Genomic and Functional Approaches to Genetic Adaptation

Elena Carnero Montoro

TESI DOCTORAL UPF / 2013

Thesis Director

Dra. ELENA BOSCH

DEPARTAMENT DE CIÈNCIES EXPERIMENTALS I DE LA SALUT

Fitxer PDF de la tesi dividit en 7 parts

Part 1 de 7	pàg 0 - 17	Introduction Cap. 1 – Cap 2
Part 2 de 7	pàg 18 - 20	Introduction Cap. 3 : 3.1, 3.2
Part 3 de 7	pàg. 20 – 25	Introduction Cap. 3 : 3.2.1
Part 4 de 7	pàg. 25 – 28	Introduction Cap. 3: 3.2.2 – 3.2.3
Part 5 de 7	pàg. 29 - 64	Introduction Cap. 4 – Cap. 6
Part 6 de 7	pàg. 65 - 183	Objectives, Results, Discussion, Concluding remarks
Part 7 de 7	pàg. 184 – 233	References, Annexes

Part 7 de 7

References

Abecasis, G.R. et al., 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422), pp.56–65. Available at:

Akey, J.M., 2009. Constructing genomic maps of positive selection in humans: where do we go from here? *Genome research*, 19(5), pp.711–22.

Altshuler, D.M. et al., 2010. Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311), pp.52–8.

Alves, I. et al., 2012. Genomic data reveal a complex making of humans. M. H. Schierup, ed. *PLoS genetics*, 8(7), p.e1002837.

Andrés, A.M. et al., 2010. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *PLoS genetics*, 6(10), p.e1001157

Haldane, J.B.S., The Causes of Evolution, 1932.

Baker, M.A. et al., 2012. Polymorphisms in the gene that encodes the iron transport protein ferroportin 1 influence susceptibility to tuberculosis.

Bamshad, M. & Wooding, S.P., 2003. Signatures of natural selection in the human genome. *Nature reviews. Genetics*, 4(2), pp.99–111.

Barreiro, L.B. et al., 2008. Natural selection has driven population differentiation in modern humans. *Nature genetics*, 40(3), pp.340–5.

Barreiro, L.B. & Quintana-Murci, Lluís, 2010. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nature reviews. Genetics*, 11(1), pp.17–30.

Barrett, R.D.H. & Hoekstra, H.E., 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nature reviews. Genetics*, 12(11), pp.767–80.

Beleza, S. et al., 2013. The timing of pigmentation lightening in Europeans. *Molecular biology and evolution*, 30(1), pp.24–35.

Bethony, J. et al., 2006. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *The Lancet*, 367(9521), pp.1521–1532.

Boyko, A.R. et al., 2008. Assessing the evolutionary impact of amino acid mutations in the human genome.

M. H. Schierup, ed. *PLoS genetics*, 4(5), p.e1000083. Cagliani, R. et al., 2013. Crohn's disease Loci are common targets of protozoa-driven selection. *Molecular biology and evolution*, 30(5), pp.1077–87.

Cagliani, R. & Sironi, M., 2013. Pathogen-driven selection in the human genome. *International journal of evolutionary biology*, 2013, p.204240.

Cai, J.J. et al., 2009. Pervasive hitchhiking at coding and regulatory sites in humans. *PLoS genetics*, 5(1), p.e1000336.

Calafell, F. et al., 2008. Evolutionary dynamics of the human ABO gene. *Human genetics*, 124(2), pp.123–35. Available at:

Casals, F. et al., 2011. Genetic adaptation of the antibacterial human innate immunity network. *BMC evolutionary biology*, 11(1), p.202.

Cole, C.R. et al., 2010. Zinc and iron deficiency and their interrelations in low-income African American and Hispanic children in Atlanta 1 – 4. *American Journal of Clinical Nutrition*, (C).

Coop, G. et al., 2009. The role of geography in human adaptation. *PLoS genetics*, 5(6), p.e1000500.

Dall'Olio, G.M. et al., 2012. Distribution of events of positive selection and population differentiation in a metabolic pathway: the case of asparagine N-glycosylation. *BMC evolutionary biology*, 12, p.98.

Dalloul, A., 2009. Autoimmunity Reviews CD5 : A safeguard against autoimmunity and a shield for cancer cells. *Autoimmunity Reviews*, 8(4), pp.349–353.

Daub, J.T. et al., 2013. Evidence for Polygenic Adaptation to Pathogens in the Human Genome. *Molecular biology and evolution*.

DiLillo, D.J., Matsushita, T. & Tedder, T.F., 2010. B10 cells and regulatory B cells balance immune responses during inflammation, autoimmunity, and cancer. *Annals of the New York Academy of Sciences*, 1183, pp.38–57.

Dunham, I. et al., 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), pp.57–74. A

Eike, M.C. et al., 2012. CIITA gene variants are associated with rheumatoid arthritis in Scandinavian populations. *Genes and immunity*, 13(5), pp.431–6.

Esteve-Codina, A. et al., 2013. Dissecting structural and nucleotide genome-wide variation in inbred Iberian pigs. *BMC genomics*, 14(1), p.148.

Everitt, A.R. et al., 2012. IFITM3 restricts the morbidity and mortality associated with influenza. *Nature*, 484(7395), pp.519–23.

Eyre, S. et al., 2012. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nature genetics*, 44(12), pp.1336–40.

Fay, J C & Wu, C.I., 2000. Hitchhiking under positive Darwinian selection. *Genetics*, 155(3), pp.1405–13.

Fay, Justin C, 2011. Weighing the evidence for adaptation at the molecular level. *Trends in genetics : TIG*, 27(9), pp.343–9.

Ferrer-Admetlla, A. et al., 2009. A natural history of FUT2 polymorphism in humans. *Molecular biology and evolution*, 26(9), pp.1993–2003.

Ferrer-Admetlla, A. et al., 2008. Balancing selection is the main force shaping the evolution of innate immunity genes. *Journal of immunology (Baltimore, Md. : 1950)*, 181(2), pp.1315–22.

Franke, A. et al., 2010. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature genetics*, 42(12), pp.1118–25.

Fraser, H.B., 2013. Gene expression drives local adaptation in humans. *Genome research*, p.gr.152710.112–.

Fu, Y.X. & Li, W.H., 1993. Statistical tests of neutrality of mutations. *Genetics*, 133(3), pp.693–709.

Fumagalli, M. et al., 2009. Parasites represent a major selective force for interleukin genes and shape the genetic predisposition to autoimmune conditions. *The Journal of experimental medicine*, 206(6), pp.1395–408.

Fumagalli, M. et al., 2011. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS genetics*, 7(11), p.e1002355.

Galvani, A.P. & Slatkin, M., 2003. Evaluating plague and smallpox as historical selective pressures for the CCR5-Delta 32 HIV-resistance allele. *Proceedings of the National Academy of Sciences of the United States of America*, 100(25), pp.15276–9.

Geiser, J. et al., 2012. A mouse model of acrodermatitis enteropathica: loss of intestine zinc transporter ZIP4 (Slc39a4) disrupts the stem cell niche and intestine integrity. *PLoS genetics*, 8(6), p.e1002766.

Genovese, G. et al., 2010. A risk allele for focal segmental glomerulosclerosis in African Americans is located within a region containing APOL1 and MYH9. *Kidney international*, 78(7), pp.698–704.

Granka, J.M. et al., 2012. Limited evidence for classic selective sweeps in African populations. *Genetics*, 192(3), pp.1049–64.

Grossman, S.R. et al., 2010. A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science (New York, N.Y.)*, 327(5967), pp.883–6.

Grossman, S.R. et al., 2013. Identifying recent adaptations in large-scale genomic data. *Cell*, 152(4), pp.703–13.

Halligan, D.L. et al., 2010. Evidence for pervasive adaptive protein evolution in wild mice. *PLoS genetics*, 6(1), p.e1000825.

Hamblin, M.T. & Di Rienzo, A, 2000. Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. *American journal of human genetics*, 66(5), pp.1669–79.

Hay, S.I. et al., 2009. A world malaria map: Plasmodium falciparum endemicity in 2007. I. Mueller, ed. *PLoS medicine*, 6(3), p.e1000048.

Haygood, R. et al., 2010. Contrasts between adaptive coding and noncoding changes during human evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 107(17), pp.7853–7.

Haygood, R. et al., 2007. Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. *Nature genetics*, 39(9), pp.1140–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17694055 [Accessed January 29, 2013].

Hedrick, P.W., Whittam, T.S. & Parham, P., 1991. Heterozygosity at individual amino acid sites: extremely high levels for HLA-A and -B genes. *Proceedings of the National Academy of Sciences of the United States of America*, 88(13), pp.5897–901. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=51985&tool=pmcentrez&rendertype=abstract [Accessed May 26, 2013].

Hood, M.I. & Skaar, E.P., 2012. Nutritional immunity: transition metals at the pathogen-host interface. *Nature reviews. Microbiology*, 10(8), pp.525–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22796883 [Accessed May 26, 2013].

Hudson, R.R., Kreitman, M. & Aguadé, M., 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics*, 116(1), pp.153–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1203113&tool=pmcentrez&rendertype=abstract [Accessed May 24, 2013].

Hughes, A.L. & Nei, M., 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, 335(6186), pp.167–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3412472 [Accessed May 26, 2013].

Hvilsom, C. et al., 2011. Extensive X-linked adaptive evolution in central chimpanzees. , pp.1–6.

Kambe, T., Weaver, B.P. & Andrews, G.K., 2008. The genetics of essential metal homeostasis during development. *Genesis (New York, N.Y. : 2000)*, 46(4), pp.214–28. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2366893&tool=pmcentrez&rendertype=abstract [Accessed July 8, 2013].

Kehl-Fie, T.E. & Skaar, E.P., 2010. Nutritional immunity beyond iron: a role for manganese and zinc. *Current opinion in chemical biology*, 14(2), pp.218–24.

Kelley, J.L. et al., 2006. Genomic signatures of positive selection in humans and the limits of outlier approaches. *Genome research*, 16(8), pp.980–9.

Kelley, J.L. & Swanson, W.J., 2008. Positive selection in the human genome: from genome scans to biological significance. *Annual review of genomics and human genetics*, 9, pp.143–60.

Kimura, M., 1968. Evolutionary rate at the molecular level. *Nature*, 217(5129), pp.624–6.

King, M.C. & Wilson, A.C., 1975. Evolution at two levels in humans and chimpanzees. *Science (New York, N.Y.)*, 188(4184), pp.107–16.

Kircher, M. & Kelso, J., 2010. High-throughput DNA sequencing--concepts and limitations. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 32(6), pp.524–36.

Kosiol, C. et al., 2008. Patterns of positive selection in six Mammalian genomes. *PLoS genetics*, 4(8), p.e1000144.

Kousathanas, A. et al., 2011. Positive and negative selection on noncoding DNA close to protein-coding genes in wild house mice. *Molecular biology and evolution*, 28(3), pp.1183–91.

Lewontin, R.C. & Krakauer, J., 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74(1), pp.175–95.

Lozano, F et al., 2000. CD5 signal transduction: positive or negative modulation of antigen receptor signaling. *Critical reviews in immunology*, 20(4), pp.347–58.

Mackay, T.F.C. et al., 2012. The Drosophila melanogaster Genetic Reference Panel. *Nature*, 482(7384), pp.173–8.

Manry, Jérémy et al., 2011. Evolutionary genetic dissection of human interferons. *The Journal of experimental medicine*, 208(13), pp.2747–59.

Manry, Jérémy & Quintana-Murci, Lluis, 2012. [Population genetics and human immunity: the interferon paradigm]. *Médecine sciences : M/S*, 28(12), pp.1095–101.

Marques-Bonet, T., Ryder, O.A. & Eichler, E.E., 2009. Sequencing primate genomes: what have we learned? *Annual review of genomics and human genetics*, 10, pp.355–86.

Martínez, V.G. et al., 2011. The Conserved Scavenger Receptor Cysteine-Rich Superfamily in Therapy and Diagnosis. , 63(4), pp.967–1000.

Messer, P.W. & Petrov, D.A., 2013. Frequent adaptation and the McDonald-Kreitman test. *Proceedings of the National Academy of Sciences of the United States of America*, 110(21), pp.8615–20.

Montanucci, L. et al., 2011. Molecular evolution and network-level analysis of the N-glycosylation metabolic pathway across primates. *Molecular biology and evolution*, 28(1), pp.813–23.

Moreno-Estrada, A. et al., 2009. Interrogating 11 fast-evolving genes for signatures of recent positive selection in worldwide human populations. *Molecular biology and evolution*, 26(10), pp.2285–97.

Mostefaoui, Y. et al., 2004. Candida albicans and Streptococcus salivarius modulate IL-6, IL-8, and TNF-alpha expression and secretion by engineered human oral mucosa cells. *Cellular microbiology*, 6(11), pp.1085–96.

Nielsen, R., 2005. Molecular signatures of natural selection. *Annual review of genetics*, 39, pp.197–218.

Pagnier, J., 1984. Evidence for the Multicentric Origin of the Sickle Cell Hemoglobin Gene in Africa. *Proceedings of the National Academy of Sciences*, 81(6), pp.1771–1773.

Prado-Martinez, J. et al., 2013. Great ape genetic diversity and population history. *Nature*, 499(7459), pp.471–5. Pritchard, J.K., Pickrell, J.K. & Coop, G., 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current biology : CB*, 20(4), pp.R208–15.

Pritchard, J.K. & Di Rienzo, Anna, 2010. Adaptation - not by sweeps alone. *Nature reviews. Genetics*, 11(10), pp.665–7.

Quach, H. et al., 2013. Different selective pressures shape the evolution of Toll-like receptors in human and African great ape populations. *Human molecular genetics*.

Quintana-Murci, Lluís & Clark, Andrew G, 2013. Population genetic tools for dissecting innate immunity in humans. *Nature reviews. Immunology*, 13(4), pp.280–93.

Raj, T. et al., 2013. Common risk alleles for inflammatory diseases are targets of recent positive selection. *American journal of human genetics*, 92(4), pp.517–29.

Richard-Miceli, C. & Criswell, L.A., 2012. Emerging patterns of genetic overlap across autoimmune disorders. *Genome medicine*, 4(1), p.6.

Rink, L. & Haase, H., 2007. Zinc homeostasis and immunity. *Trends in immunology*, 28(1), pp.1–4.

Sabeti, P C et al., 2006. Positive natural selection in the human lineage. *Science (New York, N.Y.)*, 312(5780), pp.1614–20. Sabeti, Pardis C et al., 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419(6909), pp.832–7.

Sabeti, Pardis C et al., 2005. The case for selection at CCR5-Delta32. A. Clark, ed. *PLoS biology*, 3(11), p.e378.

Schaffner, Stephen F et al., 2005. Calibrating a coalescent simulation of human genome sequence variation. *Genome research*, 15(11), pp.1576–83.

Schneider, A. et al., 2011. A method for inferring the rate of occurrence and fitness effects of advantageous mutations. *Genetics*, 189(4), pp.1427–37.

Ségurel, L. et al., 2012. The ABO blood group is a trans-species polymorphism in primates. *Proceedings of the National Academy of Sciences of the United States of America*, 109(45), pp.18493–8.

Sellick, G.S. et al., 2013. Scan of 977 nonsynonymous SNPs in CLL4 trial patients for the identification of genetic variants influencing prognosis Scan of 977 nonsynonymous SNPs in CLL4 trial patients for the identification of genetic variants influencing prognosis. , pp.1625–1633.

Serra, F. et al., 2011. Natural selection on functional modules, a genome-wide analysis. *PLoS computational biology*, 7(3), p.e1001093.

Siepel, A. et al., 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome research*, 15(8), pp.1034–50.

Simarro, P.P. et al., 2010. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *International journal of health geographics*, 9, p.57.

Sironi, M. & Clerici, M., 2010. The hygiene hypothesis: an evolutionary perspective. *Microbes and infection / Institut Pasteur*, 12(6), pp.421–7.

Smith, J.M. & Haigh, J., 1974. The hitch-hiking effect of a favourable gene. *Genetical research*, 23(1), pp.23–35.

Soldevila, G., Raman, C. & Lozano, Francisco, 2011. The immunomodulatory properties of the CD5 lymphocyte receptor in health and disease. *Current opinion in immunology*, 23(3), pp.310–8.

Tajima, F., 1989. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, 123(3), pp.585–595.

Tang, K., Thornton, K.R. & Stoneking, Mark, 2007. A new approach for using genome scans to detect recent positive selection in the human genome. K. H. Wolfe, ed. *PLoS biology*, 5(7), p.e171.

Tishkoff, S A et al., 2001. Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. *Science (New York, N.Y.)*, 293(5529), pp.455–62.

Van der Vaart, M., Spaink, H.P. & Meijer, A.H., 2012. Pathogen recognition and activation of the innate immune response in zebrafish. *Advances in hematology*, 2012, p.159807.

Vasseur, E. et al., 2012. The evolutionary landscape of cytosolic microbial sensors in humans. *American journal of human genetics*, 91(1), pp.27–37.

Veeramah, K.R. et al., 2012. An early divergence of KhoeSan ancestors from those of other modern humans is supported by an ABC-based analysis of autosomal resequencing data. *Molecular biology and evolution*, 29(2), pp.617–30.

Vera, J. et al., 2009. The CD5 ectodomain interacts with conserved fungal cell wall components and protects from zymosan-induced septic shock-like syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), pp.1506–11.

Vitti, J.J. et al., 2012. Human evolutionary genomics: ethical and interpretive issues. *Trends in genetics : TIG*, 28(3), pp.137–45.

Voight, B.F. et al., 2006. A map of recent positive selection in the human genome. *PLoS biology*, 4(3), p.e72.

Williams, G.C., 2001. Pleiotropy, Natural Selection, and the Evolution of Senescence. *Sci. Aging Knowl. Environ.*, 2001(1), p.cp13–.

Williamson, S.H. et al., 2007. Localizing recent adaptive evolution in the human genome. *PLoS genetics*, 3(6), p.e90.

Youinou, P. & Renaudineau, Y., 2011. CD5 expression in B cells from patients with systemic lupus erythematosus. *Critical reviews in immunology*, 31(1), pp.31–42.

Annexes

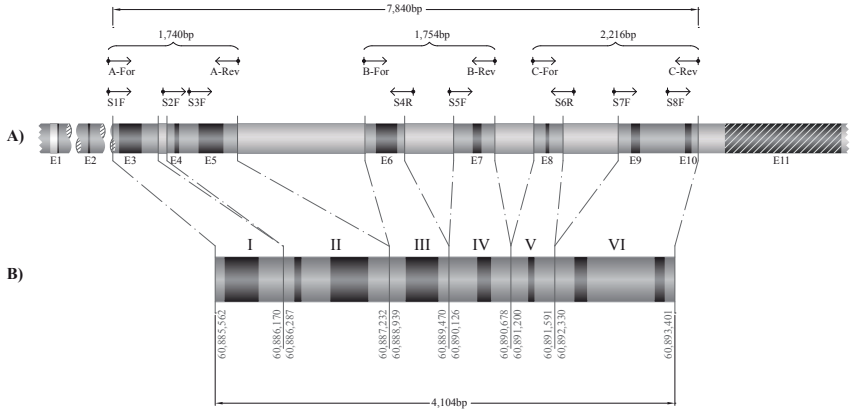
Annex 1. Supplementary Information chapter 1

Figure S1. *CD5* resequencing A) Resequencing design. Amplification and sequencing primers are shown by arrows and labeled as in supplementary table S2. Resequencing was designed towards exonic regions found in the fully-processed CD5 form. Exonic and unstranslated regions are represented by grey boxes and a grey-lined box, respectively. **B)** Construction of individual assemblies. Each final individual consensus sequence was built after the concatenation of six genomic segments (from I to VI). Genomic location for each segment is based on human assembly hg19/GRCh37 (March 2009).

Figure S2. The reference sequence

Reference sequence based on human assembly hg19/GRCh37 (March 2009) for each of the six genomic segments resequenced in all the individuals.

Suplementary Figure 1.



Suplementary table S1. Derived allele frequencies for rs22229177 in the HGDP-CEPH Human Genome Diversity Cell Panel.

Region	Population	2N	Frequency
Sub-Saharan Africa	Bantu	38	0.526
	Biaka Pygmies	52	0.385
	Mbuti Pygmies	10	0.500
	Mandenka	44	0.477
	San	12	0.000
	Yoruba	42	0.619
Middle East & North Africa	Mozabite	56	0.518
	Palestinian	90	0.467
	Bedouin	94	0.340
	Druze	82	0.500
Europe	French	56	0.518
	Basque	48	0.458
	Orcadian	28	0.393
	Sardinian	56	0.482
	Italian	42	0.500
	Adygei	34	0.500
	Russian	50	0.660
Central and South Asia	Balochi	48	0.729
	Brahui	50	0.840
	Burusho	50	0.740
	Hazara	46	0.870
	Kalash	44	0.682
	Makrani	50	0.660
	Pathan	50	0.660
	Sindhi	48	0.667
	NW China	58	0.845
East Asia	NE China	90	0.922
	S China	134	0.978
	Han	90	1.000
	Yakut	48	0.958
	Cambodian	20	0.950
	Japanese	60	0.983
Oceania	Melanesian	28	1.000
	Papuan	30	1.000
America	Pima	28	1.000
	Maya	38	0.763
	Colombian	14	0.929
	Karitiana	28	0.929
	Surui	14	1.000

Suplementary table S2. Amplification and sequencing primers for *CD5* sequencing analysis

Primer ID	Sequence (5´ - 3´)
Amplification	
A-For	GAAGGGACGAAGCTCACAAG
A-Rev	CAAGGCATTGAGTGTGGATG
B-For	AGGGAAGGGCAGAAAAGAAG
B-Rev	TTACTTGGGGCAGAAAATGG
C-For	GGAAACTGAGGCCTACGAGA
C-Rev	ACTGAGGGGAGGCATTGAGT
Sequencing	
S1F	GAAGGGACGAAGCTCACAAG
S2F	TCGCAGGAGGCTTAGAGAC
S3F	ATCACCTCCCAAGGCTAAG
S4R	TTGCCCTGTCTCCTATTATTG
S5F	TGGTATATGATGGCAAGGTG
S6R	ACTGTGTTGGGGAATACTGC
S7F	CAGTCAGATTGCTGGGTTAC
S8F	CAGGAGCGCTGTACTAAAGG

Annex 2. Supplementary Information Chapter 2

Supplemental Figure S1. Worldwide allele frequencies for the Leu372Val (rs1871534, top) and Thr357Ala (rs2272662, bottom) polymorphisms. Circles are not proportional to sample sizes. Complete list of population and sample sizes analyzed are given in Supplementary Table S1.

Supplemental Figure S2. **Neanderthal mt-DNA control for contamination.**

Supplemental Figure S3. Patterns of selection in a genomic region of 100 kb around the *ZIP4* gene (*SLC39A4)* (A). Gene context and summary of tests for positive selection obtained in the Yoruba population from the 1000 Genomes data. Those statistics which are based on the site frequency spectrum (Fay and Wu´s H, Fu and Li´s D and Tajima´s D) show weakly negative scores near *ZIP4* that do not approach genome-wide significance (not shown), so they should not be regarded as indicative of positive selection. Those statistics which are based on population differentiation (here: F_{ST}) show three SNPs (see Figure 1) with elevated values between CEU and YRI. One of them, rs1871534 (Leu372Val) is among the most highly differentiated SNPs in the genome. (B) Fine-scale recombination rate from the Yoruba population plotted in linear scale reveals a moderate recombination hotspot near *SLC39A4*. (C) Detailed view of simulated values along the 100 kb region for different statistical tests of positive selection assuming different scenarios comparable to Figure 1: (i) no selection and considering the observed recombination landscape from the Yoruba population (black lines); (ii) a selective sweep in the West African population and a constant recombination rate (red lines); and (iii) a selective sweep in the West African population and the observed recombination landscape including the hotspot (blue lines). Statistics were calculated in a sliding window approach with 25 kb windows and approximately 3 kb offset. For F_{ST} only the

maximum score for each window was considered. Straight lines indicate median values and dashed lines indicate the 5th and the 95th percentiles of 500 replicated simulations.

Supplemental Figure S4. Detection of ZIP4 isoforms by Western blot. (A) Gel was loaded with 80 µg of total protein extracts from HeLa cells transiently transfected with the different ZIP4 isoforms. Anti-HA antibody (1:1000) was used to detect the transporters and anti-beta actin (1:3000) as a loading control. (B) HeLa cells transfected with the Ala357-Leu372, Ala357-Val372, and Ala357-Pro372 isoforms were treated with 10µg/ml cyclohexamide for different time periods (1h, 3h, 6h and 8h). Total protein extracts were obtained and western blotting was performed. A representative experiment for each isoform is shown (left). The quantification analysis normalized the band intensity to the initial amount of protein before the treatment (time 0) (right). This experiment was performed three times per isoform (n=3).

Supplemental Figure S5. Retention of ZIP4 in the endoplasmatic reticulum. Immunostaining under permeabilizing conditions on cells expressing different ZIP4 variants using anti-HA (1:1000) for ZIP4 detection and anti-calnexin (1:1000) (Abcam) as an endogenous endoplasmic reticulum maker protein.

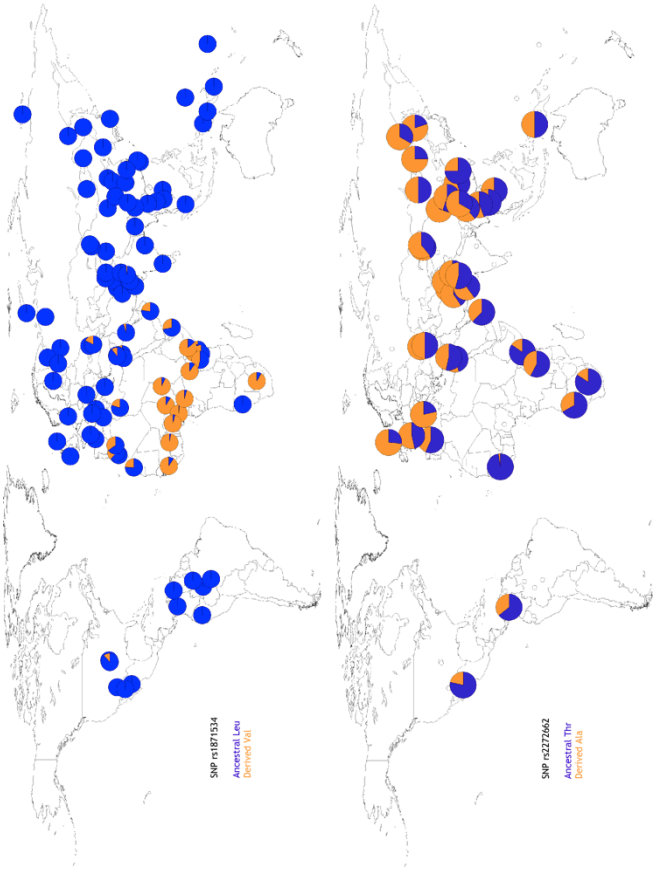
Supplemental Figure S6. Linkage disequilibrium plot for the YRI population in a 50kb window around the *ZIP4* (*SLC39A4)* gene. The plot was generated with Haploview and using HapMap 2 data (release 21).

Supplemental Figure S7. Haplotype visualization in a 40kb window around the *ZIP4* (*SLC39A4)* gene. Plots from the HapMap browser (http://hapmap.ncbi.nlm.nih.gov) are shown for the Yoruba, the Han Chinese and the French populations. There is no indication of extended haplotype patterns that could indicate a classical selective sweep in any of the three populations.

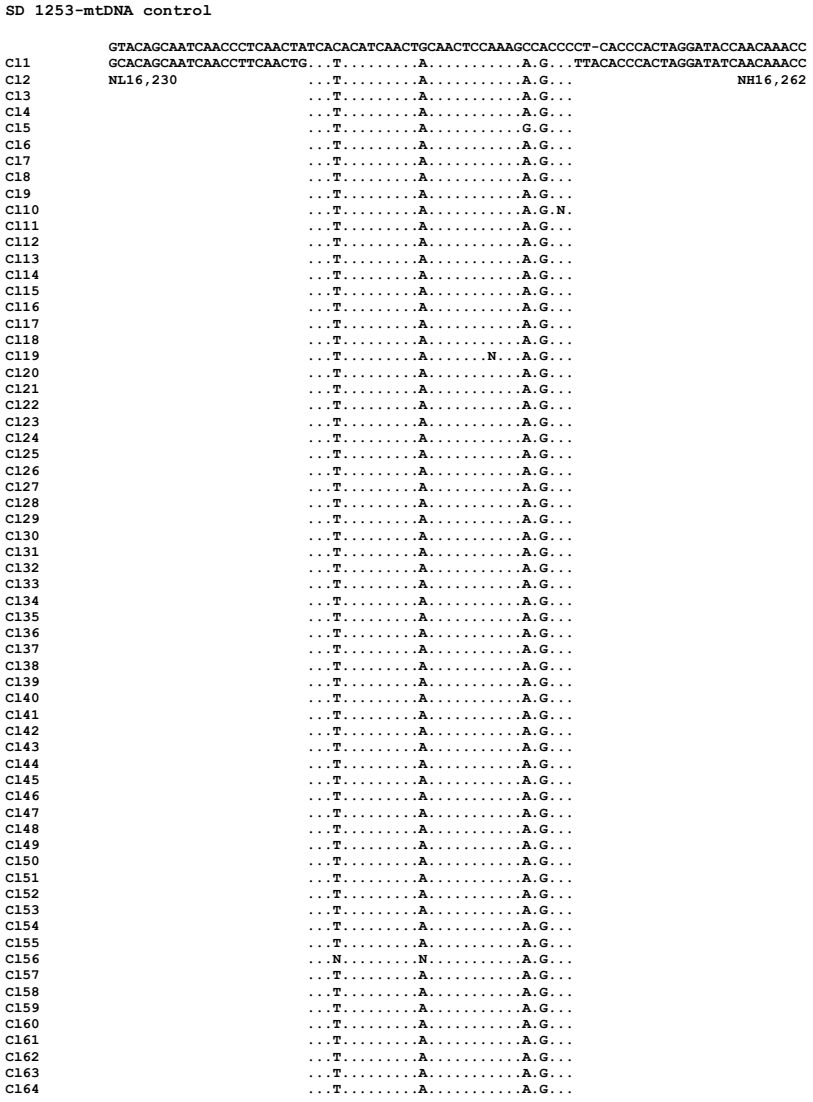
Supplemental Table S1. Worldwide allele frequencies for the Leu372Val (rs1871534) and Thr357Ala (rs2272662) polymorphisms.

Supplemental Table S2. Description of primers and hcDNA used in mutagenesis.

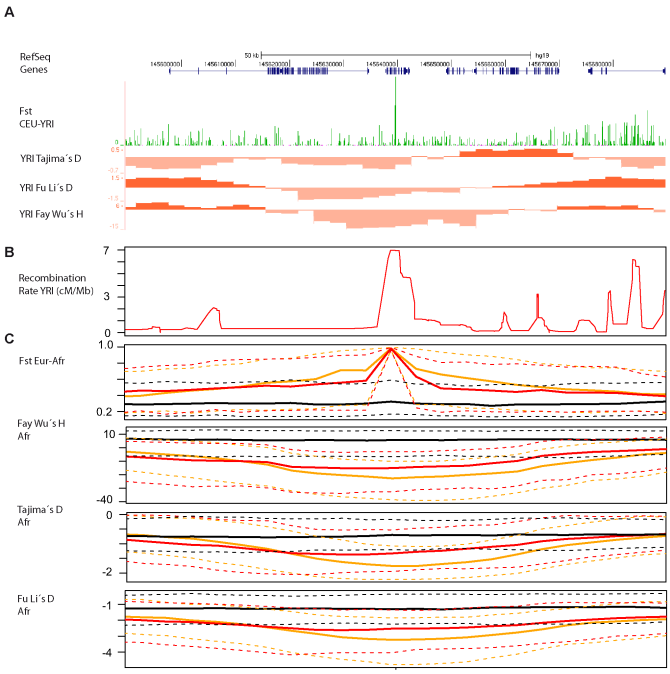
Supplemental Figure S1



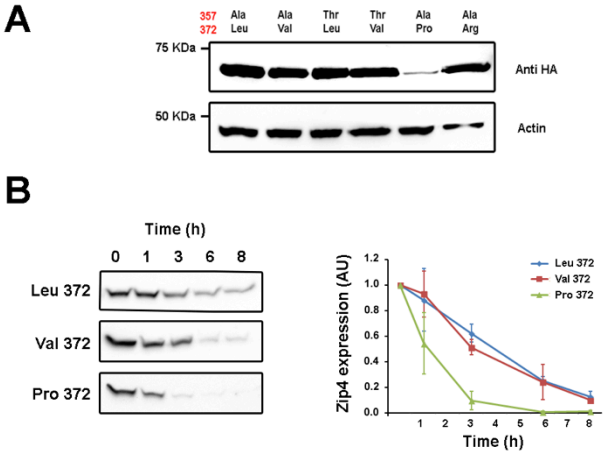
Supplemental Figure S2



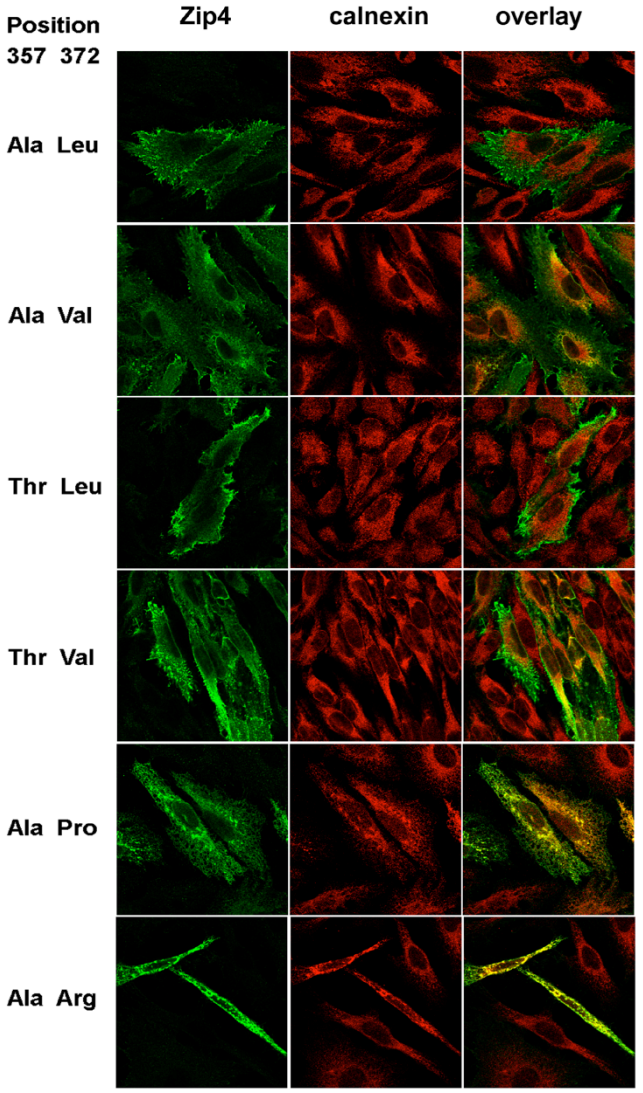
Supplemental Figure S3



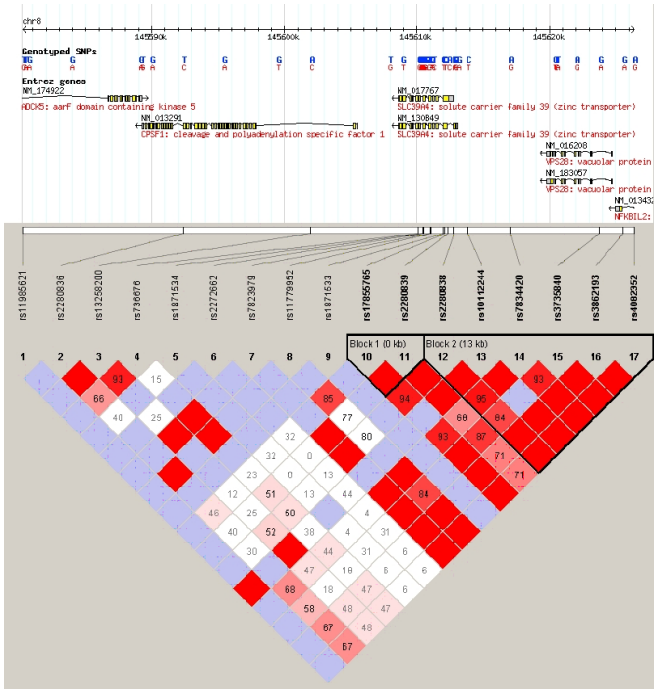
Supplemental Figure S4

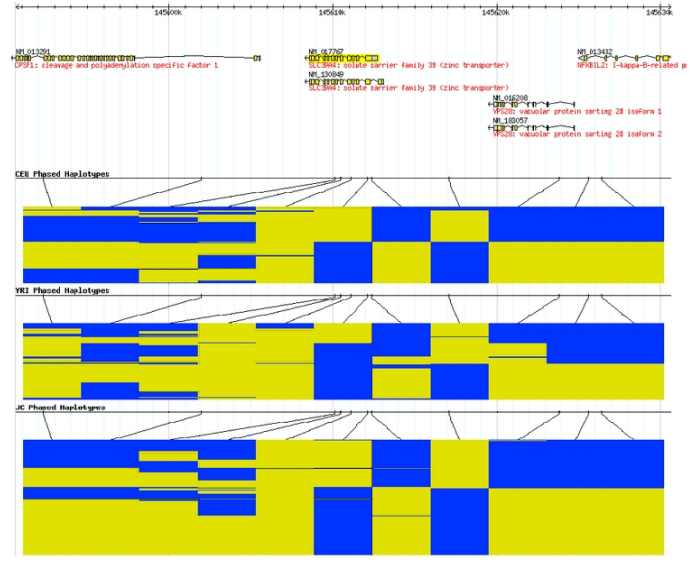


Supplemental Figure S5



Supplemental Figure S6 & S7





Supplemental Table S1

					rs1871534		rs2272662	
Order	Population	Origin	Geographic Coordinates	Source	2N	Val Allele	2N	Ala Allele
1	Morocco (Casablanca)	North Africa	33.53N, 7.58W	Present study	52	0.327		
2	Morocco (Rabat)	North Africa	34N, 6.85W	Present study	18	0.389		
3	Morocco (Nador)	North Africa	35.16N, 2.93W	Present study	20	0.350		
4	Libyans	North Africa	32.88N, 13.16E	Present study	92	0.196		
5	Saharawi	North Africa	25N, 13W	Present study	58	0.241		
6	African Americans	Africa	25-65N, 65-125W	Alfred	174	0.701		
7	Bantu	Africa	29S, 30E	HGDP	36	0.917	38	0.158
8	Chagga	Africa	2.5-3.5S, 37-38E	Alfred	88	0.750		
9	Hausa	Africa	7-18N, 4-20E	Alfred	76	0.908		
10	Ibo	Africa	5-7N, 5-10E	Alfred	94	0.989		
11	Lisongo	Africa	4-11.5N, 14-27E	Alfred	14	0.929		
12	Luhya	Africa	0.6N , 34.8E	HapMap	92	0.859	92	0.152
13	Maasai	Africa	0N, 37.9E	HapMap	90	0.456	92	0.424
14	Mandenka	Africa	12N, 12W	HGDP	42	0.905	44	0.023
15	Pygmy (Biaka)	Africa	4N,17E	Alfred	134	0.955		
16	Pygmy (Gabon)	Africa	2.13N, 12.05E	Present study	78	0.974	70	0.100
17	Pygmy (Mbuti)	Africa	1N, 29E	Alfred	74	0.892		
18	San	Africa	21S, 20E	HGDP	12	0.000	12	0.333
19	Sandawe	Africa	4-7S, 35-38E	Alfred	78	0.462		
20	Somali	Africa	12N-2S, 40-52E	Alfred	32	0.281		
21	Yoruba	Africa	6-10N, 2-8E	Alfred	148	0.959		
22	Zaramo	Africa	4-11S, 36-40E	Alfred	66	0.864		
23	Ami	Asia	22.5-24N, 121-121.5E	Alfred	78	0.000		
24	Atayal	Asia	21.75-25.5N, 120.5-122.5E	Alfred	82	0.000		
25	Balochi	Asia	30-31N, 66-67E	HGDP	48	0.042	48	0.458
26	Brahui	Asia	30-31N, 66-67E	HGDP	48	0.021	48	0.479
27	Burusho	Asia	36-37N, 73-75E	HGDP	50	0.000	48	0.521
29	Cambodian	Asia	10.5-14.5N, 102.5-107.5E	Alfred	44	0.000		
30	Dai	Asia	21N, 100E	HGDP	20	0.000	20	0.300
31	Daur	Asia	48-49N, 124E	HGDP	20	0.000	20	0.750
32	Druze	Asia	32.5-34N, 35-37E	Alfred	198	0.056		
33	Hakka	Asia	22-35N, 105-122E	Alfred	80	0.000		
34	Han	Asia	22-40N, 100-120E	Alfred	114	0.000		
35	Hazara	Asia	24-38N, 56-73E	Alfred	194	0.005		
36	Hezhen	Asia	47-48N, 132-135E	HGDP	20	0.000	20	0.800
37	Japanese	Asia	30-46N, 130-146E	Alfred	94	0.000		
38	Kachari	Asia	27-27.5N, 94-95.5E	Alfred	26	0.000		
39	Kalash	Asia	35-37N, 71-72E	HGDP	46	0.000	42	0.595
40	Keralite	Asia	8-13N, 75-77.5E	Alfred	54	0.000		
41	Khanty	Asia	59-67N, 65-88E	Alfred	98	0.000		
42	Komi-Zyrian	Asia	59-69N, 46-66E	Alfred	90	0.000		
43	Koreans	Asia	34.5-43N, 124.5-130.5E	Alfred	106	0.000		

44	Kuwaiti	Asia	28-30N, 46-49E	Alfred	22	0.045		
45	Lahu	Asia	22N, 100E	HGDP	16	0.000	16	0.563
46	Lao Loum	Asia	14-23N, 100-107.5E	Alfred	224	0.000		
47	Makrani	Asia	26N, 62-66E	HGDP	50	0.020	50	0.600
48	Malaysians	Asia	1-7N, 100-119E	Alfred	20	0.000		
49	Miaozu	Asia	28N, 109E	HGDP	18	0.000	20	0.550
50	Mohanna	Asia	23-27N, 66-68E	Alfred	96	0.000		
51	Mongola	Asia	48-49N, 118-120E	HGDP	20	0.000	20	0.500
52	Naxi	Asia	26N, 100E	HGDP	18	0.000	18	0.611
53	Negroid Makrani	Asia	23-27N, 61-68E	Alfred	48	0.167		
54	Oroqen	Asia	48-53N, 122-131E	HGDP	18	0.000	18	0.667
55	Pashtun	Asia	24-39N, 61-77E	Alfred	192	0.000		
56	Pathan	Asia	32-35N, 69-72E	HGDP	50	0.000	50	0.620
57	She	Asia	27N, 119E	HGDP	20	0.000	20	0.250
58	Sindhi	Asia	24-27N, 68-70E	HGDP	48	0.021	48	0.458
59	Thoti	Asia	13-20N, 77-84E	Alfred	24	0.000		
60	Tu	Asia	36N, 101E	HGDP	20	0.000	20	0.700
61	Tujia	Asia	29N, 109E	HGDP	20	0.000	20	0.450
62	Uygur	Asia	44N, 81E	HGDP	20	0.000	20	0.600
63	Xibo	Asia	43-44N, 81-82E	HGDP	18	0.000	18	0.611
64	Yakut	Asia	55-74N, 105-165E	Alfred	100	0.000		
65	Yizu	Asia	28N, 103E	HGDP	20	0.000	18	0.667
66	Adygei	Europe	45-44N, 39-40.5E	Alfred	106	0.000		
67	Basque	Europe	43N, 0	HGDP	48	0.000	46	0.435
68	Chuvash	Europe	54.5-56.5N, 46-48.5E	Alfred	82	0.000		
69	Danes	Europe	54.7-58N, 8-13E	Alfred	100	0.000		
70	Europeans (Mixed)	Europe	35-70N, 24W-56E	Alfred	176	0.000		
71	Finns	Europe	60-75N, 20-35E	Alfred	66	0.000		
72	French	Europe	46N, 2E	HGDP	56	0.000	54	0.537
73	Greeks	Europe	35-41.6N, 19.5-28.5E	Alfred	100	0.000		
74	Hungarian	Europe	45.5-48.5N, 16-23E	Alfred	170	0.000		
75	Irish	Europe	51-56N, 6-11W	Alfred	224	0.000		
76	Italian	Europe	37.9-47N, 7-18.5E	Alfred	178	0.006		
77	Orcadian	Europe	59N, 3W	HGDP	30	0.000	30	0.733
78	Russians	Europe	45-85N, 30-180E	Alfred	92	0.000		
79	Samaritans	Europe	31.75-32.25N, 34.5-35.5E	Alfred	76	0.000		
80	Sardinian	Europe	38.75-41.25N, 8-10E	Alfred	66	0.000		
81	Tuscan	Europe	40N, 9E	HGDP	16	0.000	14	0.786
82	Adygei	Middle East	44N, 39E	HGDP	34	0.000	34	0.794
83	Bedouin	Middle East	31N, 35E	HGDP	92	0.174	90	0.500
84	Druze	Middle East	32N, 35E	HGDP	80	0.025	84	0.548
85	Jews (Ashkenazi)	Middle East		Alfred	226	0.018		
86	Jews (Ethiopian)	Middle East	12-15N, 35-40E	Alfred	72	0.208		
87	Jews (Sephardic)	Middle East		Alfred	48	0.083		
88	Jews (Yemenite)	Middle East	12-18N, 43-53E,	Alfred	80	0.050		

Supplemental Table S2

Primer A (Val372)	ggcagtgggtgca**gtc**actggggacgctgtcctg
Primer B (Ala357)	ctggctgcaggggggtc**acc**cactacatcctgcagac
Primer C (Pro372)	gcctggcagtgggtgca**ccc**actggggacgctgtc
Primer D (Arg372)	cctggcagtgggtgca**cgc**actggggacgctgtc
hcDNA_cloned (Thr357-Leu372)	atggcgtccctggtctcgctggagctggggctgcttctggctgtgctggtggtgacg gcgacggcgtccccgcctgctggtctg ctgagcctgctcacctctggccagggcgctctggatcaagaggctctgggcggcct gttaaatacgctggcggaccgtgtgc actgcaccaacgggccgtgtggaaagtgcctgtctgtggaggacgccctgggcct gggcgagcctgaggggtcagggctgc ccccgggcccggtcctggaggccaggtacgtcgcccgcctcagtgccgccgccg tcctgtacctcagcaaccccgagggcac ctgtgaggacactcgggctggcctctgggcctctcatgcagaccacctcctggccct gctcgagagccccaaggccctgacc ccgggcctgagctggctgctgcagaggatgcaggcccgggctgccggccagacc cccaagacggcctgcgtagatatccct cagctgctggaggaggcggtgggggcgggggctccgggcagtgctggcggcgt cctggctgccctgctggaccatgtcagg agcgggtcttgcttccacgccttgccgagccctcagtacttcgtggactttgtgttcca gcagcacagcagcgaggtccctatg acgctggccgagctgtcagccttgatgcagcgcctgggggtgggcagggaggcc cacagtgaccacagtcatcggcacagg ggagccagcagccgggaccctgtgcccctcatcagctccagcaacagctccagtg

	tgtgggacacggtatgcctgagtgcc agggacgtgatggctgcatatggactgtcggaacaggctggggtgaccccggag gcctgggcccaactgagccctgccctg ctccaacagcagctgagtggagcctgcacctcccagtccaggcccccccgtccagg accagctcagccagtcagagaggtat ctgtacggctccctggccacgctgctcatctgcctctgcgcggtctttggcctcctgct gctgacctgcactggctgcaggggg gtc**gcc**cactacatcctgcagaccttcctgagcctggcagtgggtgca**ctc**actgg ggacgctgtcctgcatctgacgcccaa ggtgctggggctgcatacacacagcgaagagggcctcagcccacagcccacctg gcgcctcctggctatgctggccgggct ctacgccttcttcctgtttgagaacctcttcaatctcctgctgcccagggacccggag gacctggaggacgggccctgcggcc acagcagccatagccacgggggccacagccacggtgtgtccctgcagctggcac ccagcgagctccggcagcccaagccc ccccacgagggctcccgcgcagacctggtggcggaggagagcccggagctgct gaaccctgagcccaggagactgagccc agagttgaggctactgccctatatgatcactctgggcgacgccgtgcacaacttcgc cgacgggctggccgtgggcgccgcc ttcgcgtcctcctggaagaccgggctggccacctcgctggccgtgttctgccacga gttgccacacgagctgggggacttcgc cgccttgctgcacgcggggctgtccgtgcgccaagcactgctgctgaacctggcct ccgcgctcacggccttcgctggtctcta cgtggcactcgcggttggagtcagcgaggagagcgaggcctggatcctggcagt ggccaccggcctgttcctctacgtagca ctctgcgacatgctcccggcgatgttgaaagtacgggacccgcggccctggctcct cttcctgctgcacaacgtgggcctgct gggcggctggaccgtcctgctgctgctgtccctgtacgaggatgacatcaccttc

Annex 3. Suplementary material chapter 3.

Table S1. List of genes included within the analysed pathways (Excel file)
Table S2. Sample description and origin
Table S3. Capture design and sequencing
Table S4. Sequencing statistics for the 20 chimpanzee individuals
Table S5. Significance of descriptive statistics in CDS
Table S6. Significance of descriptive statistics in non-coding regions
Table S7. Distribution of fitness effects for all elements and pathways
Table S8. Estimated alpha (α) and omega (ω_α) values between pathways for each genomic element analyzed
Table S9. Estimated CDS alpha and omega (ω_α) values per dN/dS quartile in the Actin and Complement pathways
Table S10. Comparison of alpha (α) values between pathways for each genomic element analyzed
Table S11. Comparison of estimated alpha (α) and omega (ω_α) values in CDS per dN/dS quartile between the Actin and Complement pathways
Supplementary Note 1: Selection of Accelerated Introns

Figure S1. Fraction of substitutions due to positive selection: alpha (α) values**. A.** Alpha (α) values per genomic element and pathway. Significance values for the 95% confidence interval have been obtained by bootstrapping requiring a minimum threshold of size (bp). Values for the 2.5% and 97.5% threshold are indicated. **B**. CDS alpha (α) values comparison between the Actin and Complement pathways. The comparison is shown overall as well as between the Actin and the percentiles 0.25, 0.25-0.75 and 0.75 of the complement dN/dS gene distribution values as calculated in (Serra et al. 2011) .
Figure S2. Unfolded site frequency spectrum (SFS) for all elements and pathways

Table S2. Sample description and origin

Name	Sex	Geographical origin
Vaillant	M	Gabon (H.-O., région de Franceville)
Doris	F	Gabon (O.-M., Rabi près de Gamba)
Julie	F	Gabon (H.-O., found on the road)
Clara	F	Gabon
Aboume	M	Gabon
Amelie	F	Gabon (H.-O., région de Franceville)
Benefice	F	Gabon (CNRS)
Lalala	F	Gabon (Libreville)
Masuku	F	Gabon (H.-O., Franceville)
Chiquita	F	Gabon
Ayrton	M	Gabon, bought in Moanda
Noemie	F	Guinee Equato
Bakoumba	M	Gabon
Brigitte	F	Gabon, bought in Moanda
Fifi	F	Gabon, bought in Port-Gentil
Judy	F	Gabon (O.M., bought in POG)
Makata	M	Gabon (H.-O., village de Makatamangué)
Makokou	F	Gabon (O.-I.)
Moanda	M	Gabon (H O)
Morphee	F	Gabon (CNRS)
Mpassa	M	Gabon (CNRS, 1. Generation)

The 20 samples *Pan troglodytes troglodytes* samples are from wild-born unrelated individuals. F, female; M, Male

Table S3. Capture design and sequencing

Lane	Sample	Index	Pool Kit 1	Pool Kit 2
1	Doris	AACT	A	E
1	Clara	TTGT	B	E
1	Aboume	GGGT	A	F
1	Amelie	CCCT	C	F
1	Benefice	ACGT	B	G
1	Lalala	CAGT	C	G
1	Masuku	CGTT	C	H
1	Chiquita	ATAT	D	H
1	Ayrton	CTTT	D	I
1	Noemie	GAAT	A	I
2	Vaillant	CAGT	J	P
2	Julie	CGTT	K	P
2	Fifi	ATAT	L	P
2	Moanda	CTTT	L	Q
2	Mpassa	GAAT	M	Q
2	Morphee	TATT	M	R
2	Makokou	CCCT	J	R
2	Makata	TCAT	K	R
2	Bakamba	GCTT	0	S
2	Brigitte	TGCT	0	S

All capturing and sequencing procedures were performed at the Genomics Unit of the Center for Genomic Regulation (CRG) Core Facilities. Briefly, individual DNA libraries were tagged with a specific PE tagged genomic adapter during sample preparation. Different pools of 2-3 libraries were then hybridised with the 120 bp biotinylated RNA baits from two custom Agilent SureSelect kits. After enrichment with each individual kit, captured fragments were purified, pooled in two groups (each containing the two sets of captured regions from each of 10 different samples), and sequenced in two different lanes of an Illumina HiSeq 2000 System.

Table S4. Sequencing statistics for the 20 chimpanzee individuals

Name	Reads	Bases on target	Mean bait coverage	Mean target coverage	Bases at >=2x	Bases at >=10x	Bases at >=20x	Bases at >=30x
Doris	14,450,654	546,586,422	68.936	63.561	1.000	0.994	0.929	0.826
Clara	13,945,914	478,904,934	60.469	55.691	1.000	0.973	0.860	0.730
Aboume	12,897,146	394,389,279	49.908	45.863	1.000	0.962	0.826	0.692
Amelie	13,539,808	341,355,827	43.029	39.696	1.000	0.942	0.764	0.590
Benefice	17,769,030	549,839,922	69.593	63.940	1.000	0.989	0.907	0.804
Lalala	15,226,252	412,113,777	52.083	47.924	1.000	0.970	0.838	0.709
Masuku	12,750,774	347,524,476	43.683	40.413	1.000	0.943	0.763	0.583
Chiquita	10,681,638	494,159,663	62.256	57.465	1.000	0.992	0.912	0.796
Ayrton	18,284,096	655,512,538	82.929	76.228	1.000	0.996	0.948	0.865
Noemie	15,891,022	564,017,153	71.281	65.588	1.000	0.992	0.923	0.825
Vaillant	44,742,276	883,726,904	111.407	102.767	1.000	0.998	0.982	0.942
Julie	34,817,600	710,907,883	89.782	82.670	1.000	0.997	0.967	0.907
Fifi	22,507,120	763,702,005	95.901	88.809	1.000	0.997	0.960	0.891
Moanda	14,585,522	405,048,219	51.064	47.102	1.000	0.973	0.845	0.700
Mpassa	41,353,350	1,088,371,551	137.493	126.565	1.000	0.993	0.941	0.880
Morphee	22,331,766	763,324,833	96.025	88.766	1.000	1.000	0.989	0.945
Makata	43,135,866	1,010,291,648	127.662	117.485	1.000	0.997	0.970	0.923
Bakamba	56,608,504	921,968,810	115.590	107.214	1.000	1.000	0.988	0.947
Brigitte	32,788,494	531,310,687	66.845	61.785	1.000	0.995	0.942	0.849
Makoukou	29,140,070	522,558,984	66.230	60.767	1.000	0.987	0.919	0.824

Mean bait coverage calculated over the total bp of baits (i.e. 7,360,656 bp) included in the total callable fraction of the genome of 8,599,335 bp.

Table S5. Significance of descriptive statistics in CDS

Pathway	Actin	Complement	Acc. Introns	Amiloid	Presenilin	Parkinson
pN	***0.003 ± 0.0003***	***0.005 ± 0.0007***	***0.005 ± 0.0004***	***0.002 ± 0.0004***	***0.004 ± 0.0004***	***0.002 ± 0.0004***
Actin	0.529	0.000	0.000	0.925	0.076	0.974
Complement	1.000	0.530	0.406	1.000	1.000	1.000
Acc. Introns	1.000	0.581	0.518	1.000	1.000	1.000
Amiloid	0.029	0.000	0.000	0.428	0.002	0.704
Presenilin	0.943	0.005	0.000	0.996	0.455	1.000
Parkinson	0.005	0.000	0.000	0.238	0.000	0.510
pS	***0.007 ± 0.0007***	***0.008 ± 0.0010***	***0.009 ± 0.0006***	***0.007 ± 0.0009***	***0.007 ± 0.0008***	***0.005 ± 0.0008***
Actin	0.494	0.070	0.000	0.547	0.280	0.995
Complement	0.985	0.531	0.082	0.956	0.895	1.000
Acc. Introns	1.000	0.802	0.498	0.992	0.985	1.000
Amiloid	0.445	0.058	0.000	0.507	0.244	0.993
Presenilin	0.712	0.132	0.000	0.700	0.479	0.999
Parkinson	0.002	0.000	0.000	0.003	0.000	0.492
pN/pS	***0.442 ± 0.0675***	***0.599 ± 0.0936***	***0.553 ± 0.0578***	***0.371 ± 0.0885***	***0.496 ± 0.0547***	***0.492 ± 0.1359***
Actin	0.503	0.025	0.015	0.780	0.147	0.335
Complement	0.979	0.480	0.778	0.983	0.957	0.797
Acc. Introns	0.935	0.312	0.500	0.958	0.861	0.709
Amiloid	0.120	0.003	0.000	0.496	0.007	0.129
Presenilin	0.769	0.094	0.141	0.896	0.487	0.524
Parkinson	0.750	0.089	0.127	0.890	0.456	0.515
dN	***0.0010 ± 0.0001***	***0.0027 ± 0.0003***	***0.0019 ± 0.0002***	***0.0009 ± 0.0002***	***0.0013 ± 0.0001***	***0.0009 ± 0.0002***
Actin	0.422	0.000	0.000	0.672	0.019	0.680
Complement	1.000	0.459	1.000	1.000	1.000	1.000
Acc. Introns	1.000	0.000	0.428	1.000	1.000	1.000
Amiloid	0.150	0.000	0.000	0.415	0.001	0.432
Presenilin	0.981	0.000	0.000	0.979	0.559	0.989
Parkinson	0.150	0.000	0.000	0.415	0.001	0.432
dS	***0.0042 ± 0.0005***	***0.0042 ± 0.0007***	***0.0039 ± 0.0004***	***0.0032 ± 0.0005***	***0.0044 ± 0.0004***	***0.0036 ± 0.0006***
Actin	0.509	0.546	0.792	0.975	0.353	0.842
Complement	0.509	0.546	0.792	0.975	0.353	0.842
Acc. Introns	0.305	0.355	0.553	0.904	0.120	0.688
Amiloid	0.027	0.078	0.054	0.495	0.000	0.235
Presenilin	0.668	0.655	0.890	0.990	0.552	0.910
Parkinson	0.140	0.208	0.262	0.785	0.031	0.482
dN/dS	***0.2468 ± 0.0462***	***0.6558 ± 0.1443***	***0.4988 ± 0.0740***	***0.2889 ± 0.0801***	***0.2949 ± 0.0501***	***0.2584 ± 0.0727***
Actin	0.485	0.000	0.000	0.291	0.148	0.446
Complement	1.000	0.486	0.972	0.997	1.000	0.999
Acc. Introns	1.000	0.075	0.477	0.984	0.999	0.991
Amiloid	0.809	0.000	0.000	0.491	0.475	0.680
Presenilin	0.830	0.000	0.000	0.527	0.512	0.706
Parkinson	0.590	0.000	0.000	0.351	0.235	0.517

Table S6. Significance of descriptive statistics in non-coding regions

	Actin	Complement	Acc. Introns	Amiloid	Presenilin	Parkinson
Intron						
pPUS	***0.0089 ± 0.0004***	***0.0092 ± 0.0005***	***0.0108 ± 0.0002***	***0.0086 ± 0.0003***	***0.0091 ± 0.0005***	***0.0078 ± 0.0003***
Actin	0,513	0,243	0,000	0,824	0,402	1,000
Complement	0,805	0,507	0,000	0,976	0,605	1,000
Acc. Introns	1,000	1,000	0,421	1,000	0,999	1,000
Amiloid	0,206	0,105	0,000	0,512	0,189	0,997
Presenilin	0,717	0,429	0,000	0,952	0,532	1,000
Parkinson	0,000	0,001	0,000	0,007	0,002	0,506
pPUS/pS	***1.327 ± 0.1310***	***1.124 ± 0.1139***	***1.200 ± 0.0815***	***1.301± 0.1754***	***1.276 ± 0.1125***	***1.730 ± 0.3488***
Actin	0,510	0,943	0,925	0,562	0,641	0,059
Complement	0,044	0,483	0,159	0,099	0,058	0,001
Acc. Introns	0,151	0,751	0,494	0,246	0,215	0,009
Amiloid	0,424	0,922	0,884	0,485	0,568	0,047
Presenilin	0,356	0,900	0,811	0,435	0,467	0,037
Parkinson	0,995	1,000	1,000	0,974	0,999	0,487
dPUS	***0.0043 ± 0.0001***	***0.0046 ± 0.0002***	***0.0058 ± 0.0002***	***0.0045 ± 0.0002***	***0.0045 ± 0.0002***	***0.0041 ± 0.0001***
Actin	0,372	0,061	0,000	0,142	0,130	0,923
Complement	0,973	0,505	0,000	0,752	0,754	1,000
Acc. Introns	1,000	1,000	0,610	1,000	1,000	1,000
Amiloid	0,879	0,307	0,000	0,555	0,535	0,999
Presenilin	0,879	0,307	0,000	0,555	0,535	0,999
Parkinson	0,027	0,001	0,000	0,020	0,010	0,496
dPUS/dS	***1.041 ± 0.1319***	***1.105 ± 0.1922***	***1.485 ± 0.1492***	***1.390 ± 0.2376***	***1.026 ± 0.1084***	***1.137± 0.1934***
Actin	0,506	0,361	0,000	0,008	0,544	0,274
Complement	0,694	0,505	0,000	0,039	0,733	0,411
Acc. Introns	0,994	0,950	0,515	0,641	1,000	0,940
Amiloid	0,984	0,903	0,243	0,471	0,998	0,865
Presenilin	0,459	0,337	0,000	0,006	0,486	0,251
Parkinson	0,760	0,566	0,001	0,067	0,827	0,478
Promoter						
pPUS	***0.0078 ± 0.0004***	***0.0092 ± 0.0005***	***0.0106 ± 0.0003***	***0.0086 ± 0.0004***	***0.0081 ± 0.0004***	***0.0079 ± 0.0003***
Actin	0,520	0,003	0,000	0,030	0,264	0,399
Complement	0,999	0,540	0,000	0,892	0,996	1,000

Amiloid	0,853	0,921	0,920	0,498	0,867	0,047
Presenilin	0,416	0,541	0,305	0,100	0,507	0,003
Parkinson	1,000	1,000	1,000	0,968	1,000	0,509
dPUS	***0.0039 ± 0.0002***	***0.0043 ± 0.0002***	***0.0047 ± 0.0002***	***0.0042 ± 0.0003***	***0.0045 ± 0.0003***	***0.0042 ± 0.0002***
Actin	0,454	0,035	0,000	0,127	0,009	0,083
Complement	0,977	0,428	0,035	0,619	0,288	0,660
Acc. Introns	1,000	0,951	0,518	0,954	0,833	0,985
Amiloid	0,934	0,272	0,012	0,484	0,181	0,489
Presenilin	0,998	0,764	0,172	0,857	0,586	0,904
Parkinson	0,934	0,272	0,012	0,484	0,181	0,489
dPUS/dS	***0.941 ± 0.1347***	***1.039 ± 0.1911***	***1.210 ± 0.1342***	***1.302 ± 0.2618***	***1.020 ± 0.1187***	***1.169 ± 0.2301***
Actin	0,503	0,292	0,007	0,020	0,243	0,093
Complement	0,760	0,522	0,070	0,078	0,589	0,238
Acc. Introns	0,953	0,798	0,485	0,304	0,922	0,595
Amiloid	0,984	0,889	0,730	0,458	0,972	0,722
Presenilin	0,725	0,472	0,049	0,065	0,528	0,203
Parkinson	0,920	0,761	0,375	0,231	0,874	0,526
Trailer						
pPUS	***0.0078 ± 0.0004***	***0.0092 ± 0.0005***	***0.0106 ± 0.0003***	***0.0086 ± 0.0004***	***0.0081 ± 0.0004***	***0.0079 ± 0.0003***
Actin	0,548	0,182	0,000	0,440	0,272	0,963
Complement	0,884	0,468	0,000	0,785	0,731	0,998
Acc. Introns	1,000	0,995	0,478	1,000	1,000	1,000
Amiloid	0,648	0,243	0,000	0,523	0,386	0,979
Presenilin	0,741	0,321	0,000	0,622	0,514	0,991
Parkinson	0,065	0,022	0,000	0,055	0,006	0,559
pPUS/pS	***1.162 ± 0.1235***	***1.123 ± 0.1256***	***1.175 ± 0.0895***	***1.305 ± 0.1907***	***1.130± 0.1383***	***1.748 ± 0.3794***
Actin	0,500	0,902	0,929	0,399	0,584	0,019
Complement	0,087	0,482	0,292	0,104	0,146	0,000
Acc. Introns	0,167	0,641	0,520	0,156	0,242	0,001
Amiloid	0,626	0,939	0,970	0,485	0,680	0,039
Presenilin	0,374	0,841	0,858	0,314	0,467	0,017
Parkinson	0,997	1,000	1,000	0,983	0,996	0,515
dPUS	***0.0039 ± 0.0002***	***0.0043 ± 0.0002***	***0.0047 ± 0.0002***	***0.0042 ± 0.0003***	***0.0045 ± 0.0003***	***0.0042 ± 0.0002***
Actin	0,533	0,087	0,000	0,573	0,356	0,308
Compleme	0,976	0,545	0,006	0,925	0,935	0,911

Presenilin	0,414	0,234	0,001	0,045	0,497	0,105
Parkinson	0,901	0,700	0,174	0,335	0,955	0,486
UTR						
pPUS	***0.0067 ± 0.0005***	***0.0078 ± 0.0006***	***0.0101 ± 0.0009***	***0.0070 ± 0.0004***	***0.0070 ± 0.0004***	***0.0054 ± 0.0004***
Actin	0,481	0,039	0,000	0,236	0,178	0,999
Complement	0,992	0,446	0,000	0,965	0,976	1,000
Acc. Introns	1,000	1,000	0,555	1,000	1,000	1,000
Amiloid	0,720	0,091	0,000	0,505	0,489	1,000
Presenilin	0,720	0,091	0,000	0,505	0,489	1,000
Parkinson	0,002	0,000	0,000	0,000	0,000	0,540
pPUS/pS	***0.999 ± 0.1231***	***0.957 ± 0.1344***	***1.114 ± 0.1225***	***1.062 ± 0.1718***	***0.984± 0.11249***	***1.189 ± 0.2743***
Actin	0,487	0,613	0,173	0,309	0,511	0,179
Complement	0,354	0,494	0,085	0,221	0,373	0,124
Acc. Introns	0,819	0,864	0,520	0,588	0,819	0,355
Amiloid	0,696	0,762	0,330	0,472	0,693	0,278
Presenilin	0,440	0,570	0,139	0,280	0,458	0,157
Parkinson	0,928	0,932	0,725	0,741	0,920	0,480
dPUS	***0.0033 ± 0.0002***	***0.0041 ± 0.0004***	***0.0041 ± 0.0003***	***0.0036 ± 0.0005***	***0.0034 ± 0.0003***	***0.0030 ± 0.0004***
Actin	0,535	0,019	0,000	0,295	0,360	0,814
Complement	1,000	0,559	0,451	0,884	0,970	0,999
Acc. Introns	1,000	0,559	0,451	0,884	0,970	0,999
Amiloid	0,918	0,112	0,026	0,550	0,699	0,962
Presenilin	0,711	0,041	0,002	0,395	0,477	0,870
Parkinson	0,115	0,003	0,000	0,106	0,094	0,541
dPUS/dS	***0.786 ± 0.1149***	***0.973 ± 0.1849***	***1.064 ± 0.1520***	***1.107 ± 0.2467***	***0.783 ± 0.1033***	***0.823 ± 0.1982***
Actin	0,503	0,114	0,003	0,048	0,498	0,397
Complement	0,925	0,505	0,196	0,261	0,941	0,735
Acc. Introns	0,981	0,680	0,486	0,406	0,988	0,864
Amiloid	0,991	0,751	0,615	0,487	0,995	0,895
Presenilin	0,497	0,110	0,003	0,047	0,481	0,389
Parkinson	0,629	0,178	0,013	0,069	0,645	0,481

Under the diagonal, the percentile in which the different descriptive values of the pathways (in rows) fall in the bootstrapped distribution of descriptive values of the corresponding compared pathway (in columns). The reciprocal comparison is shown above the diagonal. Upper (dark grey) and lower (light grey) significance thresholds are set to the 0.975 and 0.025 percentiles of the bootstrapped distribution. Black cells contain the percentile of the observed descriptive value of a given pathway within its own bootstrapped distribution of descriptives values. Cells in italics and bold contain the observed values of each descriptive.

Table S7. Distribution of fitness effects for all elements and pathways

element	dataset	nearly neutral	midly deleterious	deleterious	very deleterious
0-fold	Actin	0.316	0.091	0.118	0.475
	Complement	0.421	0.286	0.272	0.021
	Acc. Introns	0.521	0.174	0.205	0.100
	Amiloid	0.211	0.125	0.196	0.468
	Presenilin	0.337	0.151	0.212	0.300
	Parkinson	0.132	0.166	0.342	0.361
UTR	Actin	0.724	0.270	0.006	0
	Complement	1	0	0	0
	Acc. Introns	1	0	0	0
	Amiloid	0.998	0.002	0	0
	Presenilin	0.893	0.086	0.021	0
	Parkinson	0.606	0.127	0.146	0.121
Intron	Actin	1	0	0	0
	Complement	1	0	0	0
	Acc. Introns	1	0	0	0
	Amiloid	1	0	0	0
	Presenilin	1	0	0	0
	Parkinson	1	0	0	0
Promoter	Actin	1	0	0	0
	Complement	1	0	0	0
	Acc. Introns	1	0	0	0
	Amiloid	1	0	0	0
	Presenilin	1	0	0	0
	Parkinson	1	0	0	0
Trailer	Actin	1	0	0	0
	Complement	1	0	0	0
	Acc. Introns	1	0	0	0
	Amiloid	1	0	0	0
	Presenilin	1	0	0	0
	Parkinson	1	0	0	0

Distribution of fitness effects of new mutations for all elements and pathways estimated as in (Keightley and Eyre-Walker 2007). Nearly neutral, $N_e s$<1; mildly deleterious, 1< $N_e s$ <10; deleterious 10< $N_e s$ <100; and very deleterious, $N_e s$ >100.

Table S8. Estimated alpha and omega values between pathways for each genomic element analyzed

Element	Alpha			Omega		
and pathway	α	2.50%	97.50%	ω_α	2.50%	97.50%
CDS						
Actin	-0.23	-6.35	0.56	-0.06	-0.36	0.22
Complement	0.71	0.42	0.89	0.89	0.35	1.50
Acc. Introns	0.13	-0.72	0.69	0.07	-0.23	0.53
Amiloid	0.38	-1.47	0.78	0.12	-0.16	0.38
Presenilin	0.16	-1.51	0.59	0.06	-0.24	0.32
Parkinson	0.72	-0.21	0.91	0.29	-0.03	0.59
Intron						
Actin	0.43	0.36	0.49	0.75	0.56	0.95
Complement	0.49	0.40	0.56	0.95	0.66	1.29
Acc. Introns	0.65	0.61	0.68	1.83	1.55	2.16
Introns w Acc. I	0.62	0.58	0.67	1.65	1.38	1.99
Only Acc. Introns	0.75	0.72	0.78	3.01	2.58	3.45
Amiloid	0.46	0.37	0.53	0.85	0.58	1.11
Presenilin	0.46	0.38	0.53	0.86	0.62	1.12
Parkinson	0.36	0.30	0.62	0.57	0.42	1.06
Promoter						
Actin	0.30	0.17	0.64	0.44	0.21	1.04
Complement	0.43	0.28	0.52	0.74	0.38	1.08
Acc. Introns	0.50	0.41	0.57	1.01	0.70	1.32
Amiloid	0.39	0.19	0.52	0.64	0.23	1.08
Presenilin	0.46	0.32	0.57	0.84	0.47	1.24
Parkinson	0.39	0.25	0.65	0.65	0.33	1.22
Trailer						
Actin	0.36	0.20	0.47	0.56	0.25	0.88
Complement	0.46	0.31	0.56	0.85	0.45	1.28
Acc. Introns	0.56	0.49	0.62	1.29	0.97	1.65
Amiloid	0.35	0.11	0.52	0.54	0.12	1.02
Presenilin	0.39	0.25	0.48	0.63	0.33	0.94
Parkinson	0.47	0.31	0.76	0.80	0.45	1.53
UTR						
Actin	0.49	-0.29	0.91	0.60	-0.22	1.40
Complement	0.35	0.00	0.67	0.53	0.00	1.34
Acc. Introns	0.37	0.15	0.50	0.58	0.18	0.98
Amiloid	0.69	-0.10	0.90	1.00	-0.08	1.84
Presenilin	0.43	-0.27	0.74	0.51	-0.19	1.22
Parkinson	0.54	0.22	0.79	0.67	0.18	1.25

Alpha (α), fraction of substitution driven to fixation due to positive selection in the chimpanzee branch; omega (ω_α) ratio of adaptive to neutral chimpanzee divergence. Significance values for the 95% confidence interval have been obtained by bootstrap requiring a minimum threshold of genome size (bp). Values for the 2.5% and 97.5% threshold are indicated.

Table S9. Estimated CDS alpha and omega values per dN/dS quartile in the Actine and Complement pathways

				Alpha			**Omega**		
Pathway	**Length**	**Subs**	**SNPs**	**α**	**2.50%**	**97.50%**	**ω_α**	**2.50%**	**97.50%**
Actin	105,779	109	314	-0.23	-6.35	0.56	-0.06	-0.36	0.22
<0.25	9,889	11	19	0.60	-4.64	5.76	0.29	-0.52	1.07
0.25-0.75	53,818	56	164	-0.01	-7.18	0.80	0.00	-0.39	0.32
>0.75	20,636	30	82	-0.42	-22.88	18.51	-0.14	-0.81	0.67
Complement	54,112	148	265	0.71	0.42	0.89	0.89	0.35	1.50
<0.25	5,272	7	22	0.71	-8.15	9.98	0.34	-0.74	1.17
0.25-0.75	24,101	63	134	0.67	0.22	0.94	0.72	0.17	1.30
>0.75	13,715	54	64	0.81	0.38	0.99	1.77	0.4	3.05

Table S10. Comparison of alpha (α) values between pathways for each genomic element analyzed

	Actin	Complement	Acc. Introns	Amiloid	Presenilin	Parkinson
CDS						
α	***-0.23***	***0.71***	***0.13***	***0.38***	***0.16***	***0.72***
Actin	0.477	0.000	0.142	0.117	0.200	0.025
Complement	0.991	0.440	0.980	0.925	0.999	0.492
Acc. Introns	0.715	0.004	0.457	0.264	0.476	0.053
Amiloid	0.891	0.017	0.795	0.488	0.772	0.108
Presenilin	0.744	0.004	0.485	0.286	0.508	0.057
Parkinson	0.991	0.481	0.981	0.933	0.999	0.512
Intron						
α	***0.43***	***0.49***	***0.65***	***0.46***	***0.46***	***0.36***
Actin	0.516	0.103	0.000	0.218	0.212	0.714
Complement	0.980	0.539	0.000	0.790	0.785	0.856
Acc. Introns	1.000	1.000	0.595	1.000	1.000	0.985
Amiloid	0.829	0.258	0.000	0.495	0.487	0.800
Presenilin	0.829	0.258	0.000	0.495	0.487	0.800
Parkinson	0.026	0.001	0.000	0.019	0.010	0.264
Promoter						
α	***0.30***	***0.43***	***0.50***	***0.39***	***0.46***	***0.39***
Actin	0.352	0.046	0.000	0.150	0.011	0.069
Complement	0.716	0.270	0.012	0.477	0.169	0.378
Acc. Introns	0.799	0.503	0.052	0.681	0.332	0.610
Amiloid	0.888	0.934	0.457	0.940	0.782	0.840
Presenilin	0.716	0.270	0.012	0.477	0.169	0.378
Parkinson	0.836	0.729	0.142	0.837	0.528	0.739
UTR						
α	***0.49***	***0.35***	***0.37***	***0.69***	***0.43***	***0.54***
Actin	0.520	0.812	0.968	0.254	0.757	0.292
Complement	0.335	0.431	0.399	0.136	0.531	0.082
Acc. Introns	0.361	0.491	0.487	0.152	0.564	0.098
Amiloid	0.700	0.983	1.000	0.585	0.943	0.880
Presenilin	0.445	0.664	0.793	0.204	0.660	0.181
Parkinson	0.562	0.896	0.997	0.324	0.825	0.402
Trailer						
α	***0.36***	***0.46***	***0.56***	***0.35***	***0.39***	***0.47***
Actin	0.525	0.086	0.000	0.559	0.349	0.081
Complement	0.453	0.063	0.000	0.522	0.285	0.066
Acc. Introns	0.962	0.519	0.003	0.902	0.917	0.391
Amiloid	0.999	0.974	0.453	0.993	0.999	0.676
Presenilin	0.973	0.568	0.010	0.919	0.950	0.434
Parkinson	0.692	0.168	0.000	0.694	0.522	0.151

Under the diagonal, the percentile in which estimated alpha values of the pathways (in rows) fall in the bootstrapped distribution of alpha values of the corresponding compared pathway (in columns). The reciprocal comparison is shown above the diagonal. Upper (dark grey) and lower (light grey) significance thresholds are set to the 0.975 and 0.025 percentiles of the bootstrapped distribution. Black cells contain the percentile of the estimated alpha value of a given pathway within its own bootstrapped distribution of alpha values. Cells in italics and bold contain the observed values of α.

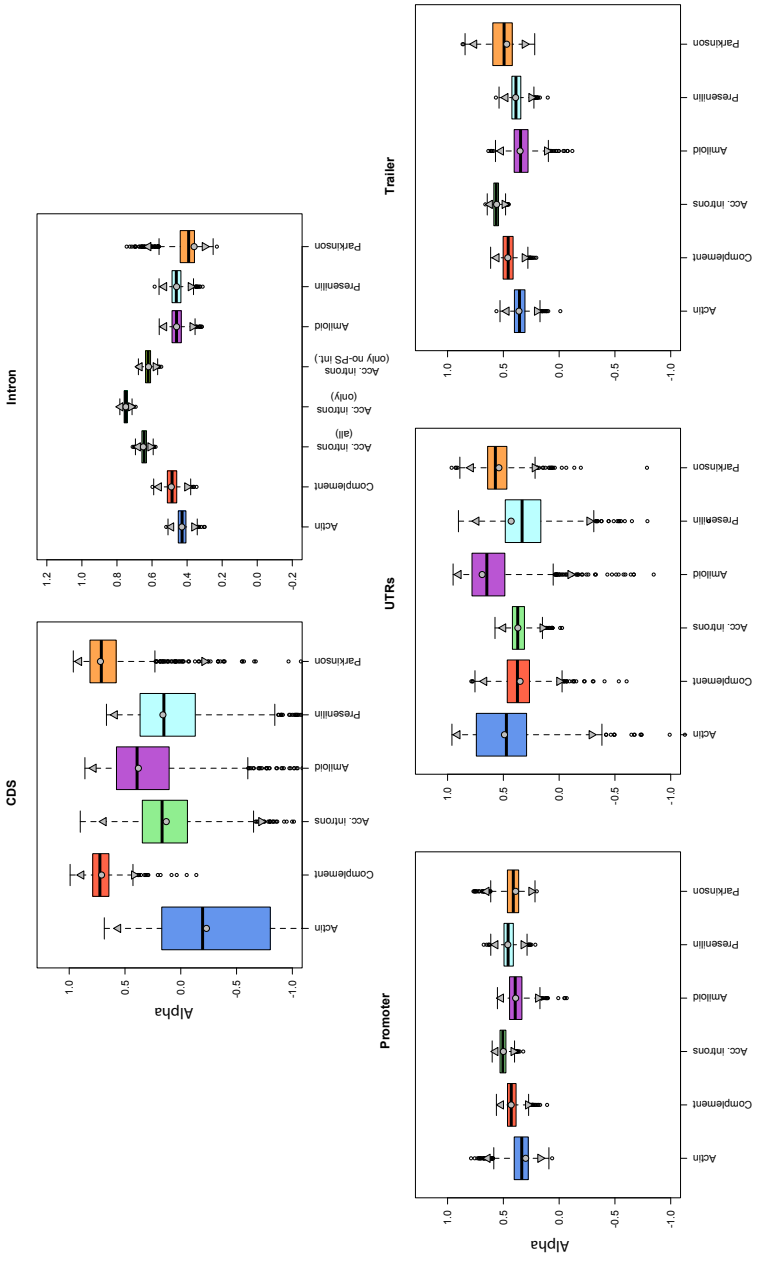
Table S11. Comparison of estimated alpha (α) and omega-alpha (ω_α) values in CDS per dN/dS quartile betweem the Actin and the Complement pathway

Figure S1

		Actin				Complement		
	Actin	0.25	0.25-0.75	0.75	**Complement**	0.25	0.25-0.75	0.75
α	***-0.23***	***0.60***	***-0.01***	***-0.42***	***0.71***	***0.71***	***0.67***	***0.81***
Actin	0.477	0.119	0.375	0.404	0	0.16	0	0.002
0.25	0.985	0.347	0.905	0.803	0.161	0.353	0.288	0.083
0.25-0.75	0.628	0.14	0.472	0.478	0.002	0.186	0.004	0.006
0.75	0.377	0.104	0.302	0.349	0	0.146	0	0.001
Complement	0.991	0.471	0.958	0.846	0.440	0.416	0.517	0.176
0.25	0.991	0.471	0.958	0.846	0.440	0.416	0.517	0.176
0.25-0.75	0.99	0.411	0.943	0.836	0.317	0.379	0.421	0.13
0.75	0.991	0.577	0.975	0.863	0.828	0.513	0.753	0.401
ω_α	***-0.06***	***0.29***	***0.00***	***-0.14***	***0.89***	***0.34***	***0.72***	***1.77***
Actin	0.482	0.199	0.381	0.555	0.001	0.239	0.001	0.006
0.25	0.996	0.431	0.956	0.842	0.015	0.479	0.067	0.016
0.25-0.75	0.648	0.226	0.497	0.612	0.002	0.266	0.005	0.006
0.75	0.26	0.16	0.241	0.474	0	0.195	0	0.002
Complement	1	0.89	1	0.996	0.451	0.894	0.669	0.083
0.25	1	0.475	0.981	0.873	0.022	0.522	0.083	0.020
0.25-0.75	1	0.758	1	0.985	0.236	0.807	0.437	0.049
0.75	1	1	1	1	1	1	1	0.437

Under the diagonal, the percentile in which estimated Actin and Complement α values per dN/dS quartile percentiles (in rows) fall in the bootstrapped distribution of alpha values of the corresponding compared percentile category (in columns). The reciprocal comparison is shown above the diagonal. Upper (dark grey) and lower (light grey) significance thresholds are set to the 0.975 and 0.025 percentiles of the bootstrapped distribution. Black cells contain the percentile of the estimated alpha value of a given category within its own bootstrapped distribution of alpha values. Cells in italics and bold contain the observed values of α and ω_α, respectively.

Figure S1
A


Promoter
UTRs
CDS
Intron
Trailer
Alpha
Actin
Complement
Acc. introns
Amiloid
Presenilin
Parkinson
Acc. introns (all)
Acc. introns (only)
Acc. introns (only no-PS Int.)

B.

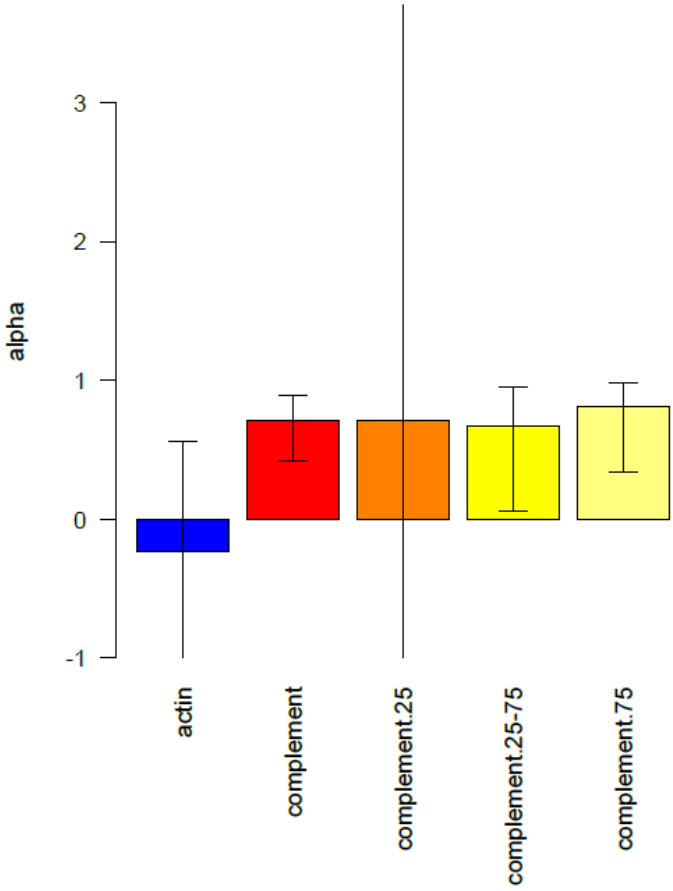
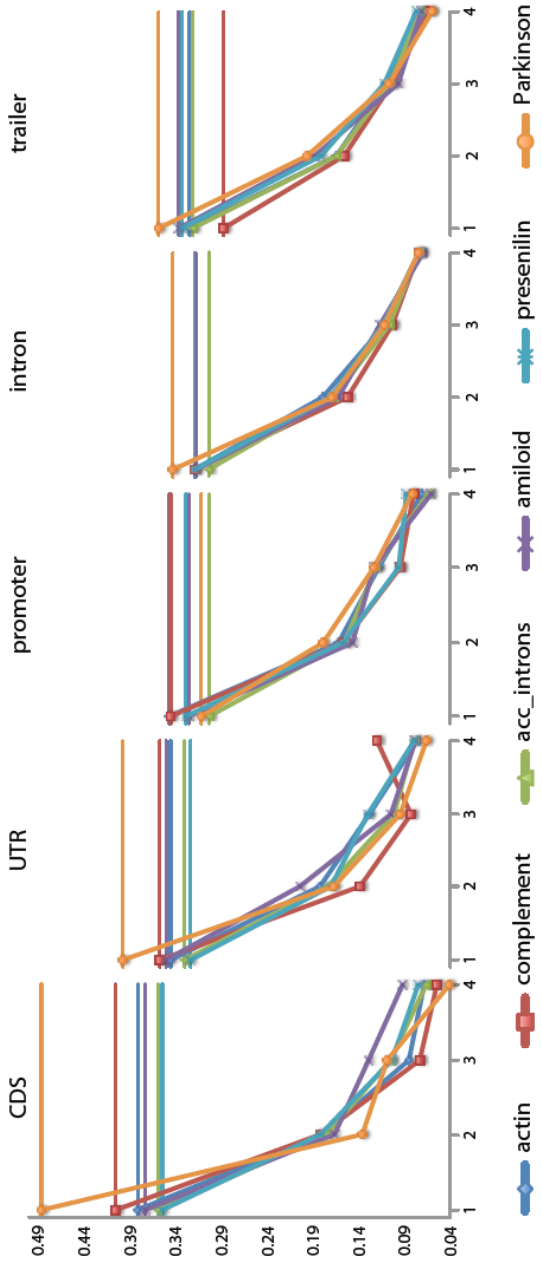

alpha
3
2
1
0
-1
actin
complement
complement.25
complement.25-75
complement.75

Figure S2


CDS
UTR
promoter
intron
trailer
0.49
0.44
0.39
0.34
0.29
0.24
0.19
0.14
0.09
0.04
1
2
3
4
actin
complement
acc_introns
amiloid
presenilin
Parkinson

Supplementary Note 1: Selection of Accelerated Introns

In order to test positive selection in *Pan troglodytes* introns we used a maximum likelihood test with the null and alternative models described by Haygood et al. (2007), fitted with HYPHY Pond et al. (2005). As neutral reference we used repeat sequences annotated in Human genome (hg18) and mapping in *Pan troglodytes* (pantro2) and *Rhesus macaca* (rhemac2) genomes (Ancestral Repeat sequences, ARs, Ponting and Hardison 2011) located in a window of 100kb surrounding each intron and not overlapping exons.

A list of 135,221 human introns coordinates (hg18) was obtained from the alignments of 14,286 genes with one-to-one defined orthology with *Rhesus macaca* and *Pan troglodytes* (Fernando, Olga PhD thesis). Based in this intervals list, sequence alignments of hg18/pantro2/rhemac2 were downloaded from UCSC web-server (http://genome.ucsc.edu/) using Galaxy tools (https://main.g2.bx.psu.edu/). In the same way hg18/pantro2/rhemac2 alignments of ARs neighbors to introns were downloaded. Alignments of ARs neighbors to each intron were concatenated obtaining a dataset of 134,599 alignments of introns (test dataset) with their respective neighbor ARs alignments.

Haygood et al. (2007) model of positive selection was tested using HYPHY software in a Linux platform for each intron, testing the alternative hypothesis of positive selection in the *Pan troglodytes* branch. In order to obtain the best likelihood for each intron, 100 replicates were performed for the null and alternative hypotheses. A log-ratio test was used to find significant differences between the best likelihood of the null and the alternative models. P- values were obtained by the chi-square test and corrected for a false discovery rate (FDR) at 0.05 using the q-value package in R (R developmental core team 2009, http://www.r-project.org/).

The alternative hypothesis of positive selection was significantly different to the null hypothesis of neutral evolution after FDR

correction for 2,033 introns belonging to 1,601 genes. Genomic sequences of these introns were downloaded individually for each species using a list of the coordinates of these introns annotated in each species (ENSEMBL v58, http://www.ensembl.org) and aligned using MUSCLE 3.6 software (Edgar, 2004). Gaps and unknown bases ('N') were eliminated of the alignments. The coverage of the alignments was calculated according with the length of intron sequence in *Pan troglodytes* and those with a coverage lesser than 80% were discarded. After this filtering we obtained a dataset of 728 introns belonging to 663 genes. The maximum likelihood test of positive selection was run again using these new alignments. Six hundred and sixty five introns remained being significant after this ran and the FDR correction.

47Finally, we chose 291 positively selected introns where the branch length estimation of the neutral reference sequence (ARs) is higher than the average estimated in their own chromosome. Thus, we eliminated possible false positive results due to conservation of the ARs sequences. Baits for sequencing could be designed for 180 of these introns, because most of them contain not unique sequences, which difficult the catch of the real sequence after sequencing process.

REFERENCES

Edgar, Robert C (2004). MUSCLE: multiple sequence alignment with high
accuracy and high throughput, *Nucleic Acids Research* 32(5), 1792-97.

Haygood R, Fedrigo O, Hanson B, Yokoyama KD, Wray GA (2007). Promoter
regions of many neural- and nutrition-related genes have experienced positive
selection during human evolution. Nat Genet. 39(9):1140-4.

Pond SL, Frost SD, Muse SV (2005). HyPhy: hypothesis testing using phylogenies. Bioinformatics. 21(5):676-9.

Ponting CP, Hardison RC (2011). What fraction of the human genome is functional? Genome Res. 21(11):1769-76.

Annex 4. More and more and more

Gracias …

A todos los **Carneros** y a los **Montoros**, que hacen que mi familia sea el más grande de mis orgullos. A mi **abu**, por todos sus "Prenda mia, todo lo que tú desees, que Dios te lo conceda¨.

A **NACHO.** Muso de mis primeras inquietudes científicas, que un dia fuiste mi padrino en la Genética y que iniciaste todo esto. También un semi-hermano. Sin duda, la persona que más lejos está y que más quisiera tener cerca. Por el café de los 30… (ya queda menos)

A **Salva** por dejarse perturbar de vez en cuando. Por todas las cosas compartidas desde el principio de esta etapa, por valorar las cosas importantes, por estar presente también en los momentos menos buenos. Por muchas ayudas. Por ser un buen amigo.

A **Rubens** y a **Sisifo**, por las preguntas incómodas. Por el aliento. Por los últimos días. Por toda la sabiduría, la filosofía. Por las ganas de vivir expandidas. Por transcender. Y a quien recibió como herencia mi escudo, **a Nunu**. Por enseñarme lo que es ser una gran mujer.

A **Iarita,** porque esto lo hemos conseguido juntas, y porque juntas vamos a conseguir muchísimas cosas más. Por nuestro poder. Por reescribir conmigo los conceptos importantes a nuestro antojo y beneficio. Por la determinación. No sabes la fuerza que me das hermanita.

A **Irenita**, por ocupar su lugar natural y recordarme lo importante.

A **Jenny**, por el aire fresco. Por los nuevos hogares. Por salvarme. Por tanto aprecio regalado. Por casita. Por ser atemporal. A **Toño,** por empezar juntos y quedarse después de la tormenta.

To **Gabe**, for your spirit that travels all the way from Brazil to here again and again to remain me what true love and friendship is. Because you are always present in my mind. For your smile. To **Danny**, for our no-matter-where reunions, because you make easily the impossible possible only by believing in. Also for the big smiles.

A mi despacho de eternos felices lisiados. Por constantemente reirnos de nosotros mismos, por el espíritu insaciable, por el cada dia. A **Mónix** y a **Paulita,** por los buenos momentos en el lab. Por la alegría! A **Itxa** y a **Urko**, que se fueron dejando el listón bien alto y a todos queriendo parecernos a ellos. A **David C.** por el Cangrejo, por ser cercano y por acercar clases . A **Arcadi** por ser siempre referente. También por posicionarse. A **Roger** y a **Nuria** por ser de mucha ayuda. A **Óscar** por las palabras cariñosas de cada dia. A l**a chungui** por ser única y no dejar de ayudar. A **Giovanni** por prestarse, ser decisivo y cuidar en el último momento de que esto salga bien. A los **naranjitos** y a los **beachbumbas**, por la lucha conjunta, por los chapuzones de tarde.

Al **E.P**, fuente de inspiracion constante y fantasia. Que exprime toda mi potencionalidad, y sabe como nadie alimentar los sueños. Porque somos parte de ese puzzle increible, inacabado e infinito de un mundo paralelo al que siempre podemos escaparnos. Por pintar los caminos!

To **Nihan**, because there are things I could only feel and live with you. I hope our flying-far bring us together sometime soon! I need somebody like you, you know , you know pequeña…it´s time to. A **Sarita** por enseñarnos cómo se vuela alto por encima de las circunstancias, y por las segundas y terceras impresiones.

Al **positive selectionator**, motor evolutivo de un Nuevo Pensamiento y forma de vivir, que sin duda está cambiando su entorno. Que al vuelo atrapa las buenas ideas, las propaga y crea nuevos conceptos. Por sacar lo mejor de todo lo que se encuentra por el camino y expandirlo, incluida yo misma. Por compartir pasiones, por ilusionarme tan fácilmente. Por la magia.

A aquel **chico del pasillo**. Que un rato caminamos juntos, y fantaseamos con esta tesis y con crear algo juntos.

Al **Sr.Rambla**, por el afecto y la amabilidad. Ha sido un placer trabajar con usted. A **Lisy**, a **Paco y** a **Rubén V.**. Por la colaboración que es más que la suma de las partes. Porque he aprendido mucho de vosotros. A **Maite** y a **Antonio** por recibirme con los abrazos abiertos, por tanta generosidad gratuita, por muchos ratos juntos. To **Jim**, for all the good thoughts.

A **Natalia** y a **Gabriel**, por pensar juntos. Por Menorca y Dublin.

Al **TUPPER CLUB**, por crear la mejor de mis (tan ansiadas) rutinas. Por ser familia. Por crear leyenda. Por nutrirnos cada dia. Qué especiales sois cada uno de vosotros tupperos!

A la **pequeña nubecilla**, hermana de leche. Por ser la otra parte de la historia de aquellas dos extranjeras que llegaron el mismo buen dia a Barcelona, cuya amistad fue un flechazo. A ellas dos que se bautizaron a sí mismas como Ricardas, cuyos pasos han ido hasta ahora y desde entonces por los mismos caminos. A ellas que brindan cada año por su tiempo juntas, que crearon juntas nuevas tradiciones en las mañanitas, que encontraron una playa con su nombre, que conducen la misma roja bicicleta por la ciudad, que ahora ya es su ciudad, que encontraron cada una a su Marc y a su familia catalane. Por la complicidad. Por la amistad.

Al anxoves **caracol** que por los suelos gusta revolotear, por las noches estiradas que se convirtieron en san fermines improvisados de medio dia. Por las conversaciones, los silencios y las lágrimas. Por los saltos. Por ser tan auténtico, por ser tan cercano, por abrir los brazos, por los atardeceres a las 11 de la noche en Cedeira.

A mi querido **Dieguito** por la vitalidad, por ser insaciable, por querer formar parte de todo, por tu fondo donde todos tienen cabida. A **Marc P.** por ser tan auténtico y por la controversia. Ya sabemos que siempre habrá un lugar ☺. A **Nino** por disfrutar tanto de todo y hacer que fácilmente se disfrute a tu lado cualquier simple momento. A **Arturo**, por ser un valiente, por perseguir sueños. A **Maria Niño,** por todas tus ganas inagotables de hacer cualquier cosa y por nuestro viaje a Galicia!. A **Alicia,** mi amiga perseguida que apareció un buen dia para echar las raíces y materializar una Amistad. Porque te espero en el próximo destino! Al **Rico Fede** por los deseos de transcender juntos más allá de nuestra generación. Por la camisa de cuadros y todo lo que le queda por vivir! A la mente brillante-impresionante-misteriosa de **mi compañero** y a **Mr. Nobody,** for doing much more than a little bit of this and that…

A **La Compi**, por poner la diferencia. Por salirse de la raya, por los lazos. Por pensarnos siempre. Por ser única entre todos los de esta especie. Por definirme. Por venir. Porque siento que has venido, vas y vienes y vendrás conmigo a todos los sitios. Porque te quiero muy cerca.

A los nuevos y ya viejos amigos que son capítulos esenciales de esta historia en Barcelona. A **Lu**, que para mi tu amistad es una conquista. Por cruzar el rio juntas. Por las perspectivas alternativas. Por transmitir tanta fuerza. Por la lucha. A **Bea**, porque no nos quitarán todo lo bailado y muchos más bailes llegarán. Por ser tan protagonista de nuestras vidas. Por la vitalidad. Por la eterna adolescencia. Porque sigamos cerquita

disfrutándonos tanto. A mi **buen amigo Steven**. Único Steven. Compañero de ya legendarias batallas desde los inicios de esta Era. Gracias por la estima compartida. Por tantos domingos. Por los viajes y las noches de a tres.
A **Pierre** por ser factor común. Por llegar más allá. Por el derroche de energía. Por las palmas, los tacones y los bailes. Por tantísimos ratos vividos tan divertidos. Por estar tan cerca y seguirnos. Por Granada. Por que te quiero mucho.

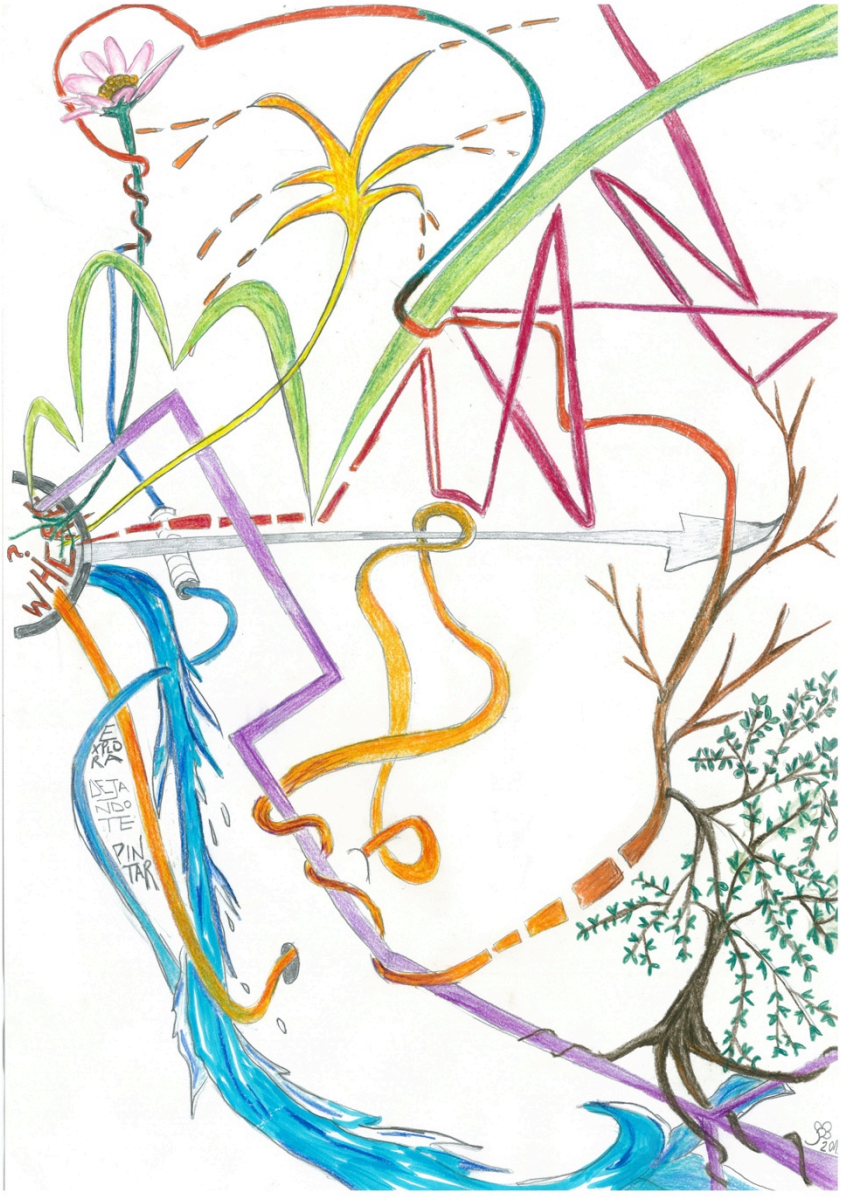
… os voy a echar mucho de menos…

Por ser constante vitales, a mi **Mariquilla**, **Carmen**, **Menchuquita**, a **Sari**, a **Cris**, a **Rubeno** y a **Jespecita**, porque no me puedo imaginar este mundo sin compartirlo con vosotras. Porque de vosotras fui hecha y a vosotras quiero volver.

A los que se han leido, y se van a leer este libro.

A **esta ciudad** con nombre propio, que ha puesto el lugar y el tiempo, y que me ha hecho tan fácil vivir aquí,
y gracias **al Pastis**, por aquel último domingo antes de primavera de 2009…

"A **la vida**, gracias a la vida, que me ha dado tanto"
…


WHERE
DEJA
NDO
TE
PIN
TAR

(Piero Pampanin, 2013)