some of these pathologies, it has been suggested that miRNAs can be used as biomarkers to develop new diagnostic tools [9,10]. Therefore, it is very important to measure the miRNA expression with high accuracy.

Northern blot has been widely used for determining and measuring miRNA expression [11]. However, latest approaches such as DNA chips (microarrays), high-throughput sequencing (HTS) and reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) are also commonly used [12,13]. HTS and microarrays are used to determine miRNA expression at a genome-wide level whereas RT-qPCR is used to measure the expression of a specific miRNA [14,15]. Furthermore, RT-qPCR

is used to validate expression studies done by microarrays and by HTS due to its high sensitivity and reproducibility [16]. Thus, RT-qPCR has become an important method to assess miRNA expression.

One of the most extensive strategies used to evaluate and compare RT-qPCR data is relative quantification [17]. This methodology normalises the expression of the genes of interest by using one or more genes, called reference genes, of which expression is stable. Data normalisation is necessary to control variables like equal mass loading which can introduce false differences in expression and can perform some experimental bias in the results. Moreover, it is essential to control other variation factors such as RNA degradation during sample processing, quality differences between samples, initial concentration variation among samples, technical variations like pipetting errors and other factors which can affect accuracy during the technique processing [18]. For this reason it is mandatory to perform a data normalisation strategy to correct these possible biases. Originally, normalisation strategies were performed using only one reference gene. However, this idea has evolved to different normalisation approximations, from using the global mean normalisation method [19] to the robust multiple reference genes normalisation

method [17] where more than one reference gene is used. The reference genes chosen (acting as endogenous controls) must not be affected by experimental parameters and they must show invariant expression to the exposed conditions of the individuals used in the study. Consequently, reference genes, as stable genes, are generally involved in basic cellular processes.

In miRNA expression studies, the most common reference genes used are ribosomal RNAs, such as 5S RNA [20–22] and small nuclear RNAs like RNU6B [10,19–24]. However, the use of miRNAs as reference genes is still not widely used; although it is very important that the references used have the same nature that the study subjects. The reference genes used should have the same length as the molecules of interest in order to assure the same efficiency during RNA isolation and reverse transcription [25]. In this sense, only a few studies have explored the stability of some miRNAs in human tissues [19–21] and the published works are largely related to cancer processes [24–28]. Focusing on miRNA expression studies in livestock species, there are few works using miRNAs as reference control in miRNA expression analysis by relative quantification [29–31] and only one report has deeply analysed the miRNA expression stability in pigs to be used as reference miRNAs [22].

The aim of this study was to analyse the miRNA expression stability in different porcine tissues and breeds. Selected tissues were skeletal muscle (structural tissue), uterus and ovary (reproductive tissues), liver (metabolic tissue) and kidney (excretory tissue). On the other hand, porcine breeds include Iberian (European breed), Meishan and Vietnamese (Asian breeds) and Landrace and Large White (European commercial breeds). The results from this work provide useful information concerning which miRNAs could be effectively used as reference genes in order to measure miRNA expression accurately through RT-qPCR studies.

## Results and Discussion

### Analysis of the Stability of the Reference miRNAs

In accordance with the most stable miRNAs described in the literature [20–22,24,26–28], ten candidate miRNAs (*Ssc-let-7a*, *Ssc-miR-103*, *Ssc-miR-17-3p*, *Hsa-miR-25*, *Hsa-miR-93*, *Ssc-miR-106a*, *Ssc-miR-191*, *Ssc-miR-16*, *Ssc-miR-26a* and *Ssc-miR-17-5p*, Table 1) were selected to study their expression stability in different porcine tissues and breeds. All candidate reference miRNAs were successfully amplified through RT-qPCR, allowing us to perform adequate genetic expression quantification [32]. Efficiencies obtained were high ranging from 90% to 110% and the standard curves correlations were at 0.995 minimum (Table 2).

Firstly, the stability was evaluated taking into account the entire data (all tissues and pig breeds) with *geNorm* [17] and *NormFinder* [33] algorithms resulting from its correlation were in a good agreement ( $r^2=0.99$ , Figure 1) and confirm the robustness of our results. *GeNorm* considers a putative reference gene when the M-value calculated is lower than 1.5 and *NormFinder* establishes a lower stability value indicating a better reference gene. All ten miRNAs evaluated had good stability values ranging from 0.64 to 0.80 in M Values (M, *geNorm*) and from 0.25 to 0.44 in stability values (SV, *NormFinder*). In this way, all analysed miRNAs can be used as reference miRNAs for miRNAs expression studies in pigs.

Although all miRNAs showed good stability values, the most stable miRNA was miR-93, followed by miR-25, miR-106a, miR-17-5p and miR-26a (Figure 1). Interestingly, these results are in accordance with a previous report in human tissues [20]. In

contrast, miR-191 (M = 0.80, SV = 0.44), a common reference miRNA used in several human studies and one of the best reference miRNA in human tissues [20,21], was determined as the least stable miRNA in pigs. Then, this result suggested that miR-191 expression stability depends on the specie studied. Comparing our data with a previous study performed also in pigs [22], there are some discordances. In our study, miR-103 was the second worst ranked whereas it was the best positioned in the study by [22]. Moreover, miR-106a (M = 0.68, SV = 0.31), a well ranked miRNA in our study, was the second least stable miRNA in [22] study. Despite of these discordances, miR-17-5p (M = 0.69, SV = 0.32) was well ranked in both studies and miR-16 (M = 0.78, SV = 0.41) had discrete M and SV values in the two studies. These variations in miRNA stability expression could be explained considering the differences of type and amount of tissues used in each study. The work performed by [22] used a total of 47 tissues where uterus was not included, but it was added in the present study. It is also important to remark that some referenced studies were from human tissues [20,21] and it could also be a source of variation in miRNA stability expression when results are compared. Thus, it is reflected that before performing a RT-qPCR study, the reference genes used must be tested experimentally due to many influencing factors. In this sense, M-values obtained by [22] ranged from 1.0 to 2.3, considerably higher than our M-values that could also be explained by the difference in type and amount of tissues used in both studies.

It is known that the stability of miRNA expression could change when factors like tissue or breed are considered. Thus, the entire data were divided by breeds. The results obtained did not differ greatly compared to when the entire data were analysed. M and SV values were calculated from 0.64 to 0.90 and from 0.22 to 0.53, respectively (data not shown). The most stable miRNAs were still ranging between those most stable in the general study. miR-93 remained the most stable miRNA in Iberian and Meishan breeds, miR-26a in Landrace and Vietnamese breeds and miR-25 in Large White breed. The least stable miRNAs were miR-191 in Iberian and Landrace breeds, miR-16 in Large White and Vietnamese breeds and miR-26a in Meishan breed (data not shown). Overall, our results showed that breed only slightly influences the stability of miRNAs.

The stability of the miRNAs was also evaluated for each of the five tissues analysed. As expected, the stability of the miRNAs varies among tissues (Table 3). M and SV values were from 0.41 to 0.90 and from 0.11 to 0.57, respectively. *Let-7a* was the most stable miRNA in ovary and skeletal muscle, miR-103 in uterus, miR-17-5p in kidney and miR-26a in liver, evidencing the specificity of each tissue developing characteristic biological functions and specific metabolic pathways. Conversely, miR-16 was the least stable miRNA in kidney, uterus and liver, miR-103 in ovary and miR-17-3p in skeletal muscle. Stability values in skeletal muscle were in accordance with the results obtained by [22] in the porcine muscle-type tissue group. Two-way analysis of variance (ANOVA) including breed, tissue and tissue by breed interaction showed that only the tissue had a significant effect on miRNA expression (p-value <0.05) in all reference miRNAs analysed. However, in case of *Ssc-miR-17-3p*, breed and tissue by breed interaction had a significant effect on miRNA expression. These results show that tissue is an important variability factor that affects the stability miRNAs expression. Thus, it is evidenced the necessity of using reference miRNAs according the tissue analysed. However, although the high divergence between breeds originated by pig breeding, results showed that miRNA expression is stable across different breeds.

**Table 1.** Primers and miRNA sequences used for the RT-qPCR design.

miRNA	Sequence (5'-3')	Forward primer (5'-3')	Reverse primer (5'-3')
Ssc-let-7a	TGAGGTAGTAGGTTGTATAGTT	GCAAGTGTAGTAGGTTGT	GGTCCAGTTTTTTTTTTTTTAACTATAC
Ssc-miR-103	AGCAGCATTGTACAGGGCTATGA	AGAGCAGCATTGTACAGG	GGTCCAGTTTTTTTTTTTTTCATAG
Ssc-miR-17-3p	ACTGCAGTGAAGGCACTTGATG	GACTGCAGTGAAGGCA	GTCAGTTTTTTTTTTTTTCTACAAG
Hsa-miR-25	CATTGCACCTGTCTCGGTCTGA	CATTGCACCTGTCTCGGT	GGTCCAGTTTTTTTTTTTTTCTACAAG
Hsa-miR-93	CAAAGTGTCTGTCGTCAGGTAG	GCAAAGTGTCTGTCGTCG	TCCAGTTTTTTTTTTTTTCTACCT
Ssc-miR-106a	AAAAGTGTCTACAGTGCAGGTAGC	GAAAAGTGTCTACAGTGCAG	TCCAGTTTTTTTTTTTTTGTCTAC
Ssc-miR-191	CAACGGAATCCCAAAGCAGCTG	AACGGAATCCCAAAGCA	TCCAGTTTTTTTTTTTTTTCAGC
Ssc-miR-16	TAGCAGCACGTAATATTGGCG	GCAAGTGTACAGTGCAG	CAGTTTTTTTTTTTTTTCGCCAA
Ssc-miR-26a	TTCAAGTAATCCAGGATAGGCT	GCAAGTGTACAGTGCAG	TCCAGTTTTTTTTTTTTTTCAGCT
Ssc-miR-17-5p	CAAAGTGTCTACAGTGCAGGTAG	CAAAGTGTCTACAGTGCAG	GGTCCAGTTTTTTTTTTTTTCTAC

doi:10.1371/journal.pone.0044413.t001

Comparing the stability values obtained in the entire data with those obtained in each tissue group, the M and SV values generally improved when tissues were treated as separated (Figure 2). This is due to the decreasing variability in the sample group analysed (Figure 2). There were some miRNAs which considerably improved their stability values when they were measured in a single tissue, such as Let-7a, miR-26a, miR-103 or miR-17-5p. These four miRNAs had discrete stability values in the entire data, but were the best stable miRNAs when data were divided by tissues. Thus, these miRNAs are very stable in a specific tissue, suggesting that the stability of miRNAs expression varies between tissues, and would be the best option for reference miRNAs if we are interested in an experimental design using only one tissue. However, they would not be the best option in multi-tissue experiments because their stability will decrease.

It is important to take into consideration the physiological status (pre and post-pubertal) of the sows because it is known that hormones could affect the gene expression. However, the ten reference miRNAs tested in this study showed no significant differences in miRNA expression between pre and post-pubertal sows. Then, it seems that the expression of these 10 miRNAs in the studied samples is stable under different hormonal environments.

### Determining the Optimal Number of Reference miRNAs

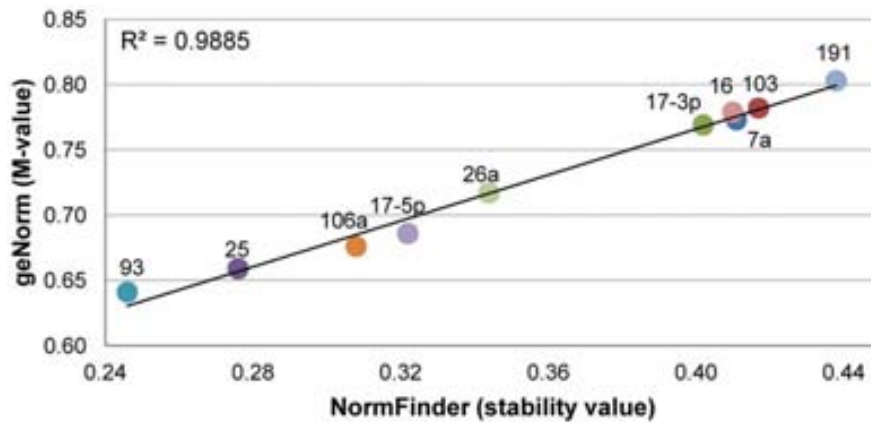
The optimal number and choice of reference miRNAs for qPCR data normalisation must be experimentally determined. Moreover, more than one reference miRNA should be used [32]. In order to determine the optimal number of reference miRNAs needed for a proper correction of RT-qPCR data, the pairwise variation between two sequential normalisation factors containing an increasing number of miRNAs were studied using the geNorm algorithm (V-values, Figure 3). A large variation means that the added gene has a significant effect and should preferably be included for calculation of a reliable normalisation factor [17]. Analysing the entire data, the lowest V-value was obtained using the ten miRNAs studied (Figure 3A). However, use of such a large number of reference miRNAs is unlikely due to experimental requirements and the economical costs. Following geNorm developer recommendations, taking a 0.15 cut-off value on pairwise variation could be enough for a reliable normalisation. In this sense, five reference miRNAs would be necessary for normalisation studies with multiple tissues (V-value = 0.11). Despite of the differences between the studies, these results could be in accordance with [20] and [22], recommending more than one reference miRNA in most situations and three reference miRNAs as the optimal number, respectively.

**Table 2.** Summary of qPCR assays for each reference miRNA studied.

Reference miRNA	Primer conc. (nM each)	cDNA dilution	qPCR efficiency mean*	Std. curve correlation mean*
Ssc-let-7a	125	1/2000	95.52% (2.11%)	0.9991 (0.0005)
Ssc-miR-103	250	1/2000	96.27% (5.15%)	0.9981 (0.0013)
Ssc-miR-17-3p	250	1/200	99.96% (7.25%)	0.9971 (0.0026)
Hsa-miR-25	250	1/2000	97.14% (3.76%)	0.9989 (0.0004)
Hsa-miR-93	200	1/2000	98.10% (3.12%)	0.9973 (0.0012)
Ssc-miR-106a	250	1/2000	99.73% (11.05%)	0.9978 (0.0015)
Ssc-miR-191	250	1/2000	97.45% (3.91%)	0.9978 (0.0009)
Ssc-miR-16	250	1/2000	98.31% (4.98%)	0.9990 (0.0004)
Ssc-miR-26a	250	1/2000	93.35% (0.62%)	0.9991 (0.0005)
Ssc-miR-17-5p	250	1/2000	98.52% (5.00%)	0.9980 (0.0009)

\*The numbers in brackets denote the standard error for the mean values.

doi:10.1371/journal.pone.0044413.t002



**Figure 1. Correlation between M-value (geNorm<sup>1</sup>) and stability value (NormFinder<sup>2</sup>) in the general data set.**<sup>1</sup>: [17], <sup>2</sup>: [33] A very good correlation between the two approximations confirms the robustness and the credibility of our results.  
doi:10.1371/journal.pone.0044413.g001

In studies considering only one tissue, the use of three reference miRNAs would be optimal, taking into account that the V-values using three reference miRNAs were below 0.10 in all tissues (Figure 3A). To contrast the results obtained, the average of stability M-value from geNorm in a stepwise exclusion of the least stable reference miRNA was calculated (Figure 3B). The necessity of including a third reference miRNA was evidenced taking into account the average expression stability variance from two reference miRNAs to three, even reaching five reference miRNAs in studies with multiple tissues in order to minimise the variation in the stability M-value. The high variation on the average expression stability M-values using only two reference miRNAs instead of three was evidenced. This variation became stable when a third or fourth reference miRNA was added. The increased variation in the M-value when a low-stability miRNA was used as an endogenous control was also proved. For example, in the last four miRNAs used in skeletal muscle, showing a high variation every time a reference miRNA was excluded, and the expression stability M-value became stable from the seventh to the third

reference miRNA used. A similar situation happened in kidney tissue. Nevertheless, the expression stability M-value variation in the entire data remained constant using from the tenth reference miRNAs to the fifth, where it started to increase.

In conclusion, this work has evaluated the stability of ten miRNAs in different porcine tissues and breeds showing that they could be used as reference miRNAs. Stability values reflect that tissue is an important variability factor and it must be taken into consideration in the experimental design. It is recommended the use of five reference miRNAs: miR-93, miR-25, miR-106a, miR-17-5p and miR-26a in studies which include multiple tissues. For studies in a specific tissue, the optimal would be the use of three reference miRNAs which is sufficient to obtain a reliable normalisation of data. The most stable reference miRNAs vary between the tissues studied. In kidney it is recommended miR-17-5p, miR-93 and miR-26a. In ovary the best options are Let-7a, miR-25 and miR-106a, while in uterus we recommend to use miR-103, Let-7a and miR-93. If the study is focused in skeletal muscle, we encourage using Let-7a, miR-17-5p and miR-25, but if we are

**Table 3.** Stability values for each microRNA calculated by using geNorm<sup>a</sup> and NormFinder<sup>b</sup> algorithms.

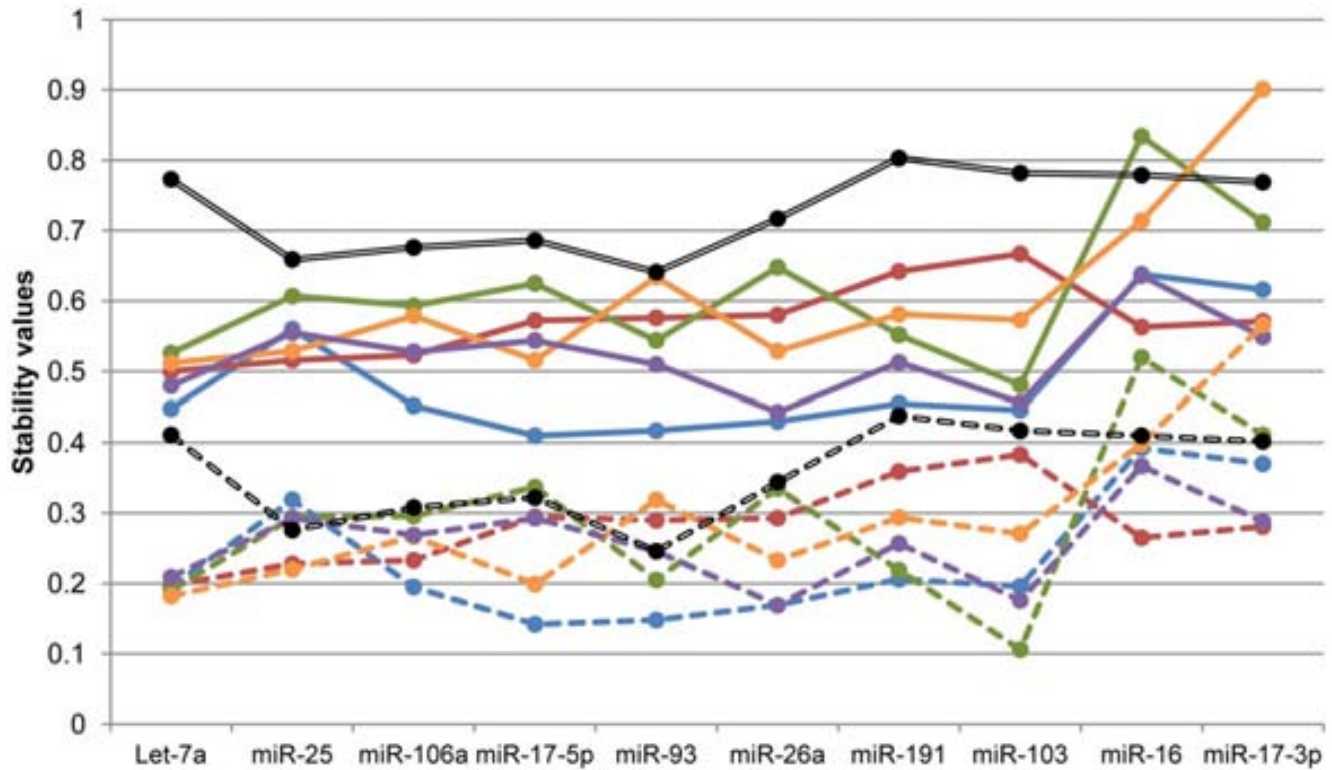
Data set	Algorithm	miR-191	miR-106a	miR-25	miR-93	miR-17-5p	miR-26a	Let-7a	miR-103	miR-16	miR-17-3p	Correlation
Kidney	gN	0.455 <sup>(7)</sup>	0.452 <sup>(6)</sup>	0.560 <sup>(8)</sup>	0.417 <sup>(2)</sup>	<b>0.410<sup>(1)</sup></b>	0.430 <sup>(3)</sup>	0.448 <sup>(5)</sup>	0.446 <sup>(4)</sup>	0.638 <sup>(10)</sup>	0.616 <sup>(9)</sup>	0.9952
	NF	0.206 <sup>(7)</sup>	0.195 <sup>(6)</sup>	0.319 <sup>(8)</sup>	0.148 <sup>(2)</sup>	<b>0.142<sup>(1)</sup></b>	0.169 <sup>(3)</sup>	0.195 <sup>(5)</sup>	0.196 <sup>(4)</sup>	0.392 <sup>(10)</sup>	0.370 <sup>(9)</sup>	
Ovary	gN	0.642 <sup>(9)</sup>	0.523 <sup>(3)</sup>	0.516 <sup>(2)</sup>	0.576 <sup>(7)</sup>	0.572 <sup>(6)</sup>	0.580 <sup>(8)</sup>	<b>0.500<sup>(1)</sup></b>	0.667 <sup>(10)</sup>	0.563 <sup>(4)</sup>	0.571 <sup>(5)</sup>	0.9897
	NF	0.359 <sup>(9)</sup>	0.233 <sup>(3)</sup>	0.228 <sup>(2)</sup>	0.290 <sup>(7)</sup>	0.295 <sup>(6)</sup>	0.293 <sup>(8)</sup>	<b>0.197<sup>(1)</sup></b>	0.383 <sup>(10)</sup>	0.265 <sup>(4)</sup>	0.281 <sup>(5)</sup>	
Uterus	gN	0.552 <sup>(4)</sup>	0.593 <sup>(5)</sup>	0.607 <sup>(6)</sup>	0.544 <sup>(3)</sup>	0.625 <sup>(7)</sup>	0.648 <sup>(8)</sup>	0.526 <sup>(2)</sup>	<b>0.481<sup>(1)</sup></b>	0.834 <sup>(10)</sup>	0.712 <sup>(9)</sup>	0.9714
	NF	0.219 <sup>(4)</sup>	0.296 <sup>(5)</sup>	0.294 <sup>(6)</sup>	0.205 <sup>(3)</sup>	0.337 <sup>(7)</sup>	0.335 <sup>(8)</sup>	0.190 <sup>(2)</sup>	<b>0.106<sup>(1)</sup></b>	0.520 <sup>(10)</sup>	0.411 <sup>(9)</sup>	
Skeletal	gN	0.581 <sup>(7)</sup>	0.579 <sup>(6)</sup>	0.529 <sup>(3)</sup>	0.634 <sup>(8)</sup>	0.516 <sup>(2)</sup>	0.529 <sup>(4)</sup>	<b>0.512<sup>(1)</sup></b>	0.573 <sup>(5)</sup>	0.714 <sup>(9)</sup>	0.901 <sup>(10)</sup>	0.9877
Muscle	NF	0.294 <sup>(7)</sup>	0.267 <sup>(6)</sup>	0.220 <sup>(3)</sup>	0.319 <sup>(8)</sup>	0.199 <sup>(2)</sup>	0.233 <sup>(4)</sup>	<b>0.182<sup>(1)</sup></b>	0.271 <sup>(5)</sup>	0.398 <sup>(9)</sup>	0.566 <sup>(10)</sup>	
Liver	gN	0.513 <sup>(5)</sup>	0.528 <sup>(6)</sup>	0.555 <sup>(9)</sup>	0.510 <sup>(4)</sup>	0.544 <sup>(7)</sup>	<b>0.443<sup>(1)</sup></b>	0.480 <sup>(3)</sup>	0.456 <sup>(2)</sup>	0.636 <sup>(10)</sup>	0.549 <sup>(8)</sup>	0.9825
	NF	0.257 <sup>(5)</sup>	0.269 <sup>(6)</sup>	0.293 <sup>(9)</sup>	0.246 <sup>(4)</sup>	0.293 <sup>(7)</sup>	<b>0.169<sup>(1)</sup></b>	0.209 <sup>(3)</sup>	0.176 <sup>(2)</sup>	0.367 <sup>(10)</sup>	0.288 <sup>(8)</sup>	
General	gN	0.803 <sup>(10)</sup>	0.676 <sup>(3)</sup>	0.659 <sup>(2)</sup>	<b>0.641<sup>(1)</sup></b>	0.686 <sup>(4)</sup>	0.717 <sup>(5)</sup>	0.773 <sup>(7)</sup>	0.782 <sup>(9)</sup>	0.779 <sup>(8)</sup>	0.769 <sup>(6)</sup>	0.9885
	NF	0.438 <sup>(10)</sup>	0.308 <sup>(3)</sup>	0.276 <sup>(2)</sup>	<b>0.246<sup>(1)</sup></b>	0.322 <sup>(4)</sup>	0.344 <sup>(5)</sup>	0.411 <sup>(7)</sup>	0.417 <sup>(9)</sup>	0.410 <sup>(8)</sup>	0.402 <sup>(6)</sup>	

<sup>a</sup>: [17], <sup>b</sup>: [33].

gN: geNorm algorithm, NF: NormFinder algorithm. Correlations between M-value (geNorm) and stability value (NormFinder) are shown. Superscript numbers into brackets show stability values sorted out for each data set group from 1 (most stable) to 10 (less stable). The stability values of the most stable miRNA for each group are marked in bold.

doi:10.1371/journal.pone.0044413.t003





**Figure 2. Stability values of each reference miRNA from geNorm<sup>1</sup> and NormFinder<sup>2</sup> algorithms.** <sup>1</sup>: [17], <sup>2</sup>: [33]. Continuous lines: M-value from geNorm software; dashed lines: stability value from NormFinder software. Black lines: general study, blue lines: kidney, red lines: ovary, green lines: uterus, orange lines: skeletal muscle, purple lines: liver. doi:10.1371/journal.pone.0044413.g002

working with liver, the most stable miRNAs to be used as reference miRNAs are miR-26a, miR-103 and Let-7a.

Overall, this study provides valuable information about the porcine reference miRNAs that can be used in order to perform a proper normalisation when relative quantification studies are undertaken. Further experiments should be made to construct a database with recommended reference miRNAs for each tissue in porcine and also for multiple tissue studies.

## Materials and Methods

### Sample Collection

Samples were collected from five different pig tissues (skeletal muscle, ovary, uterus, kidney and liver). Pigs included in the study came from different breeds: Iberian (IB), Meishan (ME), Vietnamese (VT), Landrace (LD) and Large White (LW). Two samples per breed and per tissue were analysed ( $n = 50$ ). All animals were females in different physiological state (per-pubertal: LD and LW; post-pubertal: IB, VT and ME) and age (6 month old: LD and LW; 1 year old: IB and VT; 2 year old: ME). All samples were taken from slaughterhouse (PRIMAYOR, Mollerussa, Spain) under veterinary supervision. Samples were immediately snap-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until use.

### Candidate Reference miRNAs Selection

Ten miRNAs were selected to be evaluated as reference genes according to the literature: Ssc-let-7a, Ssc-miR-103, Ssc-miR-17-3p, Hsa-miR-25, Hsa-miR-93, Ssc-miR-106a, Ssc-miR-191, Ssc-miR-16, Ssc-miR-26a and Ssc-miR-17-5p [20,22,24,26,28]. Selection was done considering their stability values on the published studies.

### RNA Isolation and cDNA Synthesis

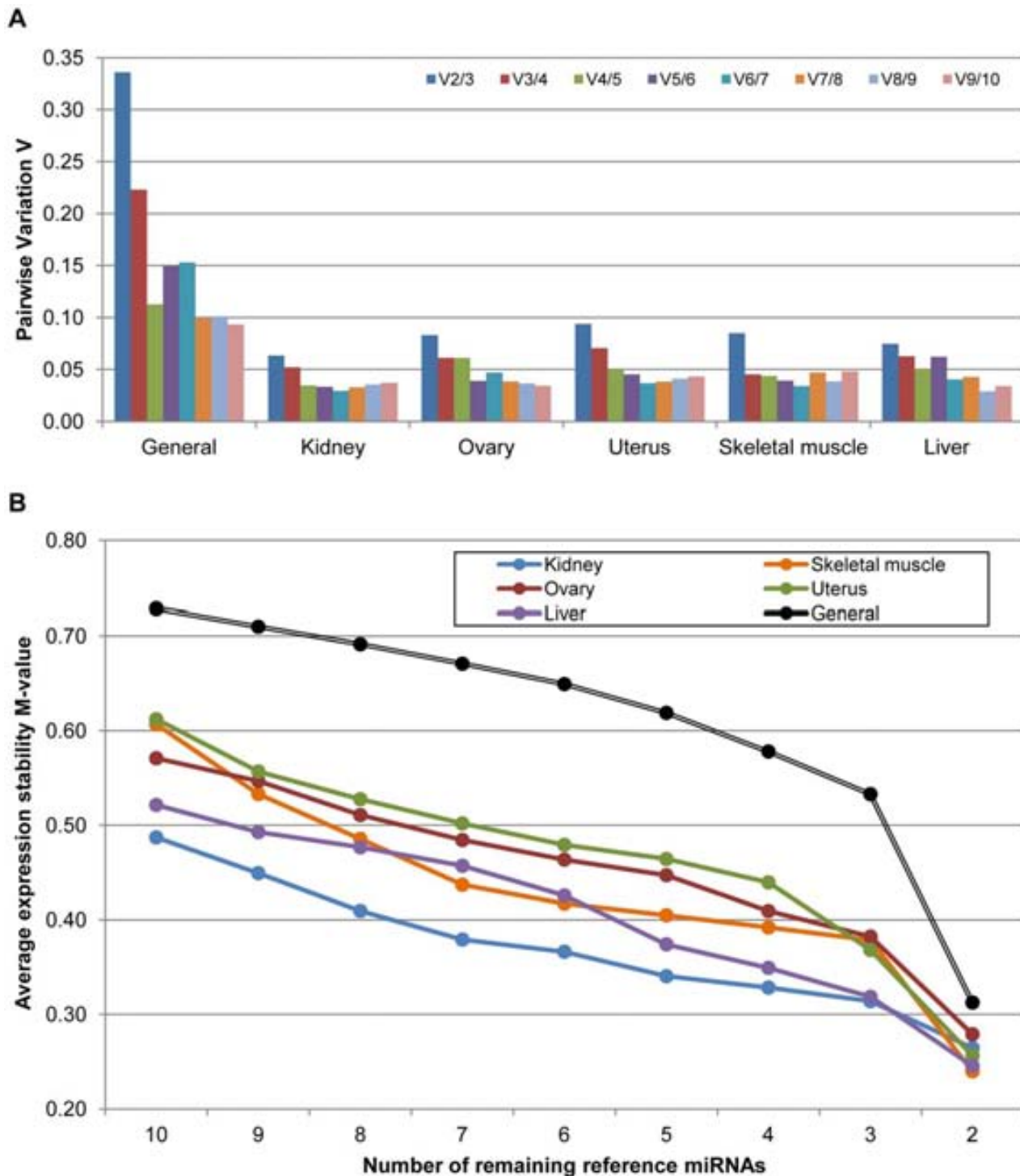
Total RNA was isolated using TRIzol<sup>®</sup> reagent following the manufacturer's recommendations (Invitrogen, Carlsbad, USA). RNA was quantified by absorbance using ND 1000 Nanodrop<sup>®</sup> Spectrophotometer (Thermo Scientific, Wilmington, USA) and checked for integrity by using RNA 600 Nano kits (Agilent Technologies, Santa Clara, USA) on an Agilent 2100 Bioanalyzer.

Reverse transcription (RT) reactions were performed in duplicate using total RNA as previously described [34]. Briefly, 600 ng of total RNA in a final volume of 20  $\mu\text{L}$  including 2  $\mu\text{L}$  of 10x poly(A) polymerase buffer, 0.1 mM of ATP, 0.1 mM of each dNTP, 1  $\mu\text{M}$  of RT-primer, 200 U of M-MuLV reverse transcriptase (New England Biolabs, USA) and 2 U of poly(A) polymerase (New England Biolabs, USA) was incubated at  $42^{\circ}\text{C}$  for 1 hour and  $95^{\circ}\text{C}$  for 5 minutes for enzyme inactivation. Minus reverse transcription (RT) and minus poly(A) polymerase controls for each tissue were included.

### Quantitative Real-time PCR Reaction

Quantitative PCR reactions were performed in a final volume of 20  $\mu\text{L}$  including 10  $\mu\text{L}$  of FastStart Universal SYBR Green Master (Roche, Germany), 250 nM of each primer (with the exception of Ssc-let-7a and Hsa-miR-93, with 125 and 200 nM, respectively) and 5  $\mu\text{L}$  of a 1:2000 dilution of the cDNA (except for Ssc-miR-17-3p where a dilution of 1:200 was used). See Table 2.

Standard curves were generated in order to calculate the RT-qPCR efficiency. All standard curves were done by using 10 fold serial dilutions from a pool of cDNA of all the samples ( $n = 50$ ) and were included per duplicate in all qPCR assays. For Ssc-miR-17-



**Figure 3. Variation of the miRNA stability expression.** (A) Pairwise Variation between two sequential normalisation factors containing an increasing number of reference genes. Data were taken from all samples and dividing it by tissue origin. According to geNorm algorithm [17], a 0.15 cut-off was accepted. For the general study with many tissues it would be recommendable the use of 5 reference miRNAs. In the studies with a single tissue it would be enough to include 3 reference miRNAs. (B) Average expression stability M-values of remaining control genes during stepwise exclusion of the least stable reference miRNA. M-values were calculated by using geNorm algorithm [17]. Black line: general study, blue line: kidney, red line: ovary, green line: uterus, orange line: skeletal muscle, purple line: liver.  
doi:10.1371/journal.pone.0044413.g003

3p, a 2-fold dilution standard curve was used. Reactions were incubated in a 96-well plate at 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min on a 7900HT Real-

Time PCR System with 7900HT SDS v2.4 (Applied Biosystems). DNA primers for each miRNA were designed following the methodology described by [34] (Table 1).

Measurements were performed in duplicate. Minus RT controls, minus poly(A) polymerase controls and no template controls were included. Moreover, melting curve analysis was performed in each assay in order to detect unspecific amplifications.

### Stability Expression Analyses

Quantities from each sample were obtained from the calibration (standard) curve added in each RT-qPCR reaction. Stability of each candidate miRNA was tested using geNorm v.3.5 algorithm [17] and NormFinder algorithm [33]. The geNorm algorithm calculates the gene expression stability M Value for each candidate reference gene based on the average pairwise variation between all studied genes. NormFinder is based on an ANOVA mathematical model and estimates intra- and intergroup variation providing the best stable candidate reference gene and also the best stable pair of them taking into account the subgroups in which data is distributed. In both programs, the lowest stability values indicate the most stably expressed reference genes allowing them to rank according to their expression stability.

RT-qPCR expression data was also analysed by a two-way analysis of variance (ANOVA) with the General Linear Models

procedure of the Statistical Package for the Social Scientists (IBM® SPSS® Statistics 19; IBM Corporation, Armonk, USA). The model used included breed and tissue as fixed factors and breed by tissue interaction. GeNorm v.3.5 software [17] was used to obtain the normalization factor (NF), necessary to normalize each sample quantity obtained from qPCR reaction. Next, fold changes were calculated in relation to the highest sample normalized value for each miRNA. Fold changes were  $\log_2$  transformed and significance threshold was set at  $\alpha < 0.05$ . Student-Newman-Keuls (SNK) and Scheffé tests were used to determine significant differential expression between breed or tissue groups.

### Acknowledgments

The authors thank Peter Busk for the RT-qPCR primer design.

### Author Contributions

Conceived and designed the experiments: IB OT AS. Performed the experiments: SC AC. Analyzed the data: OT IB SC. Wrote the paper: OT. Supervision the present work: AS IB.

### References

- Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116(2): 281–297.
- Krol J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* 11(9): 597–610.
- Glazov EA, Horwood PF, Assavalapsakul W, Kongsuwan K, Mitchell RW, et al. (2010) Characterization of microRNAs encoded by the bovine herpesvirus 1 genome. *J Gen Virol* 91(Pt 1): 32–41.
- Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nat Rev Genet* 12(2): 99–110.
- Farazi TA, Spitzer JL, Morozov P, Tuschl T (2011) miRNAs in human cancer. *J Pathol* 223(2): 102–115.
- Fendler A, Stephan C, Yousef GM, Jung K (2011) MicroRNAs as regulators of signal transduction in urological tumors. *Clin Chem* 57(7): 954–968.
- Esteller M (2011) Non-coding RNAs in human disease. *Nat Rev Genet* 12(12): 861–874.
- Jiang X, Tsitsiou E, Herrick SE, Lindsay MA (2010) MicroRNAs and the regulation of fibrosis. *FEBS J* 277(9): 2015–2021.
- Patel N, Sauter ER (2011) Body fluid micro(mi)RNAs as biomarkers for human cancer. *Journal of Nucleic Acids Investigation* 2: e1.
- Sharkey JW, Antoine DJ, Park BK (2012) Validation of the isolation and quantification of kidney enriched miRNAs for use as biomarkers. *Biomarkers*.
- de Planell-Saguer M, Rodicio MC (2011) Analytical aspects of microRNA in diagnostics: A review. *Anal Chim Acta* 699(2): 134–152.
- Meyer SU, Pfaffl MW, Ulbrich SE (2010) Normalization strategies for microRNA profiling experiments: A ‘normal’ way to a hidden layer of complexity? *Biotechnol Lett* 32(12): 1777–1788.
- Pritchard CC, Cheng HH, Tewari M (2012) MicroRNA profiling: Approaches and considerations. *Nat Rev Genet* 13(5): 358–369.
- Podolska A, Kaczkowski B, Kamp Busk P, Sokilde R, Litman T, et al. (2011) MicroRNA expression profiling of the porcine developing brain. *PLoS One* 6(1): e14494.
- Chen C, Deng B, Qiao M, Zheng R, Chai J, et al. (2012) Solexa sequencing identification of conserved and novel microRNAs in backfat of large white and chinese meishan pigs. *PLoS One* 7(2): e31426.
- Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, et al. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33(20): e179.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3(7): RESEARCH0034.
- Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. *J Mol Endocrinol* 29(1): 23–39.
- Mestdagh P, Van Vlierberghe P, De Weer A, Muth D, Westermann F, et al. (2009) A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biol* 10(6): R64.
- Peltier HJ, Latham GJ (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: Identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* 14(5): 844–852.
- Shen Y, Li Y, Ye F, Wang F, Wan X, et al. (2011) Identification of miR-23a as a novel microRNA normalizer for relative quantification in human uterine cervical tissues. *Exp Mol Med* 43(6): 358–366.
- Gu Y, Li M, Zhang K, Chen L, Jiang AA, et al. (2011) Identification of suitable endogenous control microRNA genes in normal pig tissues. *Anim Sci J* 82(6): 722–728.
- Feng R, Chen X, Yu Y, Su L, Yu B, et al. (2010) miR-126 functions as a tumour suppressor in human gastric cancer. *Cancer Lett* 298(1): 50–63.
- Wotschovsky Z, Meyer HA, Jung M, Fendler A, Wagner I, et al. (2011) Reference genes for the relative quantification of microRNAs in renal cell carcinomas and their metastases. *Anal Biochem* 417(2): 233–241.
- Schaefer A, Jung M, Miller K, Lein M, Kristiansen G, et al. (2010) Suitable reference genes for relative quantification of miRNA expression in prostate cancer. *Exp Mol Med* 42(11): 749–758.
- Chang KH, Mestdagh P, Vandesompele J, Kerin MJ, Miller N (2010) MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. *BMC Cancer* 10: 173.
- Viprey VF, Corrias MV, Burchill SA (2012) Identification of reference microRNAs and suitability of archived hemopoietic samples for robust microRNA expression profiling. *Anal Biochem* 421(2): 566–572.
- Song J, Bai Z, Han W, Zhang J, Meng H, et al. (2012) Identification of suitable reference genes for qPCR analysis of serum microRNA in gastric cancer patients. *Dig Dis Sci* 57(4): 897–904.
- Luo L, Ye L, Liu G, Shao G, Zheng R, et al. (2010) Microarray-based approach identifies differentially expressed microRNAs in porcine sexually immature and mature testes. *PLoS One* 5(8): e11744.
- Curry E, Safranski TJ, Pratt SL (2011) Differential expression of porcine sperm microRNAs and their association with sperm morphology and motility. *Theriogenology* 76(8): 1532–1539.
- Stowe HM, Curry E, Calcatera SM, Krisner RL, Paczkowski M, et al. (2012) Cloning and expression of porcine dicer and the impact of developmental stage and culture conditions on MicroRNA expression in porcine embryos. *Gene* 501(2): 198–205.
- Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55(4): 611–622.
- Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64(15): 5245–5250.
- Balcells I, Cirera S, Busk PK (2011) Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC Biotechnol* 11: 70.



## 4. DISCUSSIÓ GENERAL

---

La present tesi s'ha desenvolupat, principalment, en l'estudi dels microRNAs (miRNAS) com a mecanisme regulador de l'expressió gènica en l'espècie porcina. El porc gaudeix d'una posició privilegiada en la producció de carn i productes derivats a nivell mundial, en la que té un mercat a l'alça i un sistema de producció molt organitzat i actiu. També ha esdevingut en els últims anys un model biomèdic molt important en l'estudi de determinades malalties i també per l'estudi d'estats fisiològics i patològics, degut a la seva proximitat anatòmica, estructural i biològica amb l'espècie humana.

Arran del descobriment dels primers miRNAs ja fa 20 anys (Lee *et al.*, 1993, Pasquinelli *et al.*, 2000, Reinhart *et al.*, 2000), aquestes petites molècules d'RNA no codificant han anat agafant un protagonisme molt rellevant en el camp de l'expressió gènica a nivell post transcripcional, modulant l'expressió dels seus gens diana i esdevenint una peça regulatòria clau amb la seva regulació fina. Tot i així, els miRNAs només varen fer que obrir la porta al descobriment de tot un ventall d'RNAs no codificants de mida petita que exerceixen algun tipus d'interferència en l'expressió de la part codificant del genoma (tenOever, 2013, Costa, 2010). El coneixement sobre el funcionament d'aquest RNA interferent està en continua evolució i, recentment, s'ha descrit una nova peça en tot aquest entramat, els circRNAs, uns RNA circulars que podrien actuar de reguladors dels propis miRNAs (Memczak *et al.*, 2013, Hansen *et al.*, 2013). S'ha vist que els circRNAs poden tenir fins a 70 zones evolutivament conservades complementàries a la zona *seed* d'un miRNA, de manera que actuaria com una esponja atrapant miRNAs i, conseqüentment, evitaria la disminució de l'expressió del mRNA diana del miRNA atrapat (Kosik, 2013).

Justament aquesta contínua evolució que també pateix el camp del miRNAs, ja que és una disciplina de treball relativament nova, ha generat que durant el temps d'elaboració de la tesi hagin canviat alguns dels coneixements principals o s'hagin

posat en dubte. Un exemple és la seva ubicació al genoma. Abans es creia que un miRNA havia d'estar localitzat en un únic lloc al genoma i ara ja hi ha miRNAs descrits en més d'una localització. Un altre canvi significatiu és l'auge en la descripció dels miRNAs\* o també anomenats miRNAs-3p, que s'ha observat que de vegades també poden formar part del complex RISC i realitzar una funció reguladora. Tots aquests ajustos han provocat, de retruc, un cert desgavell en la base de dades miRBase, que a cada actualització ha anat modificant nomenclatures, afegint nous miRNAs (alguns sense referències) i també eliminant-ne alguns per falta d'evidències. La gestió dels isomiRs és una de les mancances de la base de dades, així com mantenir una nomenclatura homòloga entre espècies, que fa que el treball entre miRNAs ortòlegs a vegades sigui més complicat. Per tot això la gestió de les dades dels nostres estudis es va realitzar prioritzant la homologia entre les seqüències i no pas els noms de cada miRNA que atorga miRBase.

Els dos grans blocs en que es divideix la tesi han estat, per una banda, la caracterització de miRNAs al ronyó de diferents races porcines i, per altra banda, la descripció i estudi de miRNAs d'origen víric durant el procés d'infecció de la malaltia d'Aujeszky, així com també desxifrar les possibles interaccions entre l'expressió de miRNAs de l'hoste i el desenvolupament del procés infecció.

Degut a que el nombre de miRNAs descrits en l'espècie porcina era relativament baix comparat amb l'espècie humana o altres espècies domèstiques, el primer objectiu va ésser augmentar el nombre de miRNAs descrits al porc utilitzant la seqüenciació massiva com a aproximació. El material animal que es va utilitzar varen ser els ronyons de sis races porcines (Ibèrica, Landrace, Large White, Piétrain, Meishan i Vietnamita) de diferents orígens i el seu ancestre europeu, el porc senglar. El ronyó es va escollir ja que en l'altre bloc de la tesi, la infecció experimental *in vitro* amb el virus de la malaltia d'Aujeszky, es van realitzar cultius de la línia cel·lular PK-15 per a produir la infecció viral, cèl·lules que deriven del ronyó porcí. Així doncs, el perfil de miRNAs al ronyó de l'espècie porcina s'hi van descriure un total de 229 miRNAs, més de la meitat (119, 52%) encara no descrits al porc. Una de les observacions inicials va ser la confirmació de la gran variació als extrems de la seqüència que presentaven els miRNAs, sobretot a

l'extrem 3'. Aquestes variacions nucleotídiques a l'extrem de la seqüència madura del miRNA generen isoformes, els anomenats isomiRs. Els mecanismes pels quals es genera aquesta diversitat, que en alguns casos pot arribar a generar isoformes funcionalment diferents, són diversos, des d'una heterogenicitat en el procés de biogènesi del miRNA fins a modificacions post transcripcionals com l'edició de l'RNA, l'activitat exonucleasa o l'addició de nucleòtids (Vickers *et al.*, 2013). En alguns casos, l'isomiR més abundant descrit d'un miRNA a les dades obtingudes no concordava amb la seqüència descrita a la bases de dades de miRNAs, miRBase. Aquest fet fa pensar que l'isomiR dominant no sempre és el mateix, afegint un grau més de complexitat en la capacitat reguladora dels miRNAs. Els patrons observats varen ser que en alguns miRNAs predominava l'expressió d'un isomiR i en d'altres no hi havia una dominància clara. En alguns casos, fins i tot, la seqüència predominant entre races de porc variava, com en el cas del miR-99a i el miR-200c-3p. En el miR-99a, la isoforma predominant s'expressava en primer lloc a les races europees comercials Landrace, Large White i Piétrian, justament les races europees amb influència asiàtica. En la raça europea Ibèrica (sense influència asiàtica) i el porc senglar, i les races asiàtiques Meishan i Vietnamita s'expressava més la segona isoforma. En el cas del miR-200c-3p, la isoforma predominant ho era per a totes les races estudiades excepte per la Vietnamita, on predominava la segona isoforma més expressada. Aquest patró de variació en la predominança d'expressió de diferents isoformes entre races va aparèixer en 76 miRNAs, un 34% del total de miRNAs descrits. No es va poder, però, associar aquestes variacions amb l'origen geogràfic de les races (europees vs. asiàtiques).

Els miRNAs més expressats en mostres de ronyó segons les dades obtingues de seqüenciació massiva varen ser el miR-200b-3p, el miR-125b, el miR-23b, el miR-126 i el miR-23a. Curiosament, el miR-200b-3p representava el 23,5% de totes les seqüències alineades amb miRBase. Aquest patró d'expressió, on un miRNA predomina en una situació temporal i espacial concreta ha estat descrita en altres ocasions, com en el cas del miR-1 al múscul esquelètic de porc (Nielsen *et al.*, 2010). Són miRNAs que no es poden considerar específics de teixit ja que també s'han descrit en d'altres però sí que realitzen una funció important en aquell teixit on la seva

expressió és predominant. Tot i no resultar significatiu estadísticament, una dada rellevant fou que el miR-125b va ser el miRNA més expressat a la raça Ibèrica, en contraposició a la resta de races i el porc senglar amb el miR-200b-3p. Aquesta dada podria estar relacionada amb l'origen de la domesticació del porc i els moviments migratoris que han anat succeint, on es demostra que les races comercials actuals (les europees amb influència asiàtica) tenen força material genètic en comú amb les races asiàtiques. Concretament, Groenen *et al.*, (2012) proposa que un 35% del genoma d'aquestes races és d'origen asiàtic.

A la comparativa de les dades de seqüenciació respecte l'abundància dels miRNAs entre els grups de races (tronc europeu: raça Ibèrica i porc senglar; tronc asiàtic: races Meishan i Vietnamita i tronc europeu amb influència asiàtica: races Landrace, Large White i Piétrain) es van detectar 125 miRNAs amb una expressió diferencial en algun dels tres grups. Ja que el nombre total de seqüències era relativament baix i no permetia una anàlisi en profunditat del nivell d'expressió només basant-nos en els recomptes de seqüències, per a fer la validació de la seva expressió mitjançant la tècnica de la RT-qPCR ens varem centrar en aquells miRNAs més expressats o que presentaven unes variacions d'un grup respecte un altre més grans. Es va validar l'expressió dels 14 miRNAs escollits, entre ells el miR-200b-3p, el miR-200c-3p i el miR-193b, tots tres miRNAs ortòlegs i que, per tant, encara no estaven descrits en el porc. Es varen trobar diferències significatives ( $p$ -valor $<0,05$ ) en 8 miRNAs. Una dada a contrastar va ser que en alguns casos la tendència en l'expressió dels miRNAs no es corresponia amb la observada en les dades de seqüenciació. El miR-200b-3p i el let-7f van resultar estar sobre expressats en les races europees, mentre que a les dades de seqüenciació tenien una major expressió a les races asiàtiques. El miR-200c-3p va presentar una major expressió a les races asiàtiques quan les dades obtingudes de seqüenciació mostraven una major expressió a les races europees comercials. La resta de miRNAs diferencialment expressats varen seguir el mateix patró d'expressió o semblant. El miR-126\* i el miR-193b van tenir una major expressió a les races europees; el miR-126 i el miR-99a van presentar una menor expressió a les races asiàtiques mentre que el miR-486 va presentar una major expressió a les races europees comercials.



Aquests resultats dispars entre les dades de seqüenciació massiva i les dades d'expressió de l'RT-qPCR alimenten el debat de si la seqüenciació massiva és una eina prou fidel per a fer-ne un ús quantitatiu. No és un debat nou, però sí que evoluciona en tant que també ho fan les tècniques de seqüenciació massiva. Les plataformes de segona generació encara necessiten d'una pre-amplificació, una passa clarament desestabilitzadora en l'obtenció d'una mostra fidel a l'original i, fins ara, sempre s'ha requerit d'un prova confirmadora a l'hora de concloure qualsevol tipus d'expressió, com el Northern blot o la RT-qPCR. Durant el desenvolupament de la tesi hi ha hagut un gran avenç amb les tècniques de seqüenciació massiva, que ha propiciat l'ús de diferents plataformes que utilitzen diferents principis de seqüenciació en els diferents treballs. En la caracterització del miRNAoma al ronyó del porc es va utilitzar l'aparell 454 GS FLX de Roche, basat en la piroseqüenciació. Va ser una de les plataformes pioneres de 2a generació i el seu rendiment per seqüència generada actualment és car en comparació amb les noves plataformes que hi ha al mercat. Així, el treball de descripció de miRNAs virals i de l'hoste en una infecció experimental amb l'SHV-1 es va realitzar amb l'aparell Ion Torrent de Life Technologies basat en la seqüenciació per síntesi amb sensibilitat iònica, un sistema detecció més barat que l'òptic. Ion Torrent ofereix més seqüències i més curtes, que no és problema si treballem amb miRNAs. La nova generació de plataformes de seqüenciació massiva ja no necessiten un amplificació prèvia i, per tant, es pot obtenir una seqüenciació de dades més fidel a la realitat. Just ara es comencen a comercialitzar aquestes plataformes i caldrà veure si les tecnologies amb que s'han dissenyat són prou eficaces per obtenir un nombre de seqüències i d'una qualitat suficient per a poder tractar els resultats plenament com a quantitativus.

Per intentar dotar de sentit biològic tots aquests resultats es va realitzar una anàlisi funcional *in silico* amb tots els gens que es van determinar com a possibles dianes dels miRNAs diferencialment expressats. Degut al gran nombre d'estudis que existeixen relacionant els miRNAs i les neoplàsies i del progrés en aquest camp, moltes de les rutes que es van associar als miRNAs diferencialment expressats estaven relacionades amb el desenvolupament de neoplàsies (com el carcinoma de cèl·lules renals), cicle

cel·lular i homeòstasi cel·lular. Tanmateix, el miR-200b-3p, que en estudis previs ja ha estat relacionat amb patologies del ronyó (Oba *et al.*, 2010, Li *et al.*, 2010), també se'l va relacionar amb una major eficiència del creixement muscular en les races europees, on es troba sobre expressat. Podem associar el múscul i el ronyó per la secreció d'eritropoetina, una hormona secretada al ronyó i necessària pel creixement muscular ja que és la responsable de l'activació de l'eritropoesi i del transport d'oxigen als músculs. El miR-200c-3p se'l va relacionar amb processos reproductius com la meiosi de l'òocit o la maduració d'aquest mediat per la progesterona. Si tenim en compte que les races asiàtiques presenten millors fenotips pels caràcters reproductius i que el miR-200c-3p està sobre expressat a les races asiàtiques, aquest podria estar implicat en la regulació d'aquests processos. El miR-99a també se'l va relacionar amb aquests processos reproductors, juntament amb el miR-200b-3p. Com que aquests dos miRNAs es varen trobar sobre expressats en races europees, podrien jugar un paper antagonista en aquests processos, ja que les races europees no presenten tan bones característiques reproductives com les races asiàtiques. Tot i així, cal tenir en compte que totes aquestes associacions s'han dut a terme amb softwares que treballen amb les anotacions funcionals dels gens en l'espècie humana. Per tant, s'assumeix que les funcions dels gens porcins seran les mateixes partint de la elevada homologia que existeix entre el genoma humà i porcí. També cal tenir en compte que els gens descrits podrien tenir funcions fins ara no descrites.

Un altre objectiu del treball de caracterització de miRNAs al ronyó va ésser generar un protocol de descripció i validació de nous miRNAs a partir de dades de seqüenciació massiva, basat en les pautes que s'anaven publicant en els diferents treballs de descripció de miRNAs (Ambros *et al.*, 2003, Kozomara and Griffiths-Jones, 2011, Griffiths-Jones *et al.*, 2008, Griffiths-Jones *et al.*, 2006). El primer pas va ésser fer el mapeig de totes les seqüències obtingudes al genoma porcí considerant un alineament perfecte. De les seqüències amb un alineament perfecte positiu es varen descartar aquelles que es van alinear en més d'una regió del genoma. Després es va acotar a aquelles seqüències d'entre 19 i 23 nucleòtids, mida característica del miRNA, i es van descartar totes aquelles seqüències que ja presentaven una anotació funcional. Finalment també es varen eliminar aquelles seqüències amb un nombre de còpies

menor a 3. Les seqüències seleccionades es van agrupar tenint en compte la seva posició al genoma per generar clústers. Cada clúster va adoptar com a seqüència de referència aquella que presentava més còpies. L'últim pas abans de validar l'expressió dels nous miRNAs per RT-qPCR va ésser comprovar que el pre-miRNA realitzava el plegament típic de forqueta. Per això es varen seleccionar les regions flanquejants del miRNA madur i es va utilitzar el software MFold (Zuker, 2003). Així, es varen descriure 8 clústers putatius de ser nous miRNAs, 5 dels quals se'n va confirmar l'expressió (CI-2, CI-5, CI-25, CI-29 i CI-38) mitjançant la quantificació relativa de les dades obtingudes per RT-qPCR.

La quantificació relativa mitjançant la tècnica de l'RT-qPCR s'ha anat generalitzant aquests últims anys per a la validació de l'expressió dels miRNAs. Per poder fer la quantificació cal normalitzar l'expressió dels miRNAs d'interès fent servir un gen o varis gens de referència, l'expressió dels quals es caracteritza per ser estable. A l'hora d'escollir els gens de referència que faríem servir per la normalització de les dades ens vàrem adonar que els gens més utilitzats en estudis d'expressió de miRNAs eren RNAs ribosomals com el 5S o RNAs nuclears petits com l'RNU6B. En aquest sentit es va creure convenient estudiar l'estabilitat en l'expressió dels miRNAs en diferents teixits per a utilitzar-los com a gens de referència, ja que és important que aquests siguin de la mateixa naturalesa que els objectes d'estudi i tinguin la mateixa llargària per assegurar una eficiència similar durant l'aïllament de l'RNA i la seva retrotranscripció (Schaefer *et al.*, 2010). Vàrem fer un disseny experimental per estudiar l'estabilitat dels 10 miRNAs més utilitzats com a gens de referència en 5 teixits diferents (múscul esquelètic, ronyó, fetge, ovari i úter) i en 5 races de porc diferents (Ibèrica, Landrace, Large White, Meishan i Vietnamita). Tots els miRNAs estudiats varen presentar valors d'estabilitat correctes per ser utilitzats com a gens de referència segons els algorismes geNorm i NormFinder. L'estudi estadístic va evidenciar que el teixit és un factor de variabilitat important en els miRNAs, mentre que la raça no va generar diferències significatives. El miRNA més estable en el conjunt de teixits va ser el miR-93, tot i que cada teixit va presentar el seu propi miRNA més estable: el miR-17-5p al ronyó, el miR-103 a l'úter, el miR-26a al fetge i el let-7a a l'ovari i al múscul. També es va determinar quants miRNAs de referència calia emprar en un estudi, utilitzant el valor de variació V.

Aquest valor mostra la mitjana dels factors de normalització per a cada mostra de l'estudi. Comparant l'evolució d'aquest valor cada vegada que afegim un nou miRNA com a gen de referència de manera seqüencial, podem observar l'efecte d'estabilitat que genera el miRNA afegit. Quan aquesta variació comença a ser petita podem considerar que l'addició d'aquell miRNA com a normalitzador no té un efecte important en l'estabilitat del grup de mostres. Com a conclusions, es va determinar que en un estudi on només es treballa amb un teixit seria necessari utilitzar 3 miRNAs com a gens de referència per aconseguir una bona normalització de les dades, mentre que en un estudi amb més d'un teixit en caldrien 5 per garantir una correcta normalització d'aquestes. Aquest treball ens demostra que els miRNAs poden ser uns bons gens de referència ja que presenten uns bons valors d'estabilitat.

El segon bloc de la tesi s'ha centrat en la determinació i estudi dels miRNAs virals i de l'hoste que podrien estar implicats en el desenvolupament de la infecció amb el virus de la malaltia d'Aujeszky també a través de les tècniques de seqüenciació massiva. En el disseny experimental es varen fer servir dues soques del virus: una de virulenta (NIA-3) i una de vacunal (Begonia). La soca vacunal Begonia deriva de la soca virulenta NIA-3 i és una soca viva atenuada, modificada genèticament, que permet al virus replicar-se però no presenta signes clínics ja que la replicació del virus és molt poc eficient. Es varen fer servir dues aproximacions: una infecció *in vitro* amb la línia cel·lular PK-15 derivada del ronyó porcí, i una infecció *in vivo* amb animals de la raça Landrace de 4 setmanes de vida en el moment de la infecció. Les mostres preses en la infecció *in vivo* varen ser el bulb olfactori i el gangli trigemin, ja que el virus es replica inicialment al teixit nerviós i és on produeix la latència. Es varen prendre mostres a diferents temps d'infecció: 12, 24 i 30 hores post infecció en el cas *in vitro* i 4,5,6 i 7 dies post infecció en el cas *in vivo*. També es va afegir un grup control (no infecció) a cada aproximació.

Es va determinar el perfil de miRNAs a cada aproximació, descrivint 229 miRNAs en l'aproximació *in vitro* i 302 miRNAs en l'aproximació *in vivo*. Igual que en l'estudi del miRNAoma en el ronyó de porc, la meitat dels miRNAs descrits foren ortòlegs i, seguint el mateix patró observat en el ronyó, l'expressió d'un miRNA va predominar a cada

aproximació. En aquest cas, el miR-23a-3p va predominar en l'aproximació *in vitro*, amb un 50% de les seqüències alineades a miRBase, i la del miR-125b-5p va fer-ho a l'aproximació *in vivo*, amb un 26% de les seqüències. Comparant el perfil d'expressió del grup control de l'aproximació *in vitro* amb el perfil obtingut en el ronyó de porc del treball anterior, tenint en compte que la línia cel·lular PK-15 deriva de cèl·lules del ronyó de porc, vàrem observar certa similitud. Comparant els miRNAs més expressats a cada perfil (aquells amb un recompte mínim de 100), es va observar fins a un 42% de miRNAs compartits. Salvant les diferències entre una línia cel·lular i un teixit, vàiem, doncs, certa analogia entre les cèl·lules PK-15 i el ronyó. De fet, hi ha més similituds entre les cèl·lules PK-15 i el ronyó que entre les cèl·lules PK-15 i els teixits utilitzats per a la infecció *in vivo* amb l'HSV-1, el bulb olfactori i el gangli trigemin, d'origen nerviós. En aquest cas, de tots els miRNAs descrits a les dues aproximacions només un 33% apareixen als dos perfils. Aquesta dada pot fer pensar, assumint que el perfil d'expressió dels miRNAs varia en teixits diferents, que la resposta davant la infecció podria variar segons l'aproximació. Una altra dada que corroborava aquest canvi de resposta fou que en l'aproximació *in vitro* es van descriure 7 miRNAs virals i només 1 en l'aproximació *in vivo*. Sis dels miRNAs virals es corresponien a l'SHV-1 (miR-LLT1, miR-LLT2, miR-LLT6, miR-LLT8, miR-LLT9 i miR-LLT11a), mentre que el restant (miR-rR1-5) estava descrit en un altre herpesvirus que afecta el macaco (*Macaca mulatta*). Els 6 miRNAs descrits de l'SHV-1 a l'aproximació *in vitro* es varen trobar tant en la infecció amb la soca NIA-3 com amb la infecció amb la soca Begonia, mentre que en l'aproximació *in vivo*, l'únic miRNA descrit, el miR-LLT1, només es va detectar en les mostres infectades amb la soca NIA-3 del virus. Val a dir, que el miR-LLT1 va ser el miRNA viral més expressat en l'aproximació *in vitro*. El miR-rR1-5, curiosament, també va ser detectat en mostres control, fet que ens podria portar a pensar en algun error o una mala descripció del miRNA a la base de dades miRBase, ja que el nombre de còpies (N=415) amb que s'ha detectat no és irrellevant.

L'estudi d'expressió diferencial dels miRNAs va mostrar dinàmiques diferents entre les dues aproximacions. Mentre que en l'aproximació *in vitro* la majoria dels miRNAs diferencialment expressats més abundants ho feien presentant una major expressió al grup control, en l'aproximació *in vivo* ho feien expressant-se més als grups infectats. En

l'aproximació *in vivo*, el let-7a i el miR-378 van presentar una major expressió en els teixits infectats, especialment amb la soca NIA-3. El miR-206 només es va detectar als teixits infectats, mostrant unes diferències d'expressió molt grans entre els grups NIA-3 i Begonia, essent més expressat al grup NIA-3. El miR-133a va mostrar un augment en la seva expressió al grup NIA-3, mentre que va presentar una baixada de la seva expressió al grup Begonia. Ja a l'aproximació *in vitro*, el miR-LLT1 va ser el miRNA més expressat només a teixits infectats, mentre que el miR-92a i el miR-92b-3p presentaven un augment en la seva expressió en la infecció amb la soca Begonia i un descens en la soca NIA-3. L'anàlisi funcional va relacionar els possibles gens diana amb rutes implicades amb el procés d'infecció viral i la resposta immune, sobretot dels miRNAs amb un augment d'expressió en el grup infectat amb la soca virulenta NIA-3 (let-7a, miR-378, miR-206 i miR-133a). També es varen relacionar amb processos biològics més generals relacionats amb funcions cel·lulars. Finalment, també van aparèixer rutes relacionades amb el sistema nerviós, implicades en la formació de la xarxa neuronal o la diferenciació i supervivència de cèl·lules neuronals, degut a l'origen nerviós dels teixits estudiats en l'aproximació *in vivo*. El miR-92a i el miR-92b-3p també se'ls va relacionar amb un ampli ventall de processos biològics. El miR-92a, que s'expressava més en el grup infectat amb la soca Begonia a l'aproximació *in vitro*, va canviar la seva expressió a l'aproximació *in vivo*, expressant-se més en el grup infectat NIA-3. Aquesta disminució en l'expressió al grup Begonia d'una aproximació a l'altra es podria explicar per les diferències en la reacció defensiva de l'hoste, on el desenvolupament de la resposta immunitària en l'aproximació *in vivo* provoca un descens en l'expressió del miR-92a en el grup infectat amb la soca viva atenuada Begonia. El miR-92b-3p va mantenir la seva expressió en les dues aproximacions, estant sobre expressat en les mostres infectades amb la soca Begonia, i podria jugar un paper més general en el desenvolupament de funcions cel·lulars. Sembla ser, doncs, que els miRNAs de l'hoste podrien tenir un paper rellevant en la lluita contra la infecció viral i el desenvolupament de la resposta immunitària en els teixits infectats.

Un dels objectius principals de l'estudi de la infecció amb el virus de la malaltia d'Aujeszky era la caracterització de nous miRNAs virals implicats en la infecció. Fent servir el mateix protocol de caracterització de nous miRNAs que a l'estudi dels ronyons

de porc es varen detectar 14 clústers putatius de ser miRNAs codificats pel genoma viral (viR01-14), 6 dels quals ja estaven descrits a miRBase (els 6 miRNAs virals que varem trobar en les dades de seqüenciació massiva). Tots els clústers estaven localitzats a la regió intrònica del transcrit de latència gran del virus (LLT), com tots els miRNAs de l'SHV-1 descrits fins l'actualitat, excepte un, el viR01, però no va passar la formació de l'estructura típica del pre-miRNA i va ser descartat. Un dels 6 miRNAs ja descrits a miRBase, el miR-LLT1 (viR02), fou el més expressat dels miRNAs a l'estudi, i per això vàrem decidir mantenir-lo; els altres 5 els vàrem descartar de la validació. Així, finalment es va validar l'expressió de 8 putatius miRNAs virals mitjançant l'RT-qPCR. D'aquests 8 miRNAs virals, 5 (viR02, viR04, viR06, viR09 i viR11) estaven descrits en els estudis previs (Wu *et al.*, 2012, Anselmo *et al.*, 2011) i 3 (viR05, viR08 i viR14) foren nous miRNAs virals descrits per nosaltres. Un dels 5 miRNAs virals ja descrits en estudis previs, el viR09, restava situat en la posició d'un moRNA. Un moRNA (miRNA offset RNA) és un RNA de mida petita ubicat just a la regió proximal contigua d'un miRNA o un miRNA\* en l'estructura de plegament del pre-miRNA (Shi *et al.*, 2009). Es desconeix la seva funció, tot i que es pensa que podria aprofitar el complex silenciador RISC per a exercir també una funció reguladora. El viR09 va ser descrit com a prv-miR-4 a l'estudi d'Anselmo *et al.*, (2011) i com a moRNA (moR8) a l'estudi de Wu *et al.*, (2012). Com que va ser el segon transcrit viral més abundant en el nostre estudi es va creure convenient mantenir-lo per a l'estudi funcional. Justament comparant l'abundància dels miRNAs virals detectats entre els estudis previs i el nostre vàrem observar certa divergència. El miRNA viral més expressat en el nostre estudi (viR02) també va ser detectat als estudis previs, però no com el més abundant, mentre que el més abundant en l'estudi d'Anselmo *et al.*, (2011), es va detectar en el nostre estudi amb poques còpies (viR12, 3 còpies). A més a més, el miRNA viral més abundant en l'estudi de Wu *et al.*, (2012) no l'hem detectat al nostre estudi. Aquestes diferències podrien ser explicades en part per les diferents aproximacions *in vitro* utilitzades, sobretot en l'estudi d'Anselmo *et al.*, (2011), que fan servir cèl·lules dendrítiques. Una altra explicació per a les diferències trobades entre el nostre estudi i el de Wu *et al.*, (2012), en els que ambdós hem fet servir la línia cel·lular PK-15 podria ser la soca viral utilitzada. En el nostre estudi s'ha utilitzat la soca virulenta NIA-3, provinent d'Europa,

mentre que Wu *et al.*, (2012) ha fet servir una altra soca virulenta, l'Ea, provinent de l'Àsia.

L'anàlisi de les dades obtingudes per RT-qPCR va revelar un canvi en l'expressió dels miRNAs virals en funció de l'aproximació *in vitro* o *in vivo*. Mentre tots els miRNAs virals es varen detectar tant en les mostres *in vitro* infectades amb la soca virulenta NIA-3 com en la soca viva atenuada Begonia, en les mostres *in vivo* només es varen detectar en el grup infectat amb la soca virulenta NIA-3. Aquest fet va demostrar que l'aproximació utilitzada és determinant per al desenvolupament de la infecció viral. L'estudi estadístic va confirmar que no es varen trobar diferències d'expressió entre la infecció amb la soca NIA-3 i la soca Begonia in l'aproximació *in vitro*. La soca viva atenuada Begonia està modificada genèticament de manera que la seva capacitat infectiva és menor. El fet que aquesta soca es pogués replicar i desenvolupar la infecció amb la mateixa eficiència que la soca virulenta NIA-3 en l'aproximació *in vitro*, mentre que en l'aproximació *in vivo* no ha estat així, es podria explicar, entre altres coses, primerament per les mateixes aproximacions (línia cel·lular vs. teixit infectat), i també per una manca de resposta immune en els cultius cel·lulars.

Es va realitzar una aproximació funcional *in silico* intentant relacionar la funció dels miRNAs virals descrits i també la dels miRNAs de l'hoste diferencialment expressats durant la infecció amb els gens que s'expressen en el genoma viral. Aquest cop es va fer servir l'algoritme miRanda ja que el genoma de referència era el viral. 7 miRNAs virals varen ser relacionats amb un total de 59 gens que codifica el genoma del virus. El transcrit de latència gran (LLT) va interaccionar amb 5 dels 7 miRNAs virals. El viR04 va ser el miRNA viral amb més interaccions descrites, amb un total de 48, i se'l va relacionar principalment amb funcions estructurals relacionades amb les glicoproteïnes de l'embolcall víric i les proteïnes de la càpside i el tegument. El viR02, el més expressat en les dades de seqüenciació, se'l va relacionar amb gens involucrats amb el desenvolupament de l'embolcall i les proteïnes del tegument, juntament amb el viR014, on podrien jugar un paper important en la morfogènesi del virus i en l'entrada del virus a la cèl·lula hoste. El viR05 i el viR08 van ser associats a gens no estructurals relacionats amb el processament del DNA, com l'escissió i encapsidació



d'aquest, o la replicació i l'emballatge. El viR06 va presentar forces interaccions amb gens relacionats amb la sortida del virus de la cèl·lula hoste, així que podria jugar algun paper regulador en aquest procés. A tot això, el viR02, el viR04, el viR06 i el viR14 també se'ls va associar als gens que realitzen funcions de regulació de l'expressió gènica (*EPO*, *IE180*, *UL41*, *UL48* i *UL54*). Wu *et al.*, (2012) també va associar els gens que regulen l'expressió gènica del virus amb el viR02, el viR04 i el viR06, així com també va determinar un gran nombre d'interaccions per al viR04, com en el nostre estudi. Finalment, el viR09, el moRNA descrit a l'estudi, va presentar interaccions tant amb gens estructurals relacionats amb les proteïnes de l'emballatge i la sortida del virus de la cèl·lula de l'hoste com amb gens no estructurals relacionats amb l'escissió, replicació i embalatge del DNA. Tot i no presentar cap interacció significativa, el viR11 va ser descrit per Wu *et al.*, (2012) i associat al gen regulador *EPO* i també a altres gens, tant estructurals com no estructurals. El fet de no haver-lo associat a cap gen viral en el nostre estudi pot ser degut a les condicions aplicades a l'hora de predir les interaccions, una mica més estrictes que a l'estudi previ (Wu *et al.*, 2012).

Els miRNAs de l'hoste sobre expressats durant la infecció van presentar interaccions amb 39 gens de l'SHV-1, postulant-se per tenir algun tipus responsabilitat en la contenció del desenvolupament de la infecció. El miR-133a va ser el que va presentar més interaccions, fins a 33, i juntament amb el miR-92a, el miR-92b-3p i el miR-378 se'ls va associar a la fase de latència per presentar interaccions amb el transcrit gran de latència (LLT) i els gens reguladors de l'expressió gènica. El miRNA let-7a també va ser relacionat amb gens estructurals encarregats de la formació de les glicoproteïnes i les proteïnes de la càpside vírica. El miR-206 no va presentar interacció amb cap gen viral.

La capacitat reguladora dels miRNAs, demostrada en molts estudis, cada vegada es dibuixa més complexa. A mesura que es van descrivint nous factors que incideixen directament en la funció reguladora dels miRNAs, com poden ser la variabilitat funcional dels propis isomiRs, la possibilitat que un miRNA estigui localitzat en més d'una posició al genoma, o l'aparició de molècules reguladores dels propis miRNAs com els circRNAs, s'afegeixen noves peces en l'engranatge regulador dels miRNAs que

no fan més que potenciar la seva regulació fina. El treball desenvolupat en aquesta tesi intenta aportar més coneixement en el repertori de miRNAs en l'espècie porcina i trobar implicacions biològiques en funció de l'expressió d'aquests i les característiques fenotípiques relacionades amb l'origen de les mostres de ronyons. També s'han pogut avaluar les bones capacitats dels miRNAs per ser utilitzats com a gens de referència en la quantificació relativa mitjançant RT-qPCR. Un altre avenç ha estat intentar desxifrar l'entramat regulador durant la infecció amb el virus de la malaltia d'Aujeszky, tot descrivint nous miRNAs virals i validant la seva expressió en una infecció *in vivo*. De la mateixa manera que s'ha descrit en estudis anteriors (Cullen, 2013, Boss i Renne, 2011) un efecte regulador dels miRNAs virals en l'homeòstasi cel·lular de l'hoste per afavorir l'entrada i replicació del virus, sembla ser que també hi ha una sèrie de mecanismes que fan que els miRNAs de l'hoste intentin reprimir la capacitat infectiva del virus interactuant amb els seus gens.

# 5. CONCLUSIONS

---

1. El miRNAoma al ronyó del porc s'ha definit amb l'expressió de 229 miRNAs. El miR-200b-3p, miR-125b, miR-23b, miR-126 i miR-23a han estat els miRNAs més abundants i presenten un perfil semblant entre les races estudiades i el porc senglar europeu. S'ha determinat el miR-200b-3p com el miRNA més abundant en quasi totes les races, a excepció del miR-125b que ho ha estat per a la raça Ibèrica.
2. Un 70% dels miRNAs descrits al miRNAoma del ronyó de porc presenten variabilitat als extrems de la cadena, generant isomiRs. Aquesta variabilitat es descriu en un 34% del total de miRNAs descrits amb un canvi en la predominança de l'isomiR més abundant segons la raça. Tot i així, no s'han pogut associar aquestes variacions amb l'origen geogràfic de les races (europees vs. Asiàtiques).
3. La caracterització del miRNAoma del ronyó i posterior anàlisi mitjançant RT-qPCR ha revelat diferències d'expressió en alguns miRNAs entre els grups de races estudiats: el miR-200b-3p, el miR-126, el let-7f i el miR-193b presenten un augment de la seva expressió a les races europees, el miR-200c-3p mostra augment de la seva expressió a les races asiàtiques i el miR-486 un augment de la seva expressió a les races europees comercials. Finalment, el miR-126 i el miR-99a presenten una menor expressió a les races asiàtiques.
4. Cinc nous miRNAs porcins (Cl-2, Cl-5, Cl-25, Cl-29 i Cl-38) s'han descrit al ronyó del porc mitjançant un protocol estricte de treball amb dades de seqüenciació massiva i la seva posterior validació amb RT-qPCR.

5. S'han determinat canvis en l'expressió de miRNAs de porc en una infecció experimental amb dues soques, una virulenta (NIA-3) i una vacunal (Begonia), del virus de la malaltia d'Aujeszky (SHV-1), mitjançant una aproximació *in vitro* amb la línia cel·lular PK-15 i una aproximació *in vivo* amb mostres de bulb olfactori i gangli trigemin. En l'aproximació *in vivo*, el let-7a i el miR-378 presenten un augment de la seva expressió en la soca NIA-3, així com el miR-206 i el miR-133a presenten diferències d'expressió entre les soques del virus, també més expressats en la soca NIA-3. En l'aproximació *in vitro*, la majoria dels miRNAs presenten un descens en la seva expressió en el grups infectats. El miR-92a i el miR-92b-3p presenten una sobre expressió en la soca Begonia.
  
6. L'anàlisi funcional *in silico* dels miRNAs let-7a, miR-378, miR-206, miR-133a, miR-92a i miR-92b-3p, tots ells sobre expressats d'alguna manera en estat d'infecció, ha relacionat els seus gens diana a rutes relacionades amb el procés d'infecció viral i la resposta immune, així com amb rutes associades al desenvolupament del sistema nerviós i altres processos cel·lulars. També han presentat interaccions amb 39 gens virals i se'ls ha relacionat amb el transcrit gran de latència (LLT) i amb altres gens virals encarregats de la regulació de l'expressió gènica (*EPO*, *IE180*, *UL41*, *UL48*), així com amb gens estructurals encarregats de la formació de les glicoproteïnes.
  
7. Les dues aproximacions (*in vivo* i *in vitro*) de la infecció amb l'SHV-1 han permès descriure i validar, mitjançant RT-qPCR, 8 miRNAs virals (viR02, viR04, viR05, viR06, viR08, viR09, viR11 i viR14). Tots ells s'han expressat en els grups infectats amb les dues soques de l'HSV-1 en l'aproximació *in vitro*, i només ho han fet en el grup infectat amb la soca NIA-3 en l'aproximació *in vivo*. L'anàlisi funcional *in silico* ha relacionat 7 dels miRNAs amb 59 gens virals. El viR02, el viR04 i el viR14 estan associats a gens estructurals involucrats en la formació de glicoproteïnes de l'embolcall viral i proteïnes del tegument i la càpside, i podrien estar involucrats en el procés d'entrada i infecció de la cèl·lula hoste. El viR05, el viR08 i també el viR04 estan associats a gens no estructurals, encarregats del processament del DNA. El viR06 podria estar implicat amb el

procés de sortida del virus de la cèl·lula hoste ja que interacciona amb la majoria dels gens encarregats de dur a terme aquesta funció. Finalment, el viR09, descrit en la posició d'un moRNA, podria estar associat tant a gens estructurals com a gens reguladors del DNA.

8. Tant la descripció del perfil de miRNAs en el ronyó de porc com els perfils de miRNAs elaborats durant la infecció amb el virus de la malaltia d'Aujeszky (SHV-1) demostren, mitjançant la gran quantitat de miRNAs ortòlegs que s'han detectat, que encara queden molts miRNAs per ser descrits a l'espècie porcina.
  
9. Els miRNAs poden ser utilitzats com a gens de referència en estudis de quantificació relativa d'altres miRNAs mitjançant la tècnica de l'RT-qPCR, ja que presenten uns valors òptims d'estabilitat. El teixit és un factor que genera variabilitat en l'estabilitat dels miRNAs, per això, en estudis amb un únic teixit es recomana utilitzar tres miRNAs com a gens de referència i en estudis amb més d'un teixit es recomana utilitzar-ne cinc.



## 6. BIBLIOGRAFIA

---

- Adams M.D., Kelley J.M., Gocayne J.D., Dubnick M., Polymeropoulos M.H., Xiao H., Merril C.R., Wu A., Olde B. & Moreno R.F. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science (New York, N.Y.)* 252, 1651-1656.
- Agrawal N., Dasaradhi P.V., Mohmmmed A., Malhotra P., Bhatnagar R.K. & Mukherjee S.K. (2003) RNA interference: biology, mechanism, and applications. *Microbiology and molecular biology reviews : MMBR* 67, 657-685.
- Alves E., Ovilo C., Rodriguez M.C. & Silio L. (2003) Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. *Animal Genetics* 34, 319-324.
- Alwine J.C., Kemp D.J. & Stark G.R. (1977) Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proceedings of the National Academy of Sciences of the United States of America* 74, 5350-5354.
- Ambros V. (2008) The evolution of our thinking about microRNAs. *Nature medicine* 14, 1036-1040.
- Ambros V., Bartel B., Bartel D.P., Burge C.B., Carrington J.C., Chen X., Dreyfuss G., Eddy S.R., Griffiths-Jones S., Marshall M., Matzke M., Ruvkun G. & Tuschl T. (2003) A uniform system for microRNA annotation. *RNA (New York, N.Y.)* 9, 277-279.
- Amills M., Clop A., Ramírez O. & Perez-Enciso M. (2010) Origin and Genetic Diversity of Pig Breeds. *eLS*.
- Andersson L. (2001) Genetic dissection of phenotypic diversity in farm animals. *Nature reviews.Genetics* 2, 130-138.
- Anselmo A., Flori L., Jaffrezic F., Rutigliano T., Cecere M., Cortes-Perez N., Lefevre F., Rogel-Gaillard C. & Giuffra E. (2011) Co-expression of host and viral microRNAs in porcine dendritic cells infected by the pseudorabies virus. *PLoS one* 6, e17374.
- Arteaga-Vazquez M., Caballero-Perez J. & Vielle-Calzada J.P. (2006) A family of microRNAs present in plants and animals. *The Plant Cell* 18, 3355-3369.
- Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C., Richardson J.E., Ringwald M., Rubin G.M. & Sherlock G.

- (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25, 25-29.
- Balcells I., Cirera S. & Busk P.K. (2011) Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC biotechnology* 11, 70.
- Bartel D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215-233.
- Bartel D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-297.
- Beilby R. (1800) A General History of Quadrupeds. the Figures Engraved on Wood by T. Bewick. , Newcastle.
- Bell J.T. & Spector T.D. (2011) A twin approach to unraveling epigenetics. *Trends in genetics : TIG* 27, 116-125.
- Benes V. & Castoldi M. (2010) Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods (San Diego, Calif.)* 50, 244-249.
- Berezikov E. (2011) Evolution of microRNA diversity and regulation in animals. *Nature reviews.Genetics* 12, 846-860.
- Betel D., Wilson M., Gabow A., Marks D.S. & Sander C. (2008) The microRNA.org resource: targets and expression. *Nucleic acids research* 36, D149-53.
- Bidanel J.P. (2011) Biology and Genetics of Reproduction. In: *The Genetics of the Pig, 2n Edition*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 218-241.
- Blow N. (2007) PCR's Next Frontier. *Nature methods* 4, 869-875.
- Boss I.W. & Renne R. (2011) Viral miRNAs and immune evasion. *Biochimica et biophysica acta* 1809, 708-714.
- Boss I.W., Plaisance K.B. & Renne R. (2009) Role of virus-encoded microRNAs in herpesvirus biology. *Trends in microbiology* 17, 544-553.
- Bowers J., Mitchell J., Beer E., Buzby P.R., Causey M., Efcavitch J.W., Jarosz M., Krzymanska-Olejnik E., Kung L., Lipson D., Lowman G.M., Marappan S., McInerney P., Platt A., Roy A., Siddiqi S.M., Steinmann K. & Thompson J.F. (2009) Virtual terminator nucleotides for next-generation DNA sequencing. *Nature methods* 6, 593-595.
- Branton D., Deamer D.W., Marziali A., Bayley H., Benner S.A., Butler T., Di Ventra M., Garaj S., Hibbs A., Huang X., Jovanovich S.B., Krstic P.S., Lindsay S., Ling X.S., Mastrangelo C.H., Meller A., Oliver J.S., Pershin Y.V., Ramsey J.M., Riehn R., Soni



- G.V., Tabard-Cossa V., Wanunu M., Wiggin M. & Schloss J.A. (2008) The potential and challenges of nanopore sequencing. *Nature biotechnology* 26, 1146-1153.
- Buchanan D.S. & Stalder K. (2011) Breeds of Pigs. In: *The Genetics of the Pig, 2nd Edition*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 445-472.
- Bustin S.A., Benes V., Nolan T. & Pfaffl M.W. (2005) Quantitative real-time RT-PCR--a perspective. *Journal of Molecular Endocrinology* 34, 597-601.
- Bustin S.A. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* 25, 169-193.
- Carthew R.W. & Sontheimer E.J. (2009) Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642-655.
- Catto J.W., Alcaraz A., Bjartell A.S., De Vere White R., Evans C.P., Fussel S., Hamdy F.C., Kallioniemi O., Mengual L., Schlomm T. & Visakorpi T. (2011) MicroRNA in prostate, bladder, and kidney cancer: a systematic review. *European urology* 59, 671-681.
- Chen C., Deng B., Qiao M., Zheng R., Chai J., Ding Y., Peng J. & Jiang S. (2012) Solexa sequencing identification of conserved and novel microRNAs in backfat of Large White and Chinese Meishan pigs. *PLoS one* 7, e31426.
- Chen C., Ridzon D.A., Broomer A.J., Zhou Z., Lee D.H., Nguyen J.T., Barbisin M., Xu N.L., Mahuvakar V.R., Andersen M.R., Lao K.Q., Livak K.J. & Guegler K.J. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic acids research* 33, e179.
- Cho I.S., Kim J., Seo H.Y., Lim do H., Hong J.S., Park Y.H., Park D.C., Hong K.C., Whang K.Y. & Lee Y.S. (2010) Cloning and characterization of microRNAs from porcine skeletal muscle and adipose tissue. *Molecular biology reports* 37, 3567-3574.
- Chow T.F., Youssef Y.M., Lianidou E., Romaschin A.D., Honey R.J., Stewart R., Pace K.T. & Yousef G.M. (2010) Differential expression profiling of microRNAs and their potential involvement in renal cell carcinoma pathogenesis. *Clinical biochemistry* 43, 150-158.
- Clop A., Amills M., Noguera J.L., Fernandez A., Capote J., Ramon M.M., Kelly L., Kijas J.M., Andersson L. & Sanchez A. (2004) Estimating the frequency of Asian cytochrome B haplotypes in standard European and local Spanish pig breeds. *Genetics, selection, evolution : GSE* 36, 97-104.
- Costa F.F. (2010) Non-coding RNAs: Meet thy masters. *BioEssays : news and reviews in molecular, cellular and developmental biology* 32, 599-608.

- Cullen B.R. (2013) MicroRNAs as mediators of viral evasion of the immune system. *Nature immunology* 14, 205-210.
- Dahan O., Gingold H. & Pilpel Y. (2011) Regulatory mechanisms and networks couple the different phases of gene expression. *Trends in genetics : TIG* 27, 316-322.
- Darwin C. (1868) *The Variation of Animals and Plants Under Domestication*. John Murray, London.
- Dekkers J.C., Mathur P.K. & Knol E.F. (2011) Genetic Improvement of the Pig. In: *The Genetics of the Pig, 2nd Edition*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 390-425.
- Dekkers J.C. (2007) Marker-assisted selection for commercial crossbred performance. *Journal of animal science* 85, 2104-2114.
- Dekkers J.C. (2004) Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *Journal of animal science* 82 E-Suppl, E313-328.
- Distl O. (2007) Mechanisms of regulation of litter size in pigs on the genome level. *Reproduction in domestic animals = Zuchthygiene* 42 Suppl 2, 10-16.
- Dobney K. & Larson G. (2006) Genetics and animal domestication: new windows on an elusive process. *Journal of Zoology* 269, 261-271.
- Eid J., Fehr A., Gray J., Luong K., Lyle J., Otto G., Peluso P., Rank D., Baybayan P., Bettman B., Bibillo A., Bjornson K., Chaudhuri B., Christians F., Cicero R., Clark S., Dalal R., Dewinter A., Dixon J., Foquet M., Gaertner A., Hardenbol P., Heiner C., Hester K., Holden D., Kearns G., Kong X., Kuse R., Lacroix Y., Lin S., Lundquist P., Ma C., Marks P., Maxham M., Murphy D., Park I., Pham T., Phillips M., Roy J., Sebra R., Shen G., Sorenson J., Tomaney A., Travers K., Trulson M., Vieceli J., Wegener J., Wu D., Yang A., Zaccarin D., Zhao P., Zhong F., Korlach J. & Turner S. (2009) Real-time DNA sequencing from single polymerase molecules. *Science (New York, N.Y.)* 323, 133-138.
- Epstein H. (1971) *Domestic Animals of China*. Africana Publication Corp., New York.
- Fang M., Berg F., Ducos A. & Andersson L. (2006) Mitochondrial haplotypes of European wild boars with  $2n = 36$  are closely related to those of European domestic pigs with  $2n = 38$ . *Animal Genetics* 37, 459-464.
- Fendler A., Stephan C., Yousef G.M. & Jung K. (2011) MicroRNAs as regulators of signal transduction in urological tumors. *Clinical chemistry* 57, 954-968.
- Feng R., Chen X., Yu Y., Su L., Yu B., Li J., Cai Q., Yan M., Liu B. & Zhu Z. (2010) miR-126 functions as a tumour suppressor in human gastric cancer. *Cancer letters* 298, 50-63.

- Friedman R.C., Farh K.K., Burge C.B. & Bartel D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome research* 19, 92-105.
- Garcia D.M., Baek D., Shin C., Bell G.W., Grimson A. & Bartel D.P. (2011) Weak seed-pairing stability and high target-site abundance decrease the proficiency of Isy-6 and other microRNAs. *Nature structural & molecular biology* 18, 1139-1146.
- Ginzinger D.G. (2002) Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Experimental hematology* 30, 503-512.
- Giuffra E., Kijas J.M., Amarger V., Carlborg O., Jeon J.T. & Andersson L. (2000) The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* 154, 1785-1791.
- Glazov E.A., Cottee P.A., Barris W.C., Moore R.J., Dalrymple B.P. & Tizard M.L. (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome research* 18, 957-964.
- Goddard M.E. & Hayes B.J. (2007) Genomic selection. *Journal of animal breeding and genetics = Zeitschrift fur Tierzuchtung und Zuchtungsbiologie* 124, 323-330.
- Godwin J.G., Ge X., Stephan K., Jurisch A., Tullius S.G. & Iacomini J. (2010) Identification of a microRNA signature of renal ischemia reperfusion injury. *Proceedings of the National Academy of Sciences of the United States of America* 107, 14339-14344.
- Griffiths A., Wessler S., Lewontin R. & Carroll S. (2008) Introduction to Genetic Analysis. W.H. Freeman and Company, New York.
- Griffiths-Jones S., Saini H.K., van Dongen S. & Enright A.J. (2008) miRBase: tools for microRNA genomics. *Nucleic acids research* 36, D154-8.
- Griffiths-Jones S., Grocock R.J., van Dongen S., Bateman A. & Enright A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic acids research* 34, D140-4.
- Griffiths-Jones S. (2004) The microRNA Registry. *Nucleic acids research* 32, D109-11.
- Groenen M.A., Archibald A.L., Uenishi H., Tuggle C.K., Takeuchi Y., Rothschild M.F., Rogel-Gaillard C., Park C., Milan D., Megens H.J., Li S., Larkin D.M., Kim H., Frantz L.A., Caccamo M., Ahn H., Aken B.L., Anselmo A., Anthon C., Auvil L., Badaoui B., Beattie C.W., Bendixen C., Berman D., Blecha F., Blomberg J., Bolund L., Bosse M., Botti S., Bujie Z., Bystrom M., Capitanu B., Carvalho-Silva D., Chardon P., Chen C., Cheng R., Choi S.H., Chow W., Clark R.C., Clee C., Crooijmans R.P., Dawson H.D., Dehais P., De Sapio F., Dibbits B., Drou N., Du Z.Q., Eversole K., Fadista J., Fairley S., Faraut T., Faulkner G.J., Fowler K.E., Fredholm M., Fritz E., Gilbert J.G., Giuffra E., Gorodkin J., Griffin D.K., Harrow J.L., Hayward A., Howe K., Hu Z.L., Humphray S.J., Hunt T., Hornshoj H., Jeon J.T., Jern P., Jones M., Jurka J.,

- Kanamori H., Kapetanovic R., Kim J., Kim J.H., Kim K.W., Kim T.H., Larson G., Lee K., Lee K.T., Leggett R., Lewin H.A., Li Y., Liu W., Loveland J.E., Lu Y., Lunney J.K., Ma J., Madsen O., Mann K., Matthews L., McLaren S., Morozumi T., Murtaugh M.P., Narayan J., Nguyen D.T., Ni P., Oh S.J., Onteru S., Panitz F., Park E.W., Park H.S., Pascal G., Paudel Y., Perez-Enciso M., Ramirez-Gonzalez R., Reecy J.M., Rodriguez-Zas S., Rohrer G.A., Rund L., Sang Y., Schachtschneider K., Schraiber J.G., Schwartz J., Scobie L., Scott C., Searle S., Servin B., Southey B.R., Sperber G., Stadler P., Sweedler J.V., Tafer H., Thomsen B., Wali R., Wang J., Wang J., White S., Xu X., Yerle M., Zhang G., Zhang J., Zhang J., Zhao S., Rogers J., Churcher C. & Schook L.B. (2012) Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491, 393-398.
- Groenen M.A., Schook L.B. & Archibald A.L. (2011) Pig Genomics. In: *The Genetics of the Pig, 2nd Edition*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 179-199.
- Groves C.P. (1981) *Ancestors for the Pigs: Taxonomy and Phylogeny of the Genus Sus*. Australian National University Press, Canberra.
- Grundhoff A. & Sullivan C.S. (2011) Virus-encoded microRNAs. *Virology* 411, 325-343.
- Gu Y., Li M., Zhang K., Chen L., Jiang A.A., Wang J., Lv X. & Li X. (2011) Identification of suitable endogenous control microRNA genes in normal pig tissues. *Animal science journal = Nihon chikusan Gakkaiho* 82, 722-728.
- Gunn G.G. (2003) *First Globalization: The Eurasian Exchange, 1500–1800*. Rowman and Littlefield, Lanham.
- Hansen T.B., Jensen T.I., Clausen B.H., Bramsen J.B., Finsen B., Damgaard C.K. & Kjems J. (2013) Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384-388.
- Heid C.A., Stevens J., Livak K.J. & Williams P.M. (1996) Real time quantitative PCR. *Genome research* 6, 986-994.
- Hou X., Tang Z., Liu H., Wang N., Ju H. & Li K. (2012) Discovery of MicroRNAs associated with myogenesis by deep sequencing of serial developmental skeletal muscles in pigs. *PLoS one* 7, e52123.
- Huang P.J., Lin W.C., Chen S.C., Lin Y.H., Sun C.H., Lyu P.C. & Tang P. (2012) Identification of putative miRNAs from the deep-branching unicellular flagellates. *Genomics* 99, 101-107.
- Huang T.H., Zhu M.J., Li X.Y. & Zhao S.H. (2008) Discovery of porcine microRNAs and profiling from skeletal muscle tissues during development. *PLoS one* 3, e3225.
- John B., Enright A.J., Aravin A., Tuschl T., Sander C. & Marks D.S. (2004) Human MicroRNA targets. *PLoS biology* 2, e363.

- Jones G.F. (1998) Genetic aspects of domestication, common breeds and their origin. In: *The Genetics of the Pig*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 17-50.
- Kanehisa M. & Goto S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27-30.
- Kanis E., De Greef K.H., Hiemstra A. & van Arendonk J.A. (2005) Breeding for societally important traits in pigs. *Journal of animal science* 83, 948-957.
- Kijas J.M. & Andersson L. (2001) A phylogenetic study of the origin of the domestic pig estimated from the near-complete mtDNA genome. *Journal of Molecular Evolution* 52, 302-308.
- Kim K.I., Lee J.H., Li K., Zhang Y.P., Lee S.S., Gongora J. & Moran C. (2002) Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. *Animal Genetics* 33, 19-25.
- Klupp B.G., Hengartner C.J., Mettenleiter T.C. & Enquist L.W. (2004) Complete, annotated sequence of the pseudorabies virus genome. *Journal of virology* 78, 424-440.
- Korlach J., Bjornson K.P., Chaudhuri B.P., Cicero R.L., Flusberg B.A., Gray J.J., Holden D., Saxena R., Wegener J. & Turner S.W. (2010) Real-time DNA sequencing from single polymerase molecules. *Methods in enzymology* 472, 431-455.
- Kosik K.S. (2013) Molecular biology: Circles reshape the RNA world. *Nature* 495, 322-324.
- Kozomara A. & Griffiths-Jones S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic acids research* 39, D152-7.
- Krol J., Loedige I. & Filipowicz W. (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nature reviews.Genetics* 11, 597-610.
- Kubista M., Andrade J.M., Bengtsson M., Forootan A., Jonak J., Lind K., Sindelka R., Sjoback R., Sjogreen B., Strombom L., Stahlberg A. & Zoric N. (2006) The real-time polymerase chain reaction. *Molecular aspects of medicine* 27, 95-125.
- Kumar P., Johnston B.H. & Kazakov S.A. (2011) miR-ID: a novel, circularization-based platform for detection of microRNAs. *RNA (New York, N.Y.)* 17, 365-380.
- Kuzmuk K.N. & Schook L.B. (2011) Pig as a Model for Biomedical Sciences. In: *The Genetics of the Pig, 2nd Edition*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 426-444.
- Lagos-Quintana M., Rauhut R., Lendeckel W. & Tuschl T. (2001) Identification of novel genes coding for small expressed RNAs. *Science (New York, N.Y.)* 294, 853-858.

- Lall S., Grun D., Krek A., Chen K., Wang Y.L., Dewey C.N., Sood P., Colombo T., Bray N., Macmenamin P., Kao H.L., Gunsalus K.C., Pachter L., Piano F. & Rajewsky N. (2006) A genome-wide map of conserved microRNA targets in *C. elegans*. *Current biology : CB* 16, 460-471.
- Larson G., Cucchi T. & Dobney K. (2011) Genetic Aspects of Pig Domestication. In: *The Genetics of the Pig, 2nd Edition*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 14-37.
- Larson G., Liu R., Zhao X., Yuan J., Fuller D., Barton L., Dobney K., Fan Q., Gu Z., Liu X.H., Luo Y., Lv P., Andersson L. & Li N. (2010) Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America* 107, 7686-7691.
- Larson G., Albarella U., Dobney K., Rowley-Conwy P., Schibler J., Tresset A., Vigne J.D., Edwards C.J., Schlumbaum A., Dinu A., Balacescu A., Dolman G., Tagliacozzo A., Manaseryan N., Miracle P., Van Wijngaarden-Bakker L., Masseti M., Bradley D.G. & Cooper A. (2007a) Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences of the United States of America* 104, 15276-15281.
- Larson G., Cucchi T., Fujita M., Matisoo-Smith E., Robins J., Anderson A., Rolett B., Spriggs M., Dolman G., Kim T.H., Thuy N.T., Randi E., Doherty M., Due R.A., Boltt R., Djubiantono T., Griffin B., Intoh M., Keane E., Kirch P., Li K.T., Morwood M., Pedrina L.M., Piper P.J., Rabett R.J., Shooter P., Van den Bergh G., West E., Wickler S., Yuan J., Cooper A. & Dobney K. (2007b) Phylogeny and ancient DNA of *Sus* provides insights into neolithic expansion in Island Southeast Asia and Oceania. *Proceedings of the National Academy of Sciences of the United States of America* 104, 4834-4839.
- Larson G., Dobney K., Albarella U., Fang M., Matisoo-Smith E., Robins J., Lowden S., Finlayson H., Brand T., Willerslev E., Rowley-Conwy P., Andersson L. & Cooper A. (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science (New York, N.Y.)* 307, 1618-1621.
- Lau N.C., Lim L.P., Weinstein E.G. & Bartel D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science (New York, N.Y.)* 294, 858-862.
- Lawrence J. (1727) *A New System of Agriculture*. T. Woodward, Dublin.
- Lee R.C. & Ambros V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science (New York, N.Y.)* 294, 862-864.
- Lee R.C., Feinbaum R.L. & Ambros V. (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-854.

- Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P., Radmark O., Kim S. & Kim V.N. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-419.
- Lee Y., Jeon K., Lee J.T., Kim S. & Kim V.N. (2002) MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO journal* 21, 4663-4670.
- Levine M. & Tjian R. (2003) Transcription regulation and animal diversity. *Nature* 424, 147-151.
- Li J.Y., Yong T.Y., Michael M.Z. & Gleadle J.M. (2010) Review: The role of microRNAs in kidney disease. *Nephrology (Carlton, Vic.)* 15, 599-608.
- Li S., Ran X.Q., Xu L. & Wang J.F. (2011) microRNA and mRNA expression profiling analysis of dichlorvos cytotoxicity in porcine kidney epithelial PK15 cells. *DNA and cell biology* 30, 1073-1083.
- Lian C., Sun B., Niu S., Yang R., Liu B., Lu C., Meng J., Qiu Z., Zhang L. & Zhao Z. (2012) A comparative profile of the microRNA transcriptome in immature and mature porcine testes using Solexa deep sequencing. *The FEBS journal*.
- Liang P. & Pardee A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science (New York, N.Y.)* 257, 967-971.
- Lim L.P., Glasner M.E., Yekta S., Burge C.B. & Bartel D.P. (2003) Vertebrate microRNA genes. *Science (New York, N.Y.)* 299, 1540.
- Lin W.C., Li S.C., Lin W.C., Shin J.W., Hu S.N., Yu X.M., Huang T.Y., Chen S.C., Chen H.C., Chen S.J., Huang P.J., Gan R.R., Chiu C.H. & Tang P. (2009) Identification of microRNA in the protist *Trichomonas vaginalis*. *Genomics* 93, 487-493.
- Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C(T)}$  Method. *Methods (San Diego, Calif.)* 25, 402-408.
- Luetkemeier E.S., Sodhi M., Schook L.B. & Malhi R.S. (2010) Multiple Asian pig origins revealed through genomic analyses. *Molecular phylogenetics and evolution* 54, 680-686.
- Lunney J.K. (2007) Advances in swine biomedical model genomics. *International journal of biological sciences* 3, 179-184.
- Lush J.L. (1943) *Animal Breeding Plans*. The Iowa State College Press, Ames.
- Malcolmson R. & Mastoris S. (1998) *The English Pig: A History*. Hambledon Press, London.

- Manunza A., Zidi A., Yeghoyan S., Balteanu V.A., Carsai T.C., Scherbakov O., Ramirez O., Eghbalsaied S., Castello A., Mercade A. & Amills M. (2013) A high throughput genotyping approach reveals distinctive autosomal genetic signatures for European and Near Eastern wild boar. *PLoS one* 8, e55891.
- Maragkakis M., Alexiou P., Papadopoulos G.L., Reczko M., Dalamagas T., Giannopoulos G., Goumas G., Koukis E., Kourtis K., Simossis V.A., Sethupathy P., Vergoulis T., Koziris N., Sellis T., Tsanakas P. & Hatzigeorgiou A.G. (2009a) Accurate microRNA target prediction correlates with protein repression levels. *BMC bioinformatics* 10, 295.
- Maragkakis M., Reczko M., Simossis V.A., Alexiou P., Papadopoulos G.L., Dalamagas T., Giannopoulos G., Goumas G., Koukis E., Kourtis K., Vergoulis T., Koziris N., Sellis T., Tsanakas P. & Hatzigeorgiou A.G. (2009b) DIANA-microT web server: elucidating microRNA functions through target prediction. *Nucleic acids research* 37, W273-6.
- Mardis E.R. (2008) Next-generation DNA sequencing methods. *Annual review of genomics and human genetics* 9, 387-402.
- Margulies M., Egholm M., Altman W.E., Attiya S., Bader J.S., Bemben L.A., Berka J., Braverman M.S., Chen Y.J., Chen Z., Dewell S.B., Du L., Fierro J.M., Gomes X.V., Godwin B.C., He W., Helgesen S., Ho C.H., Irzyk G.P., Jando S.C., Alenquer M.L., Jarvie T.P., Jirage K.B., Kim J.B., Knight J.R., Lanza J.R., Leamon J.H., Lefkowitz S.M., Lei M., Li J., Lohman K.L., Lu H., Makhijani V.B., McDade K.E., McKenna M.P., Myers E.W., Nickerson E., Nobile J.R., Plant R., Puc B.P., Ronan M.T., Roth G.T., Sarkis G.J., Simons J.F., Simpson J.W., Srinivasan M., Tartaro K.R., Tomasz A., Vogt K.A., Volkmer G.A., Wang S.H., Wang Y., Weiner M.P., Yu P., Begley R.F. & Rothberg J.M. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.
- Mata J., Marguerat S. & Bahler J. (2005) Post-transcriptional control of gene expression: a genome-wide perspective. *Trends in biochemical sciences* 30, 506-514.
- Memczak S., Jens M., Elefsinioti A., Torti F., Krueger J., Rybak A., Maier L., Mackowiak S.D., Gregersen L.H., Munschauer M., Loewer A., Ziebold U., Landthaler M., Kocks C., le Noble F. & Rajewsky N. (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333-338.
- Mestdagh P., Van Vlierberghe P., De Weer A., Muth D., Westermann F., Speleman F. & Vandesompele J. (2009) A novel and universal method for microRNA RT-qPCR data normalization. *Genome biology* 10, R64.
- Metzker M.L. (2010) Sequencing technologies - the next generation. *Nature reviews.Genetics* 11, 31-46.



- Meuwissen T.H., Hayes B.J. & Goddard M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819-1829.
- Meyers B.C., Axtell M.J., Bartel B., Bartel D.P., Baulcombe D., Bowman J.L., Cao X., Carrington J.C., Chen X., Green P.J., Griffiths-Jones S., Jacobsen S.E., Mallory A.C., Martienssen R.A., Poethig R.S., Qi Y., Vaucheret H., Voinnet O., Watanabe Y., Weigel D. & Zhu J.K. (2008) Criteria for annotation of plant MicroRNAs. *The Plant Cell* 20, 3186-3190.
- Morin R.D., O'Connor M.D., Griffith M., Kuchenbauer F., Delaney A., Prabhu A.L., Zhao Y., McDonald H., Zeng T., Hirst M., Eaves C.J. & Marra M.A. (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome research* 18, 610-621.
- Nielsen M., Hansen J.H., Hedegaard J., Nielsen R.O., Panitz F., Bendixen C. & Thomsen B. (2010) MicroRNA identity and abundance in porcine skeletal muscles determined by deep sequencing. *Animal Genetics* 41, 159-168.
- Oba S., Kumano S., Suzuki E., Nishimatsu H., Takahashi M., Takamori H., Kasuya M., Ogawa Y., Sato K., Kimura K., Homma Y., Hirata Y. & Fujita T. (2010) miR-200b precursor can ameliorate renal tubulointerstitial fibrosis. *PLoS one* 5, e13614.
- Ogata H., Goto S., Sato K., Fugibuchi W., Bono H. & Kanehisa M. (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 27, 29-34.
- Okamura K., Phillips M.D., Tyler D.M., Duan H., Chou Y.T. & Lai E.C. (2008) The regulatory activity of microRNA\* species has substantial influence on microRNA and 3' UTR evolution. *Nature structural & molecular biology* 15, 354-363.
- Orphanides G. & Reinberg D. (2002) A unified theory of gene expression. *Cell* 108, 439-451.
- Ottoni C., Flink L.G., Evin A., Georg C., De Cupere B., Van Neer W., Bartosiewicz L., Linderholm A., Barnett R., Peters J., Decorte R., Waelkens M., Vanderheyden N., Ricaut F.X., Cakirlar C., Cevik O., Hoelzel A.R., Mashkour M., Karimlu A.F., Seno S.S., Daujat J., Brock F., Pinhasi R., Hongo H., Perez-Enciso M., Rasmussen M., Frantz L., Megens H.J., Crooijmans R., Groenen M., Arbuckle B., Benecke N., Vidarsdottir U.S., Burger J., Cucchi T., Dobney K. & Larson G. (2013) Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics. *Molecular biology and evolution* 30, 824-832.
- Overton M. (1996) *Agricultural Revolution in England: The Transformation of the Agrarian Economy 1500-1850*. Cambridge University Press, Cambridge.
- Pareek C.S., Smoczynski R. & Tretyn A. (2011) Sequencing technologies and genome sequencing. *Journal of Applied Genetics* 52, 413-435.

- Pasquinelli A.E., Reinhart B.J., Slack F., Martindale M.Q., Kuroda M.I., Maller B., Hayward D.C., Ball E.E., Degnan B., Muller P., Spring J., Srinivasan A., Fishman M., Finnerty J., Corbo J., Levine M., Leahy P., Davidson E. & Ruvkun G. (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408, 86-89.
- Patel N. & Sauter E.R. (2011) Body fluid micro(mi)RNAs as biomarkers for human cancer. *Journal of Nucleic Acids Investigation* 2:e1.
- Pejsak Z. & Truszczyński M. (2006) Aujeszky's disease (Pseudorabies). In: *Diseases of Swine, Ninth Edition*. (ed. by B. Straw, J. Zimmerman, S. D'Allaire & D. Taylor), pp. 419-433. Blackwell Science, Oxford, UK.
- Peltier H.J. & Latham G.J. (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA (New York, N.Y.)* 14, 844-852.
- Pfaffl M.W. (2004) Quantification Strategies in Real-Time PCR. In: *A-Z of Quantitative PCR*. (ed. by S.A. Bustin), pp. 87-112. International University Line, La Jolla.
- Pfaffl M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* 29, e45.
- Pfeffer S., Zavolan M., Grasser F.A., Chien M., Russo J.J., Ju J., John B., Enright A.J., Marks D., Sander C. & Tuschl T. (2004) Identification of virus-encoded microRNAs. *Science (New York, N.Y.)* 304, 734-736.
- Podolska A., Anthon C., Bak M., Tommerup N., Skovgaard K., Heegaard P.M., Gorodkin J., Cirera S. & Fredholm M. (2012) Profiling microRNAs in lung tissue from pigs infected with *Actinobacillus pleuropneumoniae*. *BMC genomics* 13, 459-2164-13-459.
- Pomeranz L.E., Reynolds A.E. & Hengartner C.J. (2005) Molecular biology of pseudorabies virus: impact on neurovirology and veterinary medicine. *Microbiology and molecular biology reviews : MMBR* 69, 462-500.
- Pourmand N., Karhanek M., Persson H.H., Webb C.D., Lee T.H., Zahradnikova A. & Davis R.W. (2006) Direct electrical detection of DNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 103, 6466-6470.
- Pritchard C.C., Cheng H.H. & Tewari M. (2012) MicroRNA profiling: approaches and considerations. *Nature reviews.Genetics* 13, 358-369.
- Ramirez O., Ojeda A., Tomas A., Gallardo D., Huang L.S., Folch J.M., Clop A., Sanchez A., Badaoui B., Hanotte O., Galman-Omitogun O., Makuza S.M., Soto H., Cadillo J., Kelly L., Cho I.C., Yeghoyan S., Perez-Enciso M. & Amills M. (2009) Integrating Y-chromosome, mitochondrial, and autosomal data to analyze the origin of pig breeds. *Molecular biology and evolution* 26, 2061-2072.

- Reddy A.M., Zheng Y., Jagadeeswaran G., Macmil S.L., Graham W.B., Roe B.A., Desilva U., Zhang W. & Sunkar R. (2009) Cloning, characterization and expression analysis of porcine microRNAs. *BMC genomics* 10, 65.
- Reinhart B.J., Slack F.J., Basson M., Pasquinelli A.E., Bettinger J.C., Rougvie A.E., Horvitz H.R. & Ruvkun G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901-906.
- Ribeiro-dos-Santos A., Khayat A.S., Silva A., Alencar D.O., Lobato J., Luz L., Pinheiro D.G., Varuzza L., Assumpcao M., Assumpcao P., Santos S., Zanette D.L., Silva W.A., Jr, Burbano R. & Darnet S. (2010) Ultra-deep sequencing reveals the microRNA expression pattern of the human stomach. *PLoS one* 5, e13205.
- Rosenberg M. & Redding R. (1998) Early Pig Husbandry in Southwestern Asia and Its Implications for Modeling the Origins of Food Production. In: *Ancestors for the Pigs: Pigs in Prehistory*. (ed. by S. Nelson), pp. 55.
- Rothberg J.M., Hinz W., Rearick T.M., Schultz J., Mileski W., Davey M., Leamon J.H., Johnson K., Milgrew M.J., Edwards M., Hoon J., Simons J.F., Marran D., Myers J.W., Davidson J.F., Branting A., Nobile J.R., Puc B.P., Light D., Clark T.A., Huber M., Branciforte J.T., Stoner I.B., Cawley S.E., Lyons M., Fu Y., Homer N., Sedova M., Miao X., Reed B., Sabina J., Feierstein E., Schorn M., Alanjary M., Dimalanta E., Dressman D., Kasinskas R., Sokolsky T., Fidanza J.A., Namsaraev E., McKernan K.J., Williams A., Roth G.T. & Bustillo J. (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475, 348-352.
- Ruan Y., Le Ber P., Ng H.H. & Liu E.T. (2004) Interrogating the transcriptome. *Trends in biotechnology* 22, 23-30.
- Ruvinsky A., Rothschild M.F., Larson G. & Gongora J. (2011) Systematics and Evolution of the Pig. In: *The Genetics of the Pig, 2nd Edition*. (ed. by M.F. Rothschild & A. Ruvinsky), pp. 1-13.
- Sanger F., Nicklen S. & Coulson A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74, 5463-5467.
- Sanger F. & Coulson A.R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* 94, 441-448.
- Schaefer A., Jung M., Miller K., Lein M., Kristiansen G., Erbersdobler A. & Jung K. (2010) Suitable reference genes for relative quantification of miRNA expression in prostate cancer. *Experimental & molecular medicine* 42, 749-758.
- Schena M., Shalon D., Davis R.W. & Brown P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, N.Y.)* 270, 467-470.

- Schmittgen T.D. & Livak K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols* 3, 1101-1108.
- Schroeder A., Mueller O., Stocker S., Salowsky R., Leiber M., Gassmann M., Lightfoot S., Menzel W., Granzow M. & Ragg T. (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology* 7, 3.
- Shao N.Y., Hu H.Y., Yan Z., Xu Y., Hu H., Menzel C., Li N., Chen W. & Khaitovich P. (2010) Comprehensive survey of human brain microRNA by deep sequencing. *BMC genomics* 11, 409.
- Sharbati S., Friedlander M.R., Sharbati J., Hoeke L., Chen W., Keller A., Stahler P.F., Rajewsky N. & Einspanier R. (2010) Deciphering the porcine intestinal microRNA transcriptome. *BMC genomics* 11, 275.
- Sharkey J.W., Antoine D.J. & Park B.K. (2012) Validation of the isolation and quantification of kidney enriched miRNAs for use as biomarkers. *Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals*.
- Shen Y., Li Y., Ye F., Wang F., Wan X., Lu W. & Xie X. (2011) Identification of miR-23a as a novel microRNA normalizer for relative quantification in human uterine cervical tissues. *Experimental & molecular medicine* 43, 358-366.
- Shendure J. & Ji H. (2008) Next-generation DNA sequencing. *Nature biotechnology* 26, 1135-1145.
- Shi R. & Chiang V.L. (2005) Facile means for quantifying microRNA expression by real-time PCR. *BioTechniques* 39, 519-525.
- Shi W., Hendrix D., Levine M. & Haley B. (2009) A distinct class of small RNAs arises from pre-miRNA-proximal regions in a simple chordate. *Nature Structural & Molecular Biology, Nat Struct Mol Biol* 2, 183-189.
- Simoons F.J. (1991) *Food in China: A Cultural and Historical Inquiry*. CRC Press, Boca Raton.
- Skalsky R.L. & Cullen B.R. (2010) Viruses, microRNAs, and host interactions. *Annual Review of Microbiology* 64, 123-141.
- Stranneheim H. & Lundberg J. (2012) Stepping stones in DNA sequencing. *Biotechnology journal* 7, 1063-1073.
- Swanson K.S., Mazur M.J., Vashisht K., Rund L.A., Beever J.E., Counter C.M. & Schook L.B. (2004) Genomics and clinical medicine: rationale for creating and effectively evaluating animal models. *Experimental biology and medicine (Maywood, N.J.)* 229, 866-875.

- Taberlet P., Fumagalli L., Wust-Saucy A.G. & Cosson J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular ecology* 7, 453-464.
- Tarver J.E., Donoghue P.C. & Peterson K.J. (2012) Do miRNAs have a deep evolutionary history? *BioEssays : news and reviews in molecular, cellular and developmental biology* 34, 857-866.
- tenOever B.R. (2013) RNA viruses and the host microRNA machinery. *Nature reviews.Microbiology* 11, 169-180.
- Trow-Smith R. (2006) A History of British Livestock Husbandry 1700-1900. Routledge.
- Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A. & Speleman F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* 3, RESEARCH0034.
- Veedu R.N., Vester B. & Wengel J. (2007) Enzymatic incorporation of LNA nucleotides into DNA strands. *Chembiochem : a European journal of chemical biology* 8, 490-492.
- Velculescu V.E., Zhang L., Vogelstein B. & Kinzler K.W. (1995) Serial analysis of gene expression. *Science (New York, N.Y.)* 270, 484-487.
- Vickers K.C., Sethupathy P., Baran-Gale J. & Remaley A.T. (2013) Complexity of microRNA function and the role of isomiRs in lipid homeostasis. *Journal of lipid research* 54, 1182-1191.
- Walters E.M., Wolf E., Whyte J.J., Mao J., Renner S., Nagashima H., Kobayashi E., Zhao J., Wells K.D., Critser J.K., Riley L.K. & Prather R.S. (2012) Completion of the swine genome will simplify the production of swine as a large animal biomedical model. *BMC medical genomics* 5, 55-8794-5-55.
- Watanobe T., Okumura N., Ishiguro N., Nakano M., Matsui A., Sahara M. & Komatsu M. (1999) Genetic relationship and distribution of the Japanese wild boar (*Sus scrofa leucomystax*) and Ryukyu wild boar (*Sus scrofa riukiuanus*) analysed by mitochondrial DNA. *Molecular ecology* 8, 1509-1512.
- White S. (2011) From Globalized Pig Breeds to Capitalist Pigs: a Study in Animal Cultures and Evolutionary History. *Environmental History* 16, 94-120.
- Wightman B., Ha I. & Ruvkun G. (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862.

- Woese C.R., Kandler O. & Wheelis M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America* 87, 4576-4579.
- Wotschovsky Z., Meyer H.A., Jung M., Fendler A., Wagner I., Stephan C., Busch J., Erbersdobler A., Disch A.C., Mollenkopf H.J. & Jung K. (2011) Reference genes for the relative quantification of microRNAs in renal cell carcinomas and their metastases. *Analytical Biochemistry* 417, 233-241.
- Wu G.S., Yao Y.G., Qu K.X., Ding Z.L., Li H., Palanichamy M.G., Duan Z.Y., Li N., Chen Y.S. & Zhang Y.P. (2007) Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. *Genome biology* 8, R245.
- Wu Y.Q., Chen D.J., He H.B., Chen D.S., Chen L.L., Chen H.C. & Liu Z.F. (2012) Pseudorabies virus infected porcine epithelial cell line generates a diverse set of host microRNAs and a special cluster of viral microRNAs. *PloS one* 7, e30988.
- Xie S.S., Li X.Y., Liu T., Cao J.H., Zhong Q. & Zhao S.H. (2011) Discovery of porcine microRNAs in multiple tissues by a Solexa deep sequencing approach. *PloS one* 6, e16235.
- Yamey G. (2000) Scientists unveil first draft of human genome. *BMJ (Clinical research ed.)* 321, 7.
- Yi R., Qin Y., Macara I.G. & Cullen B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & development* 17, 3011-3016.
- Yuan J. & Flad R. (2002) Pig Domestication in Ancient China. *Antiquity* 76, 724-732.
- Zhang B., Kirov S. & Snoddy J. (2005) WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research* 33, W741-8.
- Zheng Y., Cai X. & Bradley J.E. (2013) microRNAs in parasites and parasite infection. *RNA biology* 10.
- Zuker M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* 31, 3406-3415.

## AGRAÏMENTS

---

La teoria diu que en aquest apartat s'escriuen unes paraules d'agraïment a tothom qui d'alguna manera o altra t'ha ensenyat, t'ha inspirat, t'ha ajudat o t'ha animat a arribar al final d'aquest camí. La realitat ensenya que queda més polit fer un recordatori de totes aquelles persones i les coses positives que n'has après, ja siguin professionals o personals. I és que tot i ser la tesi d'un estudiant de doctorat, moltes persones han contribuït a que ara mateix estigui escrivint aquestes últimes pàgines de la tesi. Moltes gràcies a tots!

Primer de tot vull agrair tots els coneixements i la confiança que el doctor **Armand Sánchez**, el director de la tesi, ha dipositat en mi durant tot aquest temps. També vull agrair-li totes les oportunitats que m'ha ofert per assistir i participar a congressos, conferències i cursos per poder viure en primera persona la ciència que es fa arreu del món.

Mereix una dedicatòria especial la doctora **Anna Castelló**, la meva *mare* adoptiva al laboratori. T'has preocupat per mi i m'has format com a científic i com a persona, i tots els becaris sabem com és d'important aquesta figura quan fas una tesi. Moltes gràcies Anna, ja saps el que t'estic d'agraït per tota la feina que has fet en mi i, de vegades, tota la paciència que t'ha calgut 😊.

He de confessar que he tingut una gran sort que la meva etapa d'estudiant de doctorat hagi coincidit amb una altra estudiant, ara ja doctora. Amiga **Ingrid Balcells** (amiga va abans que doctora), estic molt, molt i molt content d'haver compartit amb tu tot el que hem viscut. M'has ensenyat tot el que has pogut i més, tot el que a biotecnologia teniu més que après i a veterinària anem una mica més justos. M'has ensenyat que un text es pot capgirar de dat a baix una vegada i una altra, i també m'has ensenyat a escatxigar i a embogonar. Moltes gràcies per haver pogut fer equip amb tu, de debò. Espero que algun dia tinguem la oportunitat de repetir-ho 😊.

També he tingut el plaer de compartir l'etapa doctoral amb en **Xavi Quílez**. Moltes gràcies per tots els moments que hem compartit (al laboratori no gaires 😊). Sovint,

molts dels grans avenços en el camí d'un pre-doc passen fora del laboratori i tu has fet que sigui capaç de veure'ls. També m'has demostrat que en 5 minuts un és capaç de posar-se a dormir, sigui quina sigui la situació! Els grans moments als congressos i conferències (i post congressos i post conferències) juntament amb l'Ingrid són els que guardarem sempre de record.

Vull donar les gràcies a la doctora **Natàlia Sastre**. Les complicitats que hem creat i tot el que m'has ensenyat i aconsellat també formen part d'aquesta tesi i de la meva persona. La teva vitalitat ajuda molt en els moments més complicats del camí. Moltes gràcies per la teva visió (d'una dècada més gran) i tots els secrets que hem compartit!  
😊

També agrair haver-me topat amb l'energia positiva i el caràcter decidit de la futura doctora **Verónica Martínez**. No saps la de vegades que he pensat "tots hauríem de tenir, com a mínim, la meitat de la mentalitat incisiva que té la Vero. Tot seria molt millor" 😊. M'emporto molts bons moments compartits amb tu que espero que es segueixin repetint! Et desitjo el millor en la teva nova etapa personal i també professional.

La meva formació científica també ha estat possible gràcies als doctors **Marcel Amills**, **Àlex Clop**, **Sebastián Ramos**, **Josep Maria Folch**, **Jordi Jordana**, **Jesús Piedrafita** i **Miguel Pérez-Enciso**, que amb els seus coneixements transmesos a les classes de doctorat, seminaris i consells m'han ajudat a assolir els coneixements necessaris per desenvolupar la tesi. També un especial agraïment a la doctora **Anna Tomàs**, per ensenyar-me i aconsellar-me tant a l'inici com al final de la tesi, i al doctor **José Ignacio Núñez**, del Centre de Recerca en Sanitat Animal, per tota la formació en matèria vírica.

Vull donar les gràcies també a tots els companys i companyes amb qui he compartit dinars, xerrades, excursions i ambient de grup: a la doctora **Anna Mercadé**, per les converses i xerrades mentre els *fidusials* sortien correctes. Gràcies a l'**Anna Esteve** (ja doctora), per les etapes de bogeria a l'inici de la tesi, inoblidables. A l'**Ali Zidi** (ja doctor), per les xerrades intentant arreglar el planeta, tant a nivell humà com espiritual. Al **Jordi Coromines**, per discutir una vegada i una altra la dualitat Espanyol-Barça (mira quin equip he posat primer). A l'**Anna Puig**, per divagar sobre els grans



misteris de la vida. Gràcies al suport del **Yulixis Ramayo**, d'aquí a res serà el teu dia! I gràcies també al **William Burgos**, a la doctora **Arianna Manunza** i la doctora **Maria Ballester** pels moments compartits.

També vull agrair els moments compartits amb els companys que em van acollir a l'inici de la tesi i que ja són fora: a la **María Salinas**, l'**Abid Badaui**, el **Jordi Estellé** i a l'**Anna Ojeda**. I també vull donar les gràcies a la nova fornada de futurs doctors que pugem al departament. A la **Sarai Córdoba**, a l'**Erica Bianco**, al **Manuel Revilla**, al **Jorge Petroli**, a la **Sandra Contreras** i a l'**Antonia Noce**. Moltes gràcies Sarai pels ànims i pel treball conjunt en el projecte, i també moltes gràcies pel disseny de la portada, ets una artista!

Crec, sincerament, que el clima de confiança que he mantingut amb tu, **Betlem Cabrera**, m'ha ajudat a no perdre l'alegria en moments importants. No perdis aquesta capacitat, perquè per les faltes d'ortografia sempre podràs mirar un diccionari.

Moltes gràcies també a la **Maria del Mar Barrera** per les hores compartides al despatx, i a la **Tània Guil**, perquè agilitzar i solucionar tràmits en el feixuc món de la burocràcia del becari és vital.

A tots el companys i companyes del Servei Veterinari de Genètica Molecular i de VetGenomics. A les doctores **Olga Francino**, la **Laura Altet** i **Elisenda Sánchez**, pels consells i els moments compartits. També agrair els riures i riures compartits amb la **Rebeca Turmo**, la **Griselda Bassols**, la **Carme Sagrera**, la **Lorena Serrano** i el **David Gallardo**.

També m'agradaria dedicar un paràgraf a l'aventura australiana del 2010. Moltes gràcies al doctor **Toni Reverter** per haver-me acollit al CSIRO a la Universitat de Queensland i haver pogut treballar amb tot el teu equip. I moltes gràcies al **Bernat Coll**, al **Ramón Ganigué** i a l'**Anna Pedró**. 3 mesos inoblidables!

Ja fora de l'àmbit científic, voldria agrair la paciència immensa de l'**Anna**, la **Carla** i la **Irene**. Sabeu que sóc el que sóc en part per vosaltres i no me'n penedeixo pas. Poder viure tot aquest temps amb vosaltres a Ca la Quima és una de les millors decisions que he pres mai! Ja fa molts anys que ens coneixem i sempre que us he necessitat heu

estat allà per aconsellar-me, animar-me o picar-me les mans, us estimo molt. Vull fer extensiu aquest agraïment a tota la colla del cau/associació/federació (SKUES/AEC/FCEG) que m'han ajudat a compaginar la feina que comporta fer una tesi amb tot el treball associatiu voluntari que sempre he intentat fer el millor que he pogut, en especial a la resta del grup: **Samantha, Maria, Samu, Agnès, Carles, Micky i Natàlia** i a la última junta d'Acció de la que he format part: **Aida, Anna, Míriam, Marta i Pili (#JuntaStyle)**.

Unes gràcies immenses a les nenes de la carrera! **Marina, Iris, Laura, Glòria, Carla, Tíscar i Georgina**. Per cuidar-me tant entre tanta dona i compartir les alegries i les penes del món de la veterinària.

Gràcies també als amics del "cole" de sempre, els de L'Hospi ☺: a l'**Alba**, al **Sergio**, al **Dani**, al **Sergi** i a l'**Helena**. Pels moments de complicitat que encara duren i poder veure com ens fem grans i cadascú tira pel seu camí compartint també els dels altres.

I, finalment, vull donar les gràcies a qui han invertit molts anys d'esforços per a que el seu fill pugui posar la paraula doctor davant del seu nom. Segurament no us ho haguéssiu imaginat mai (ni jo), però si he arribat fins aquí és perquè vosaltres, **Coral** i **Joan**, m'heu educat amb els valors de l'esforç i la perseverança. Moltes gràcies per recolzar-me en tot allò que he volgut fer fins ara, us agradés més o menys. Estic convençut que sempre ho fareu. I també vull donar les gràcies al meu germanet **Arnau**, per què tot i ser una mica pesat (sovint), a mesura que passen els anys em demostres que sempre podré confiar en tu.



