



Evolution/Évolution

Comparative study on synteny between yeasts and vertebrates

Étude comparative de la synténie chez les levures et chez les vertébrés

Guénola Drillon, Gilles Fischer*

CNRS UMR7238, laboratoire de génomique des microorganismes, université Pierre-et-Marie-Curie, institut des Cordeliers, 15, rue de l'École-de-médecine, 75006 Paris, France

ARTICLE INFO

Article history:

Received 7 November 2010

Accepted after revision 29 March 2011

Available online 5 July 2011

Keywords:

Yeast
Vertebrate
Synteny
Genome
Evolution
Chromosome
Rearrangements

Mots clés :

Levures
Vertébrés
Synténie
Génome
Evolution
Chromosome
Réarrangements

ABSTRACT

We studied synteny conservation between 18 yeast species and 13 vertebrate species in order to provide a comparative analysis of the chromosomal plasticity in these 2 phyla. By computing the regions of conserved synteny between all pairwise combinations of species within each group, we show that in vertebrates, the number of conserved synteny blocks exponentially increases along with the divergence between orthologous protein and that concomitantly; the number of genes per block exponentially decreases. The same trends are found in yeasts but only when the mean protein divergence between orthologs remains below 36%. When the average protein divergence exceeds this threshold, the total number of recognizable synteny blocks gradually decreases due to the repeated accumulation of rearrangements. We also show that rearrangement rates are on average 3-fold higher in vertebrates than in yeasts, and are estimated to be of 2 rearrangements/Myr. However, the genome sizes being on average 200 times larger in vertebrates than in yeasts, the normalized rates of chromosome rearrangements (per Mb) are about 50-fold higher in yeast than in vertebrate genomes.

© 2011 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

RÉSUMÉ

Nous avons étudié la conservation de la synténie entre toutes les combinaisons deux à deux de 13 génomes de vertébrés et de 18 génomes de levures dans le but de fournir une analyse comparative de la plasticité chromosomique de ces 2 Phyla. En calculant les régions de synténie conservée entre toutes les paires d'espèces au sein de chaque groupe, nous montrons que chez les vertébrés, le nombre de blocs synténie augmente de façon exponentielle avec la divergence entre protéines orthologues et que de façon concomitante, le nombre de gènes par bloc décroît de façon exponentielle. Chez les levures, on observe les mêmes tendances mais lorsque la divergence protéique dépasse 36 %, le nombre de blocs diminue graduellement. Nous montrons également que les taux de réarrangements sont en moyenne 3 fois plus élevé chez les vertébrés que chez les levures et correspondent à une valeur de 2 réarrangements/Ma. Cependant, les génomes étant en moyenne 200 fois plus gros chez les vertébrés que chez les levures ; les taux normalisés de réarrangements chromosomiques (par Mb) sont environ 50 fois plus élevés dans les génomes de levures que dans les génomes de vertébrés.

© 2011 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

* Corresponding author.

E-mail address: gilles.fischer@upmc.fr (G. Fischer).

1. Synteny, an old genetic concept with a new meaning in comparative genomics

1.1. Synteny in the “linkage” sense

The first use of the word synteny dates back to the early seventies (Fig. 1) when new methods for gene mapping based on somatic hybrid cell lines were developed. Synteny originally described the colocalization of several markers on the same chromosome. As human chromosomes were preferentially lost in man-rodent hybrid cells, two genes could be attributed to the same chromosome when simultaneously present or absent from a hybrid cell population whatever the genetic distance separating them. These physically linked, but not necessarily genetically linked, genes were called syntenic genes. Etymologically, the term synteny means “on the same ribbon” (from the Greek *syn* = together and *taenia* = ribbon). Although relatively limited in number until the 1990s, nearly all published scientific papers referring to synteny involved gene mapping studies based on hybrid somatic cells in human and also in many primate, cattle and rodent species [1–4]. These methods led to the development of high-density radiation hybrid maps during the 1990s [5]. In the last 20 years, the number of synteny-related papers published each year has linearly increased to reach more than 200 scientific reports in the year 2009. It is interesting to note that in yeast, the number of publications dealing

with synteny has always been quantitatively negligible since this term was first invented (Fig. 1). However, several experimental studies based on electrophoretic karyotyping and later on, on comparative genomic hybridization, have allowed an exploration of the chromosome structures and their evolution in yeast [6–11].

1.2. Synteny in the conserved gene order sense

Chromosomes do not remain collinear over evolutionary time because rearrangements such as translocations, inversions, duplications and deletions shuffle the order and orientation of large genomic segments between genomes. When genetic maps became available for several related species, researchers started to compare genomes in order to understand how chromosomes are evolving. In this context, the notion of shared-synteny (or synteny conservation) was increasingly used in the literature. However, this notion was employed with a meaning different from the original definition of synteny. Instead of describing the linkage of genes on chromosomes in different species, the concept of shared-synteny rather described the preservation of gene order between homologs along chromosome segments in different species. Some geneticists rejected this use of the term synteny and noticed that a majority of the scientific papers did not use the term synteny according to its original meaning [12]. It is probably because a term of reference was lacking to

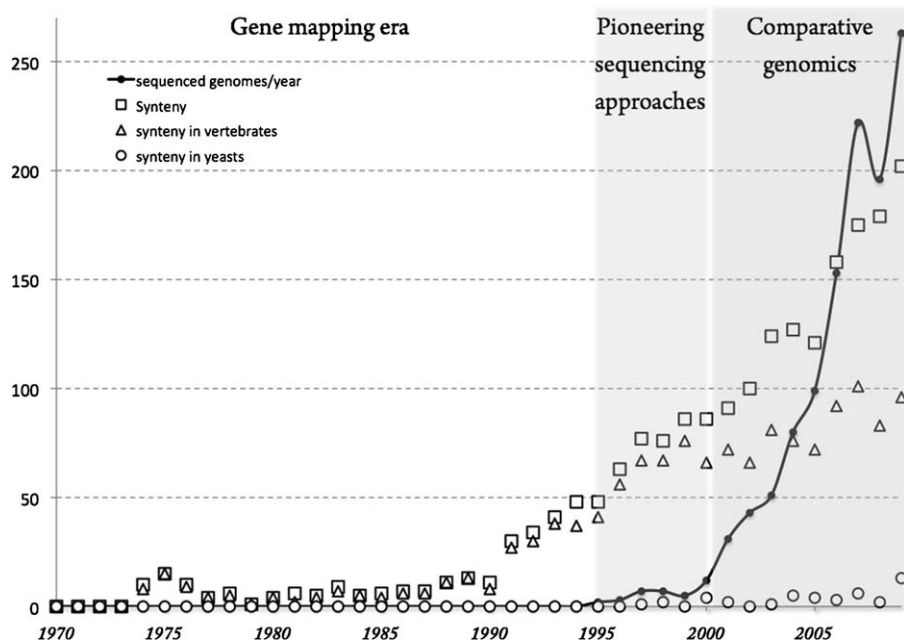


Fig. 1. The use of the term synteny in the scientific literature. The ‘Synteny’ plot (open squares) corresponds to the total number of publications citing the word synteny in either the title or the abstract sections identified in PubMed between 1970 and 2009. The ‘synteny in vertebrates’ plot (open triangles) corresponds to the fraction of these publications that in addition comprises one of the following terms: mammal* or mouse or human or primate or fish or cattle or rodent or dog or rat or mouse or vertebrate*, in either the title or the abstract sections. The ‘synteny in yeasts’ plot (open circles) corresponds to the fraction of the total synteny publications that comprises one of the following terms: yeast* or *Saccharomyces* or *Candida* or *Kluyveromyces*, in either the title or the abstract sections, followed by manual curation to remove publications citing yeast for methodological reasons (such as YAC). The black curve represents the number of completely sequenced genomes (eukaryotes, bacteria and archaea) published and referenced in the Genome OnLine Database (<http://www.genomesonline.org/>). From the year 2000 where the number of completely sequenced genomes rapidly increased, the relative prominence of vertebrates (open triangles) in the synteny-related literature has partly declined probably to the profit of plant and bacteria studies while the total number of publications dealing with yeast in the field of synteny (open circles) has always remained anecdotic.

describe the conserved order of common markers in different species that the term “shared synteny” has been diverted from its original meaning. Subsequently, this term was gradually stripped of the word “shared” (or conserved) and in today’s researcher’s vocabulary, synteny, on its own, (abusively) means conserved gene order between different species rather than linkage of two or more markers on a chromosome per se.

In the last decade, sequencing technologies have taken over traditional methods of gene mapping. With the growing availability of genome sequences, the large prominence of vertebrates in the synteny-related literature has partly declined (Fig. 1) probably to the profit of plant and bacteria studies (Fig. 1). Concomitantly, synteny studies have moved from the experimental field to the bioinformatics field. Although the total number of publications dealing with yeast in the field of synteny has remained anecdotic (Fig. 1), pioneering genome-wide explorations of gene content and gene order based on sequencing data only were first developed between related yeast species [13–16]. These studies paved the road for the birth of a new field called comparative genomics aiming at understanding the mechanisms of genome evolution through the comparative analysis of chromosomes between related species. Comparative genomics was concomitantly developed in vertebrates, with the sequencing of a compact fish genome, *Tetraodon nigroviridis* [17], to help for the annotation of the human genome [18,19], as well as in yeast with the Génolevures program [20] which

represented the first large exploratory sequencing project between related species aiming at deciphering the mechanisms of genome evolution. Among other things, the Génolevures 1 program sought for the mechanisms of chromosome map reorganization through the study of synteny conservation [21]. Since then, the study of synteny has been the tool of choice, both in yeasts and vertebrates, to unravel major conceptual advances in our understanding of genome evolution such as orthology/paralogy relationships and the relative contributions of segmental vs whole genome duplication (WGD) events. Synteny has also allowed the determination of the relative rates of chromosome rearrangements in individual lineages of yeast and vertebrate as well as the reconstruction of ancestral genomes. Finally, the study of the structure and the repartition of synteny breakpoints gives access the mechanisms of chromosome rearrangements and to the models of genome evolution. However, no study has so far put into perspective the relative levels and rates of chromosomal reorganization between yeast and vertebrates.

2. The evolution of synteny in yeasts and vertebrates

2.1. Major structural and functional differences between yeast and vertebrate genomes

Yeasts and vertebrates harbor very different genome characteristics in terms of size (a 200-fold difference on average, Table 1), number of genes, proportion and size of

Table 1
List of the 18 yeast and 13 vertebrate species with completed genome sequences.

Class	Species	Genome size (Mb)	Chromosome number	Scaffold number	Reference
Saccharomycetes	<i>Candida albicans</i>	14.3	8	8	[44]
Saccharomycetes	<i>Candida dubliniensis</i>	14.6	8 ^a	8 ^a	[45]
Saccharomycetes	<i>Candida glabrata</i>	12.3	13	13	[35]
Saccharomycetes	<i>Candida tropicalis</i>	14.6	8	23	[46]
Saccharomycetes	<i>Clavispora lusitaniae</i>	12.1	8	9	[46]
Saccharomycetes	<i>Debaryomyces hansenii</i>	12.2	7	7	[35]
Saccharomycetes	<i>Eremothecium gossypii</i>	8.7	7	7	[47]
Saccharomycetes	<i>Kluyveromyces lactis</i>	10.7	6	6	[35]
Saccharomycetes	<i>Lachancea kluyveri</i>	11.3	8	8	[48]
Saccharomycetes	<i>Lachancea thermotolerans</i>	10.4	8	8	[48]
Saccharomycetes	<i>Lachancea waltii</i>	10.7	8	10	[49]
Saccharomycetes	<i>Lodderomyces elongisporus</i>	15.5	9	27	[46]
Saccharomycetes	<i>Pichia guilliermondii</i>	10.6	8	9	[46]
Saccharomycetes	<i>Pichia pastoris</i>	9.4	4	6	[50,51]
Saccharomycetes	<i>Pichia stipitis</i>	15.4	8	9	[52]
Saccharomycetes	<i>Saccharomyces cerevisiae</i>	12.1	16	16	[53]
Saccharomycetes	<i>Yarrowia lipolytica</i>	20.5	6	6	[35]
Saccharomycetes	<i>Zygosaccharomyces rouxii</i>	9.8	7	7	[48]
Mammalia	<i>Canis familiaris</i>	2400	39	39	[54]
Actinopterygii	<i>Danio rerio</i>	1700	25	25	Unpublished
Mammalia	<i>Equus caballus</i>	2689	32	32	[55]
Aves	<i>Gallus gallus</i>	1000	40 ^b	30	[56]
Mammalia	<i>Homo sapiens</i>	3080	23	23	[18,19]
Mammalia	<i>Macaca mulatta</i>	2871	22	21	[57]
Mammalia	<i>Mus musculus</i>	2644	20	20	[58]
Marsupialia	<i>Opos monodelphis</i>	3475	9	9	[59]
Actinopterygii	<i>Oryzias latipes</i>	800	24	24	[60]
Mammalia	<i>Pan troglodytes</i>	3100	24	22	[61]
Mammalia	<i>Ratus Norvegicus</i>	3000	21	21	[62]
Aves	<i>Taeniopygia guttata</i>	2644	28	29	[63]
Actinopterygii	<i>Tetraodon nigroviridis</i>	350	21	21	[36]

^a Pseudochromosomes obtained by mapping onto *C. albicans* chromosomes [45].

^b Including microchromosomes that were not assembled.

introns, number of transposable elements and repeat sequences, gene density and proportion of coding and noncoding DNA (see [22] and [23] for a review of yeast and vertebrate genome architectures, respectively). In addition, major functional properties that can have a profound impact onto genome dynamics also differ between yeasts and vertebrates. Firstly, outcrossing between germ lines is the only mode of propagation of vertebrates, implying that the chromosome rearrangements that can be transmitted to the next generation and eventually reach fixation in populations are restricted to the meiotic divisions and the subsequent mitotic amplification of the gamete cell lines. The life cycle of wild yeasts is more complex, including clonal reproduction, outcrossing, and inbreeding. Yeast reproduction is principally characterized by a rapid clonal expansion when the environmental conditions are favorable. The proportion of sexual reproduction varies between lineages. Many lineages seem to be completely asexual while for those that undergo meiosis, mating mainly occur between ascospores originating from the same tetrad (inbreeding), hence limiting the level of outcrossing. It was calculated that *Saccharomyces* species undergo one sexual cycle every 1000 asexual divisions and that the proportion of outcrossing would be limited to once in every 50,000 to 100,000 asexual generations [24,25]. The rates of meiotic recombination are also very different because 1 centimorgan corresponds to approximately 3 kb in yeast but to about 1 Mb in human [26]. This implies that the two organisms have similar genome sizes in centimorgans. Secondly, it is well known that mitotic mutation rates vary between organisms [27,28]. From recent sequencing data, the intergeneration substitution rate is estimated to 1.1×10^{-8} per base per human haploid genome [29] and about 3×10^{-10} per base per division in either diploid or haploid cells of *Saccharomyces cerevisiae* [30,31]. These figures correspond to a 36-fold difference in the per-base probability of mutation. This difference is probably due to the cell divisions that occur in the germ line between two generations in human, while in yeast, one cell division corresponds to one asexual generation. In human, the number of cell divisions in the germ line per generation is limited to 30 cell divisions in women because oogenesis ceases replication during fetal life but is close to 200 divisions in a 20 year old man where spermatogenesis takes place throughout life [32]. Finally, another major functional difference between yeasts and vertebrates is the generation time that could differ by several orders of magnitude (few hours in yeasts compared to few months or years in vertebrates). This implies that for a similar evolutionary time the number of generations would be much higher in yeasts than in vertebrates although the average generation time for yeast populations in natural environments must be much longer than a few hours because they would often have to face critical growth conditions (such as long periods of starvation, low temperatures, etc.).

2.2. Chromosome evolution in yeasts and vertebrates

Because of these radically different structural and functional properties and also because important efforts to

understanding genome evolution have been made so far separately in yeasts and vertebrates, it was interesting to compare the dynamics of chromosome map reshuffling between these two groups of eukaryotes. Large sequencing data sets are presently available for 51 vertebrates (<http://www.ensembl.org/index.html>) and 32 yeasts from the *Saccharomycotina* subphylum [33]. However, there is a great diversity in the completeness of genome sequences. Because fragmented genome assemblies would introduce a high number of artificial synteny breakpoints, we excluded species where the genome sequence is broken into too many small contigs and focused on the 13 vertebrate genomes and the 18 yeast genomes for which chromosomes are represented by a single or a limited number of sequencing scaffolds (Table 1).

To look for common or different evolutionary themes and to test whether there exists some sort of molecular clock for chromosome rearrangements, we computed the blocks of conserved synteny between all pairs of species applying exactly the same criteria (see legend of Fig. 2) to the 78 and 153 possible pairwise comparisons of species within the groups of vertebrates and yeasts, respectively. A unit to measure evolutionary time that would be common to both yeast and vertebrate is nevertheless needed in order to compare the evolution of the number and the size of synteny blocks in these two groups of species. Estimations of evolutionary time in Myr for yeast are weak due to the absence of reliable fossil records. In addition, generation times are very different between yeasts and vertebrates. Therefore, we decided to use the average protein divergence between orthologs as the common unit of evolutionary range. Previous analyses using the global level of divergence of orthologous proteins revealed that the evolutionary range covered by the *Saccharomycotina* yeasts exceeds that of vertebrates and is similar to the span covered by the entire phylum of *Chordata* [34–36].

In vertebrates, the number of synteny blocks increases exponentially with increasing divergence time, varying from a very small number of blocks, 43 between human and chimpanzee, to more than 1900 blocks between dog and zebrafish (Fig. 2a). The highest numbers of blocks are found for comparisons involving a fish genome (circled in black on Fig. 2). Such large numbers are in good accordance with the large phylogenetic distance that separates fish from tetrapodes. However, Actinopterygii species have undergone a lineage specific WGD event that was subsequently followed by a massive loss of gene duplicates. Some synteny blocks could result from these local deletion events rather than from large chromosomal rearrangements per se (see below). It is also possible that these large numbers could partly result from an increase of rearrangement rates after the WGD event. In yeasts, the number of synteny blocks is more restrained, varying from 26 between *Candida albicans* and *C. dubliniensis* up to 744 between *Debaryomyces hansenii* and *Pichia pastoris*. The number of blocks also exponentially increases along with protein divergence but only between 8 and 36% of divergence. At increasing phylogenetic distances, the number of synteny blocks decreases (Fig. 2a). This trend is most likely due to the repeated accumulation of

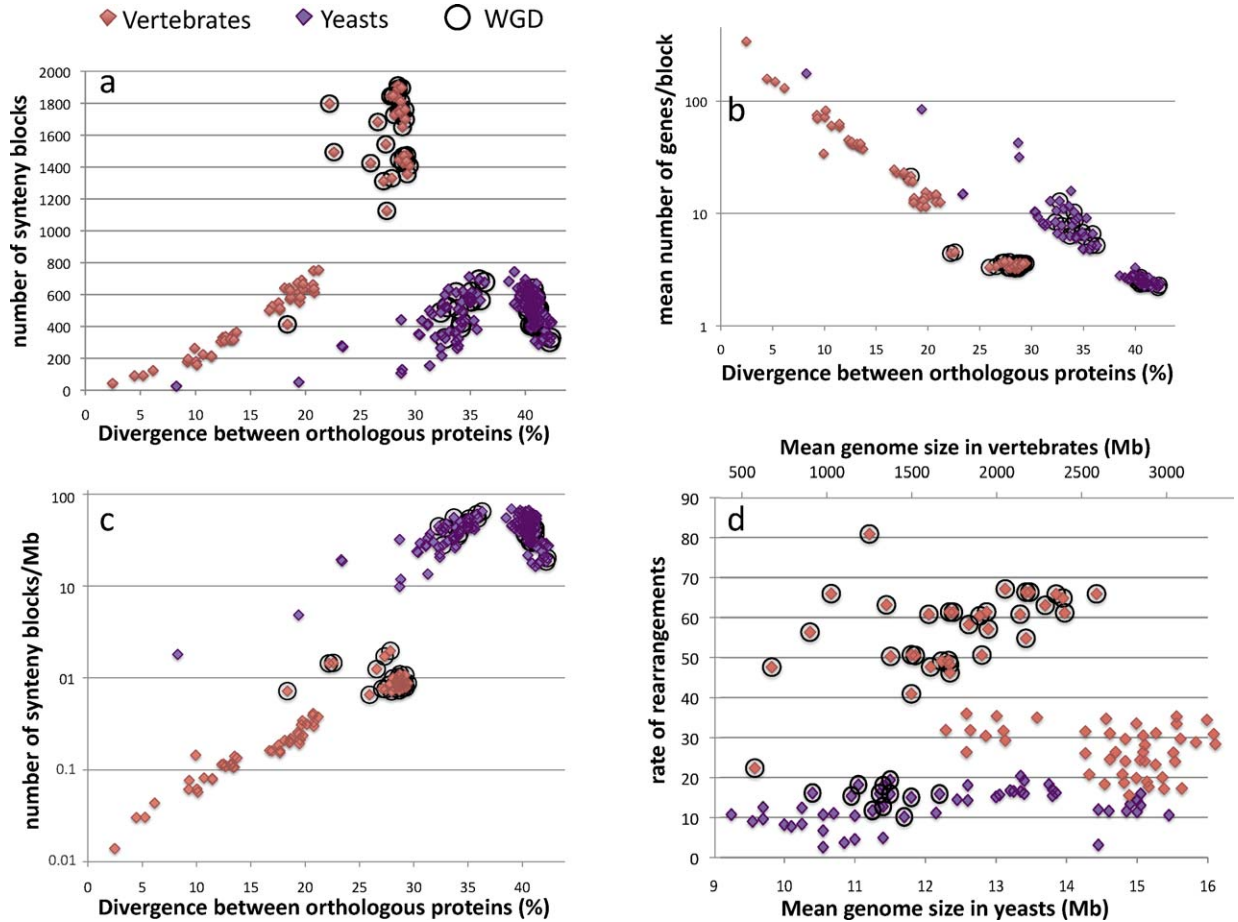


Fig. 2. Comparative analysis of genome reorganization in 13 vertebrate and 18 yeast species (Table 1). Pairs of genes were considered as orthologs if their products were reciprocal best-hits with at least 40% similarity in sequence and their sequences were less than 30% different in length as previously described [39]. Syntenic blocks were defined as series of neighboring pairs of orthologs separated by less than 5 nonneighboring reciprocal best-hits in the two compared genomes. Syntenic blocks were constructed for the 78 and 153 possible pairwise comparisons between the 13 vertebrate (orange diamonds) and 18 yeast (purple diamonds) species, respectively. Black circles indicate pairwise comparisons involving at least one species that undergone a lineage-specific ancestral whole genome duplication (WGD) event (*D. rerio*, *O. latipes* and *T. nigroviridis* in vertebrates and *S. cerevisiae* and *C. glabrata* in yeasts). Protein divergence values correspond to the mean divergence between syntenic reciprocal best hits for each pair of compared genomes. **a.** Evolution of the number of syntenic blocks as a function of protein divergence in vertebrates and yeasts. **b.** Evolution of the number of genes per block with increasing phylogenetic distances. **c.** The number of syntenic blocks is normalized by the mean size of the 2 compared genomes and plotted as log-scale. **d.** The number of syntenic blocks is used to approximate the number of rearrangements (comprising more than 5 genes) accumulated between 2 genomes for all comparisons involving a level of protein divergence lower than 36%. For higher level of divergence, the number of syntenic blocks cannot be used to approximate the number of rearrangements because it decreases with increasing evolutionary distances (see **a.**). Rearrangement rates correspond to the number of rearrangements divided by mean ortholog divergence between the compared paired of species. All 78 possible pairwise comparisons were taken into account for vertebrates while only 55 out of the 153 pairwise comparisons were considered in yeast (below the threshold of 36% divergence).

breakpoints that lead to the reduction of the size of the syntenic blocks below the minimal threshold of 2 neighboring genes (Fig. 2b) and also to a less efficient recognition of orthologous protein when divergence increases (not shown). Two yeast genomes (*S. cerevisiae* and *Candida glabrata*) have also undergone a WGD event followed by rediploidization (circled in black in Fig. 2). But, as opposed to vertebrates, all the comparisons that involve either of these 2 species are scattered throughout the plot because of their intermediate phylogenetic position relative to other yeast species.

For comparable evolutionary distances, where ranges of protein divergence overlap between yeast and vertebrate (i.e. between 8 and 30% of protein divergence), the number

of syntenic blocks between 2 vertebrate genomes is about 6 to 8-fold higher than between 2 yeast genomes (Fig. 2a). This shows that despite a lower evolutionary range, the raw level of chromosome map reorganization is much higher in vertebrate than in yeast. This result shows that, for comparable evolutionary distances, more chromosomal rearrangements occurred on average between 2 vertebrate genomes than between 2 yeast genomes. However, the genome sizes being on average 200 times larger in vertebrates, the physical density of syntenic breakpoints along chromosomes (measured by the number of syntenic blocks per Mb) is consistently higher in yeasts (between 5 and 65 blocks per Mb) than in vertebrates (between 0.01 and 2 blocks per Mb, Fig. 2c).

For both yeast and vertebrate, the average number of shared orthologs per synteny block decreases exponentially with increasing evolutionary distance until it asymptotically reaches the threshold of 2 genes below which it is impossible to recognize conserved synteny blocks (Fig. 2b). Surprisingly, in the overlapping evolutionary range (i.e. between 8 and 30% of divergence), the number of genes per block is higher in yeasts than in vertebrates (54 vs 21 on average, respectively). This higher number of genes per synteny block is best explained by the conjunction of a higher gene density in yeast (only 4 times as many genes in vertebrates than in yeasts while genome sizes are on average 200 times larger) and a higher number of rearrangements in vertebrates that is limited to only 6 to 8 times that of the yeast genomes.

Then, we estimated the rates of rearrangements by approximating the number of synteny blocks to the number of chromosomal rearrangements that occurred since two species diverged from their last common ancestor. Our analysis only accounts for rearrangements involving more than 5 orthologous genes because we tolerate up to 5 consecutive nonsyntenic homologs within a synteny block. For instance, small inversions involving less than 5 genes are not counted here. In yeast, approximating the number of rearrangements to the number of synteny blocks holds true only for pairwise comparisons involving average protein divergence below 36%. For higher levels of divergence, the superimposition of numerous rearrangements leads to the progressive destruction of recognizable synteny blocks and therefore to a strong underestimation of the number of rearrangements that actually occurred (see Fig. 2a and legend of Fig. 2d). The rates of rearrangements correspond to the number of rearrangements that occurred per unit of evolutionary time, which corresponds here to 1% of divergence between orthologous proteins (Fig. 2d). Mean rates of rearrangements are statistically different between the two groups (40 ± 4 vs 13 ± 1 rearrangements/%divergence in vertebrates and yeasts, respectively; T-test P-val = 5.4×10^{-23}). On average, rearrangement rates are 3-fold higher in vertebrates than in yeasts.

In yeast, rearrangement rates do not convincingly correlate with genome sizes ($R^2 = 0.11$, P-val = 0.02) while in vertebrate, rearrangement rates appear to be anti-correlated with genome sizes ($R^2 = -0.60$, P-val = 5.8×10^{-9} , Fig. 2d) because small genomes seem to be more rearranged. However, this anti-correlation uniquely relies on the presence of the small duplicated fish genomes (all 3 fish used in the analysis have the smallest vertebrate genomes) and vanishes when the corresponding data points (circled in black in Fig. 2d) are removed from the analysis ($R^2 = -0.23$; p-value = 0.12). In fish genomes, rearrangement rates are confounded by the lineage specific rediploidisations subsequent to the WGD, which only involve local deletions, not gene-reordering rearrangements. In reality, these fish genomes are remarkably stable and show little rearrangements. For example, Medaka (*Oryzias latipes*) has been subjected to zero interchromosomal event since it splits from the pufferfish (*Tetraodon nigroviridis*) lineage more than 100 Myr ago (Hugues Roest Crolius, pers. com.). Therefore approxi-

imating the number of rearrangements by the number of synteny blocks for these postduplicated genomes might lead to an overestimation of the rearrangement rates in vertebrates. When comparisons involving duplicated fish (*O. latipes*, *D. rerio* and *T. nigroviridis*) and yeast (*S. cerevisiae* and *C. glabrata*) genomes are excluded from the analysis, the mean rearrangement rate remains significantly 2-fold higher in vertebrates than in yeasts (27 ± 2 vs 13 ± 1 rearrangements/%divergence, respectively). It has been shown that both in yeasts and in vertebrates, rearrangement rates are variable between individual lineages [37–40]. For instance, rearrangement rates are smaller between *S. cerevisiae* and *Lachancea waltii* (12.7) than between *S. cerevisiae* and *C. glabrata* (15.9) and also smaller between human and dog (20.9) than between human and mouse (26.5), as previously reported [40,41]. Despite these lineage-specific variations, we show here that the global rates of rearrangements are higher in vertebrates than in yeasts, arguing against the hypothesis of a molecular clock for rearrangements. However, because of very large genome sizes in vertebrates, the average rearrangement rate per Mb is about 50-fold higher in yeasts than in vertebrates (1.04 vs 0.02 rearrangements/%divergence/Mb in yeasts and vertebrates, respectively).

Because vertebrates have emerged within the *Chordata* phylum approximately 450 Myr ago [42], the average rate of 40 ± 4 rearrangements/%divergence can be translated into time unit and would correspond to a rate of 2 rearrangements/Myr (918 blocks on average divided by 450), close to previous estimates on mammalian genome evolution (3.2 chromosomal rearrangements per million years on the mouse branch from the murid rodent ancestor; 3.5 chromosomal rearrangements per million years on the rat branch; and 1.6 chromosomal rearrangements per million years on the human branch [37]). A similar translation would be less reliable in yeast because estimated emergence time for the *Saccharomycotina* subphylum vary between 400 and 1000 Myr ago [43] and also because at large evolutionary distance (ortholog divergence greater than 36%) the number of synteny blocks cannot be used to approximate the number of rearrangements that actually happened.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

Acknowledgements

We thank Hugues Roest Crolius for critical reading of the manuscript and for our regular scientific discussions that have contributed to the realization of this work. We are highly grateful to Jean-Luc Souciet, Bernard Dujon and Claude Gaillardin for having given rise to the Genolevures adventure and for allowing us to contribute.

References

- [1] R.S. Kucherlapati, R.P. Creagan, E.A. Nichols, D.S. Borgaonkar, F.H. Ruddle, Synteny relationships of four human genes: mannose phosphate isomerase to pyruvate kinase-3 and triose phosphate isomerase to lactate dehydrogenase-B, *Cytogenet. Cell. Genet.* 14 (1975) 364–367.

- [2] J.C. McAvin, D. Patterson, J.E. Womack, Mapping of bovine PRGS and PAIS genes in hybrid somatic cells: syntenic conservation with human chromosome 21, *Biochem. Genet.* 26 (1988) 9–18.
- [3] J.D. Minna, P.A. Lalley, U. Francke, Comparative mapping using somatic cell hybrids, *In Vitro* 12 (1976) 726–733.
- [4] H.V. van Someren, H. Beyersbergen van, J. de Wit, Proceedings: evidence for synteny between the human loci for fumarate hydratase, UDP glucose pyrophosphorylase, 6-phosphogluconate dehydrogenase, phosphoglucomutase1, and peptidase-C in man-Chinese hamster somatic cell hybrids, *Cytogenet. Cell. Genet.* 13 (1974) 150–152.
- [5] R.J. Leach, P. O'Connell, Mapping of mammalian genomes with radiation (Goss and Harris) hybrids, *Adv. Genet.* 33 (1995) 63–99.
- [6] H. Muller, A. Thierry, J.Y. Coppee, C. Gouyette, C. Hennequin, O. Sismeiro, E. Talla, B. Dujon, C. Fairhead, Genomic polymorphism in the population of *Candida glabrata*: gene copy-number variation and chromosomal translocations, *Fungal. Genet. Biol.* 46 (2009) 264–276.
- [7] E. Naumova, G. Naumov, P. Fournier, H.V. Nguyen, C. Gaillardin, Chromosomal polymorphism of the yeast *Yarrowia lipolytica* and related species: electrophoretic karyotyping and hybridization with cloned genes, *Curr. Genet.* 23 (1993) 450–454.
- [8] R.F. Petersen, T. Nilsson-Tillgren, J. Piskur, Karyotypes of *Saccharomyces sensu lato* species, *Int J Syst Bacteriol* 49 (Pt 4) (1999) 1925–1931.
- [9] S. Polakova, C. Blume, J.A. Zarate, M. Mentel, D. Jorck-Ramberg, J. Stenderup, J. Piskur, Formation of new chromosomes as a virulence mechanism in yeast *Candida glabrata*, *Proc Natl Acad Sci U S A* 106 (2009) 2688–2693.
- [10] M. Spirek, J. Yang, C. Groth, R.F. Petersen, R.B. Langkjaer, E.S. Naumova, P. Sulo, G.L. Naumov, J. Piskur, High-rate evolution of *Saccharomyces sensu lato* chromosomes, *FEMS Yeast Res* 3 (2003) 363–373.
- [11] A. Vaughan-Martini, A. Martini, G. Cardinali, Electrophoretic karyotyping as a taxonomic tool in the genus *Saccharomyces*, *Antonie Van Leeuwenhoek* 63 (1993) 145–156.
- [12] E. Passarge, B. Horsthemke, R.A. Farber, Incorrect use of the term synteny, *Nat. Genet.* 23 (1999) 387.
- [13] K. Hartung, D. Frishman, A. Hinnen, S. Wolf, Single-read sequence tags of a limited number of genomic DNA fragments provide an inexpensive tool for comparative genome analysis, *Yeast* 14 (1998) 1327–1332.
- [14] R.S. Keogh, C. Seoighe, K.H. Wolfe, Evolution of gene order and chromosome number in *Saccharomyces*, *Kluyveromyces* and related fungi, *Yeast* 14 (1998) 443–457.
- [15] R.B. Langkjaer, M.L. Nielsen, P.R. Daugaard, W. Liu, J. Piskur, Yeast chromosomes have been significantly reshaped during their evolutionary history, *J. Mol. Biol.* 304 (2000) 217–288.
- [16] O. Ozier-Kalogeropoulos, A. Malpertuy, J. Boyer, F. Tekaia, B. Dujon, Random exploration of the *Kluyveromyces lactis* genome and comparison with that of *Saccharomyces cerevisiae*, *Nucleic Acids Res.* 26 (1998) 5511–5524.
- [17] H. Roest Crolius, O. Jaillon, A. Bernot, C. Dasilva, L. Bouneau, C. Fischer, C. Fizames, P. Wincker, P. Brottier, F. Quetier, W. Saurin, J. Weissenbach, Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence, *Nat. Genet.* 25 (2000) 235–238.
- [18] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczyk, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minooshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H. Pan, J. Ramsler, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordtsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kasprzyk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, A. Patrino, M.J. Morgan, J. Szustakowski, P. de Jong, J.J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y.J. Chen, Initial sequencing and analysis of the human genome, *Nature* 409 (2001) 860–921.
- [19] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huse, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhall, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M.L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferreira, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ilegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfankoch, E. Pratts, V. Puri, H. Qureshi, M. Rearson, R. Rodriguez, Y.H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N.N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J.F. Abril, R. Guigo, M.J. Campbell, K.V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yoosheph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y.H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, G. Glasser, A. Glodde, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh, X. Zhu, The sequence of the human genome, *Science* 291 (2001) 1304–1351.
- [20] J. Souciet, M. Aigle, F. Artiguenave, G. Blandin, M. Bolotin-Fukuhara, E. Bon, P. Brottier, S. Casaregola, J. de Montigny, B. Dujon, P. Durrrens, C. Gaillardin, A. Lepingle, B. Llorente, A. Malpertuy, C. Neuveglise, O. Ozier-Kalogeropoulos, S. Potier, W. Saurin, F. Tekaia, C. Toffano-Nioche, M. Wesolowski-Louvel, P. Wincker, J. Weissenbach, Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies, *FEBS Lett.* 487 (2000) 3–12.
- [21] B. Llorente, A. Malpertuy, C. Neuveglise, J. de Montigny, M. Aigle, F. Artiguenave, G. Blandin, M. Bolotin-Fukuhara, E. Bon, P. Brottier, S. Casaregola, P. Durrrens, C. Gaillardin, A. Lepingle, O. Ozier-Kalogeropoulos, S. Potier, W. Saurin, F. Tekaia, C. Toffano-Nioche, M. Wesolowski-Louvel, P. Wincker, J. Weissenbach, J. Souciet, B. Dujon, Genomic exploration of the hemiascomycetous yeasts: 18. Comparative analysis of chromosome maps and synteny with *Saccharomyces cerevisiae*, *FEBS Lett.* 487 (2000) 101–112.
- [22] B. Dujon, Evolutionary genomics of yeasts, in: Caetano-Anolles (Ed.), *Evolutionary genomics and systems biology*, Wiley-Blackwell, 2010.
- [23] J.N. Volff, Vertebrate genomes, Karger, 2006.

- [24] D.M. Ruderfer, S.C. Pratt, H.S. Seidel, L. Kruglyak, Population genomic analysis of outcrossing and recombination in yeast, *Nat. Genet.* 38 (2006) 1077–1081.
- [25] I.J. Tsai, D. Bensasson, A. Burt, V. Koufopanou, Population genomics of the wild yeast *Saccharomyces paradoxus*: quantifying the life cycle, *Proc. Natl. Acad. Sci. U S A* 105 (2008) 4957–4962.
- [26] C. Seoighe, K.H. Wolfe, Extent of genomic rearrangement after genome duplication in yeast, *Proc. Natl. Acad. Sci. U S A* 95 (1998) 4447–4452.
- [27] J.W. Drake, B. Charlesworth, D. Charlesworth, J.F. Crow, Rates of spontaneous mutation, *Genetics* 148 (1998) 1667–1686.
- [28] K.T. Nishant, N.D. Singh, E. Alani, Genomic mutation rates: what high-throughput methods can tell us, *Bioessays* 31 (2009) 912–920.
- [29] J.C. Roach, G. Glusman, A.F. Smit, C.D. Huff, R. Hubley, P.T. Shannon, L. Rowen, K.P. Pant, N. Goodman, M. Bamshad, J. Shendure, R. Drmanac, L.B. Jorde, L. Hood, D.J. Galas, Analysis of genetic inheritance in a family quartet by whole-genome sequencing, *Science* 328 (2010) 636–639.
- [30] M. Lynch, W. Sung, K. Morris, N. Coffey, C.R. Landry, E.B. Dopman, W.J. Dickinson, K. Okamoto, S. Kulkarni, D.L. Hartl, W.K. Thomas, A genome-wide view of the spectrum of spontaneous mutations in yeast, *Proc. Natl. Acad. Sci. U S A* 105 (2008) 9272–9277.
- [31] K.T. Nishant, W. Wei, E. Mancera, J.L. Argueso, A. Schlattl, N. Delhomme, X. Ma, C.D. Bustamante, J.O. Korbel, Z. Gu, L.M. Steinmetz, E. Alani, The baker's yeast diploid genome is remarkably stable in vegetative growth and meiosis, *PLoS Genet.* 6 (2010).
- [32] N. Arnheim, P. Calabrese, Understanding what determines the frequency and pattern of human germline mutations, *Nat. Rev. Genet.* 10 (2009) 478–488.
- [33] B. Dujon, Yeast evolutionary genomics, *Nat. Rev. Genet.* 11 (2010) 512–524.
- [34] B. Dujon, Yeasts illustrate the molecular mechanisms of eukaryotic genome evolution, *Trends Genet.* 22 (2006) 375–387.
- [35] B. Dujon, D. Sherman, G. Fischer, P. Durrens, S. Casaregola, I. Lafontaine, J. De Montigny, C. Marck, C. Neuvéglise, E. Talla, N. Goffard, L. Frangeul, M. Aigle, V. Anthouard, A. Babour, V. Barbe, S. Barnay, S. Blanchin, J.M. Beckerich, E. Beyne, C. Bleykasten, A. Boisrame, J. Boyer, L. Cattolico, F. Confaniolieri, A. De Daruvar, L. Despons, E. Fabre, C. Fairhead, H. Ferry-Dumazet, A. Groppi, F. Hantraye, C. Hennequin, N. Jauniaux, P. Joyet, R. Kachouri, A. Kerrest, R. Koszul, M. Lemaire, I. Lesur, L. Ma, H. Muller, J.M. Nicaud, M. Nikolski, S. Oztas, O. Ozier-Kalogeropoulos, S. Pellenz, S. Potier, G.F. Richard, M.L. Straub, A. Suleau, D. Swennen, F. Tekaiia, M. Wesolowski-Louvel, E. Westhof, B. Wirth, M. Zeniou-Meyer, I. Zivanovic, M. Bolotin-Fukuhara, A. Thierry, C. Bouchier, B. Caudron, C. Scarpelli, C. Gaillardin, J. Weissenbach, P. Wincker, J.L. Souciet, Genome evolution in yeasts, *Nature* 430 (2004) 35–44.
- [36] O. Jaillon, J.M. Aury, F. Brunet, J.L. Petit, N. Stange-Thomann, E. Mauceli, L. Bouneau, C. Fischer, C. Ozouf-Costaz, A. Bernot, S. Nicaud, D. Jaffe, S. Fisher, G. Lutfalla, C. Dossat, B. Segurens, C. Dasilva, M. Salanoubat, M. Levy, N. Boudet, S. Castellano, V. Anthouard, C. Jubin, V. Castelli, M. Katinka, B. Vacherie, C. Biemont, Z. Skalli, L. Cattolico, J. Poulain, V. De Berardinis, C. Cruaud, S. Duprat, P. Brottier, J.P. Coutanceau, J. Gouzy, G. Parra, G. Lardier, C. Chappelle, K.J. McKernan, P. McEwan, S. Bosak, M. Kellis, J.N. Volf, R. Guigo, M.C. Zody, J. Mesirov, K. Lindblad-Toh, B. Birren, C. Nusbaum, D. Kahn, M. Robinson-Rechavi, V. Laudet, V. Schachter, F. Quetier, W. Saurin, C. Scarpelli, P. Wincker, E.S. Lander, J. Weissenbach, H. Roest Crolius, Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype, *Nature* 431 (2004) 946–957.
- [37] G. Bourque, E.M. Zdobnov, P. Bork, P.A. Pevzner, G. Tesler, Comparative architectures of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages, *Genome Res.* 15 (2005) 98–110.
- [38] D.W. Burt, C. Bruley, I.C. Dunn, C.T. Jones, A. Ramage, A.S. Law, D.R. Morrice, I.R. Paton, J. Smith, D. Windsor, A. Sazanov, R. Fries, D. Waddington, The dynamics of chromosome evolution in birds and mammals, *Nature* 402 (1999) 411–413.
- [39] G. Fischer, S.A. James, I.N. Roberts, S.G. Oliver, E.J. Louis, Chromosomal evolution in *Saccharomyces*, *Nature* 405 (2000) 451–454.
- [40] G. Fischer, E.P. Rocha, F. Brunet, M. Vergassola, B. Dujon, Highly variable rates of genome rearrangements between hemiascomycetous yeast lineages, *PLoS Genet.* 2 (2006) e32.
- [41] W.J. Murphy, D.M. Larkin, A. Everts-van der Wind, G. Bourque, G. Tesler, L. Auviel, J.E. Beever, B.P. Chowdhary, F. Galibert, L. Gatzke, C. Hitte, S.N. Meyers, D. Milan, E.A. Ostrander, G. Pape, H.G. Parker, T. Raudsepp, M.B. Rogatcheva, L.B. Schook, L.C. Skow, M. Welge, J.E. Womack, J.S. O'Brien, P.A. Pevzner, H.A. Lewin, Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps, *Science* 309 (2005) 613–617.
- [42] S.B. Hedges, The origin and evolution of model organisms, *Nat. Rev. Genet.* 3 (2002) 838–849.
- [43] J.W. Taylor, M.L. Berbee, Dating divergences in the Fungal Tree of Life: review and new analyses, *Mycologia* 98 (2006) 838–849.
- [44] T. Jones, N.A. Federspiel, H. Chibana, J. Dungan, S. Kalman, B.B. Magee, G. Newport, Y.R. Thorstenson, N. Agabian, P.T. Magee, R.W. Davis, S. Scherer, The diploid genome sequence of *Candida albicans*, *Proc Natl Acad Sci USA* 101 (2004) 7329–7334.
- [45] A.P. Jackson, J.A. Gamble, T. Yeomans, G.P. Moran, D. Saunders, D. Harris, M. Aslett, J.F. Barrell, G. Butler, F. Citiulo, D.C. Coleman, P.W. de Groot, T.J. Goodwin, M.A. Quail, J. McQuillan, C.A. Munro, A. Pain, R.T. Poulter, M.A. Rajandream, H. Renauld, M.J. Spiering, A. Tivey, N.A. Gow, B. Barrell, D.J. Sullivan, M. Berriman, Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*, *Genome Res.* 19 (2009) 2231–2244.
- [46] G. Butler, M.D. Rasmussen, M.F. Lin, M.A. Santos, S. Sakthikumar, C.A. Munro, E. Rheinbay, M. Grabherr, A. Forche, J.L. Reedy, I. Agrafioti, M.B. Arnaut, S. Bates, A.J. Brown, S. Brunke, M.C. Costanzo, D.A. Fitzpatrick, P.W. de Groot, D. Harris, L.L. Hoyer, B. Hube, F.M. Klis, C. Kodira, N. Lennard, M.E. Logue, R. Martin, A.M. Neiman, E. Nikolaou, M.A. Quail, J. Quinn, M.C. Santos, F.F. Schmitzberger, G. Sherlock, P. Shah, K.A. Silverstein, M.S. Skrzypek, D. Soll, R. Staggs, I. Stansfield, M.P. Stumpf, P.E. Sudbery, T. Srikantha, Q. Zeng, J. Berman, M. Berriman, J. Heitman, N.A. Gow, M.C. Lorenz, B.W. Birren, M. Kellis, C.A. Cuomo, Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes, *Nature* 459 (2009) 657–662.
- [47] F.S. Dietrich, S. Voegeli, S. Brachat, A. Lerch, K. Gates, S. Steiner, C. Mohr, R. Pohlmann, P. Luedi, S. Choi, R.A. Wing, A. Flavier, T.D. Gaffney, P. Philippsen, The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome, *Science* 304 (2004) 304–307.
- [48] S. Génolevures Consortium, J. L. B. Dujon, C. Gaillardin, M. Johnston, P.V. Baret, P. Cliften, D.J. Sherman, J. Weissenbach, E. Westhof, P. Wincker, C. Jubin, J. Poulain, V. Barbe, B. Segurens, F. Artiguenave, V. Anthouard, B. Vacherie, M.E. Val, R.S. Fulton, P. Minx, R. Wilson, P. Durrens, G. Jean, C. Marck, T. Martin, M. Nikolski, T. Rolland, M.L. Seret, S. Casaregola, L. Despons, C. Fairhead, G. Fischer, I. Lafontaine, V. Leh, M. Lemaire, J. de Montigny, C. Neuvéglise, A. Thierry, I. Blanc-Lenle, C. Bleykasten, J. Diffels, E. Fritsch, L. Frangeul, A. Goefflon, N. Jauniaux, R. Kachouri-Lafond, C. Payen, S. Potier, L. Pribylova, C. Ozanne, G.F. Richard, C. Sacerdot, M.L. Straub, E. Talla, Comparative genomics of protoploid *Saccharomycetaceae*, *Genome Res.* (2009).
- [49] M. Kellis, B.W. Birren, E.S. Lander, Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*, *Nature* 428 (2004) 617–624.
- [50] K. De Schutter, Y.C. Lin, P. Tiels, A. Van Hecke, S. Glinka, J. Weber-Lehmann, P. Rouze, Y. Van de Peer, N. Callewaert, Genome sequence of the recombinant protein production host *Pichia pastoris*, *Nat Biotechnol* 27 (2009) 561–566.
- [51] D. Mattanovich, A. Graf, J. Stadlmann, M. Dragosits, A. Redl, M. Maurer, M. Kleinheinz, M. Sauer, F. Altmann, B. Gasser, Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*, *Microb. Cell Fact.* 8 (2009) 29.
- [52] T.W. Jeffries, I.V. Grigoriev, J. Grimwood, J.M. Laplaza, A. Aerts, H. Salamov, J. Schmutz, E. Lindquist, P. Dehal, H. Shapiro, Y.S. Jin, V. Passoth, P.M. Richardson, Genome sequence of the lignocellulose-converting and xylose-fermenting yeast *Pichia stipitis*, *Nat. Biotechnol.* 25 (2007) 319–326.
- [53] A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S.G. Oliver, Life with 6000 genes, *Science* 274 (546) (1996) 547–563.
- [54] K. Lindblad-Toh, C.M. Wade, T.S. Mikkelsen, E.K. Karlsson, D.B. Jaffe, M. Kamal, M. Clamp, J.L. Chang, E.J. Kulbokas 3rd, M.C. Zody, E. Mauceli, X. Xie, M. Breen, R.K. Wayne, E.A. Ostrander, C.P. Ponting, F. Galibert, D.R. Smith, P.J. DeJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C.W. Chin, A. Cook, J. Cuff, M.J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber, C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K.P. Koepfli, H.G. Parker, J.P. Pollinger, S.M. Searle, N.B. Sutter, R. Thomas, C. Webber, J. Baldwin, A. Abebe, A. Abouelleil, L. Aftuck, M. Ait-Zahra, T. Aldredge, N. Allen, P. An, S. Anderson, C. Antoine, H. Arachchi, A. Aslam, L. Ayyotte, P. Bachantsang, A. Barry, T. Bayul, M. Benamara, A. Berlin, D. Bessette, B. Blitshteyn, T. Bloom, J. Blye, L. Boguslavskiy, C. Bonnet, B. Boukhgalter, A. Brown, P. Cahill, N. Calixte, J. Camarata, Y. Cheshatsang, J. Chu, M. Citroen, A. Collymore, P. Cooke, T. Dawoe, R. Daza, K. Decktor, S. DeGray, N. Dhargay, K. Dooley, P. Dorje, K. Dorjee, L. Dorris, N. Duffey, A. Dupes, O. Egbriemolen, R. Elong, J. Falk, A. Farina, S. Faro, D. Ferguson, P. Ferreira, S. Fisher, M. FitzGerald, K. Foley, C. Foley, A. Franke, D. Friedrich, D. Gage, M. Garber, G. Gearin, G. Giannoukos, T. Goode, A. Goyette, J. Graham, E. Grandbois, K. Gyaltsen, N. Hafez, D. Hagopian, B. Hagos, J. Hall, C. Healy, R. Hegarty, T. Honan, A. Horn, N. Houde, L. Hughes, L. Hunnicutt, M. Husby, B. Jester, C. Jones, A. Kamat, B. Kanga, C. Kells, D. Khazanovich, A.C. Kieu, P.

- Kisner, M. Kumar, K. Lance, T. Landers, M. Lara, W. Lee, J.P. Leger, N. Lennon, L. Leuper, S. LeVine, J. Liu, X. Liu, Y. Lokvitsang, T. Lokvitsang, A. Lui, J. Macdonald, J. Major, R. Marabella, K. Maru, C. Matthews, S. McDonough, T. Mehta, J. Meldrim, A. Melnikov, L. Meneus, A. Mihalev, T. Mihova, K. Miller, R. Mittelman, V. Mengla, L. Mulrain, G. Munson, A. Navidi, J. Naylor, T. Nguyen, N. Nguyen, C. Nguyen, R. Nicol, N. Norbu, C. Norbu, X. Novod, T. Nyjima, P. Olandt, B. O'Neill, K. O'Neill, S. Osman, L. Oyono, C. Patti, D. Perrin, P. Phunkhang, F. Pierre, M. Priest, A. Rachupka, S. Raghuraman, R. Rameau, V. Ray, C. Raymond, F. Rege, C. Rise, J. Rogers, P. Rogov, J. Sahalie, S. Settipalli, T. Sharpe, T. Shea, M. Sheehan, N. Sherpa, J. Shi, D. Shih, J. Sloan, C. Smith, T. Sparrow, J. Stalker, N. Stange-Thomann, S. Stavropoulos, C. Stone, S. Stone, S. Sykes, P. Tchuinga, P. Tenzing, S. Tesfaye, D. Thoulutsang, Y. Thoulutsang, K. Topham, I. Topping, T. Tsamla, H. Vassiliev, V. Venkataraman, A. Vo, T. Wangchuk, T. Wangdi, M. Weiland, J. Wilkinson, A. Wilson, S. Yadav, S. Yang, X. Yang, G. Young, Q. Yu, J. Zainoun, L. Zembek, A. Zimmer, E.S. Lander, Genome sequence, comparative analysis and haplotype structure of the domestic dog, *Nature* 438 (2005) 803–819.
- [55] C.M. Wade, E. Giulotto, S. Sigurdsson, M. Zoli, S. Gnerre, F. Imsland, T.L. Lear, D.L. Adelson, E. Bailey, R.R. Bellone, H. Blocker, O. Distl, R.C. Edgar, M. Garber, T. Leeb, E. Mauceli, J.N. MacLeod, M.C. Penedo, J.M. Raison, T. Sharpe, J. Vogel, L. Andersson, D.F. Antczak, T. Biagi, M.M. Binns, B.P. Chowdhary, S.J. Coleman, G. Della Valle, S. Fryc, G. Guerin, T. Hasegawa, E.W. Hill, J. Jurka, A. Kiiialainen, G. Lindgren, J. Liu, E. Magnani, J.R. Mickelson, J. Murray, S.G. Nergadze, R. Onofrio, S. Pedroni, M.F. Piras, T. Raudsepp, M. Rocchi, K.H. Roed, O.A. Ryder, S. Searle, L. Skow, J.E. Swinburne, A.C. Syvanen, T. Tozaki, S.J. Valberg, M. Vaudin, J.R. White, M.C. Zody, E.S. Lander, K. Lindblad-Toh, Genome sequence, comparative analysis, and population genetics of the domestic horse, *Science* 326 (2009) 865–867.
- [56] Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution, *Nature* 432 (2004) 695–716.
- [57] R.A. Gibbs, J. Rogers, M.G. Katze, R. Bumgarner, G.M. Weinstock, E.R. Mardis, K.A. Remington, R.L. Strausberg, J.C. Venter, R.K. Wilson, M.A. Batzer, C.D. Bustamante, E.E. Eichler, M.W. Hahn, R.C. Hardison, K.D. Makova, W. Miller, A. Milosavljevic, R.E. Palermo, A. Siepel, J.M. Sikela, T. Attaway, S. Bell, K.E. Bernard, C.J. Buhay, M.N. Chandrabose, M. Dao, C. Davis, K.D. Delehaunty, Y. Ding, H.H. Dinh, S. Dugan-Rocha, L.A. Fulton, R.A. Gabisi, T.T. Garner, J. Godfrey, A.C. Hawes, J. Hernandez, S. Hines, M. Holder, J. Hume, S.N. Jiangiani, Y. Joshi, Z.M. Khan, E.F. Kirkness, A. Cree, R.G. Fowler, S. Lee, L.R. Lewis, Z. Li, Y.S. Liu, S.M. Moore, D. Muzny, L.V. Nazareth, D.N. Ngo, G.O. Okwuonu, G. Pai, D. Parker, H.A. Paul, C. Pfannkoch, C.S. Pohl, Y.H. Rogers, S.J. Ruiz, A. Sabo, J. Santibanez, B.W. Schneider, S.M. Smith, E. Sodergren, A.F. Svatek, T.R. Utterback, S. Vattathil, W. Warren, C.S. White, A.T. Chinwalla, Y. Feng, A.L. Halpern, L.W. Hillier, X. Huang, P. Minx, J.O. Nelson, K.H. Pepin, X. Qin, G.G. Sutton, E. Venter, B.P. Walenz, J.W. Wallis, K.C. Worley, S.P. Yang, S.M. Jones, M.A. Marra, M. Rocchi, J.E. Schein, R. Baertsch, L. Clarke, M. Csuros, J. Glasscock, R.A. Harris, P. Havlak, A.R. Jackson, H. Jiang, Y. Liu, D.N. Messina, Y. Shen, H.X. Song, T. Wylie, L. Zhang, E. Birney, K. Han, M.K. Konkel, J. Lee, A.F. Smit, B. Ullmer, H. Wang, J. Xing, R. Burhans, Z. Cheng, J.E. Karro, J. Ma, B. Raney, X. She, M.J. Cox, J.P. Demuth, L.J. Dumas, S.G. Han, J. Hopkins, A. Karimpour-Fard, Y.H. Kim, J.R. Pollack, T. Vinar, C. Addo-Quaye, J. Degenhardt, A. Denby, M.J. Hubisz, A. Indap, C. Kosiol, B.T. Lahn, H.A. Lawson, A. Marklein, R. Nielsen, E.J. Vallender, A.G. Clark, B. Ferguson, R.D. Hernandez, K. Hirani, H. Kehrer-Sawatzki, J. Kolb, S. Patil, L.L. Pu, Y. Ren, D.G. Smith, D.A. Wheeler, I. Schenck, E.V. Ball, R. Chen, D.N. Cooper, B. Giardina, F. Hsu, W.J. Kent, A. Lesk, D.L. Nelson, E. O'Brien, W. K. Prufer, P.D. Stenson, J.C. Wallace, H. Ke, X.M. Liu, P. Wang, A.P. Xiang, F. Yang, G.P. Barber, D. Haussler, D. Karolchik, A.D. Kern, R.M. Kuhn, K.E. Smith, A.S. Zwiag, Evolutionary and biomedical insights from the rhesus macaque genome, *Science* 316 (2007) 222–234.
- [58] R.H. Waterston, K. Lindblad-Toh, E. Birney, J. Rogers, J.F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S.E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M.R. Brent, D.G. Brown, S.D. Brown, C. Bult, J. Burton, J. Butler, R.D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A.T. Chinwalla, D.M. Church, M. Clamp, C. Clee, F.S. Collins, L.L. Cook, R.R. Copley, A. Coulson, O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K.D. Delehaunty, J. Deri, E.T. Dermitzakis, C. Dewey, N.J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D.M. Dunn, S.R. Eddy, L. Eltnitski, R.D. Emes, P. Eswar, E. Eyraas, A. Felsenfeld, G.A. Fedell, P. Flicek, K. Foley, W.N. Frankel, L.A. Fulton, R.S. Fulton, T.S. Furey, D. Gage, R.A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Graffham, T.A. Graves, E.D. Green, S. Gregory, R. Guigo, M. Guyer, R.C. Hardison, D. Haussler, Y. Hayashizaki, L.W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D.B. Jaffe, L.S. Johnson, M. Jones, T.A. Jones, A. Joy, M. Kamal, E.K. Karlsson, D. Karolchik, A. Kasprzyk, J. Kawai, E. Keibler, C. Kells, W.J. Kent, A. Kirby, D.L. Kolbe, I. Korf, R.S. Kucherlapati, E.J. Kulbokas, D. Kulp, T. Landers, J.P. Leger, S. Leonard, I. Letunic, R. Levine, J. Li, M. Li, C. Lloyd, S. Lucas, B. Ma, D.R. Maglott, E.R. Mardis, L. Matthews, E. Mauceli, J.H. Mayer, M. McCarthy, W.R. McCombie, S. McLaren, K. McClay, J.D. McPherson, J. Meldrim, B. Meredith, J.P. Mesirov, W. Miller, T.L. Miner, E. Mongin, K.T. Montgomery, M. Morgan, R. Mott, J.C. Mullikin, D.M. Muzny, W.E. Nash, J.O. Nelson, M.N. Nhan, R. Nicol, Z. Ning, C. Nusbaum, M.J. O'Connor, Y. Okazaki, K. Oliver, E. Overton-Larty, L. Pachter, G. Parra, K.H. Pepin, J. Peterson, P. Pevzner, R. Plumb, C.S. Pohl, A. Poliakov, T.C. Ponce, C.P. Ponting, S. Potter, M. Quail, A. Reymond, B.A. Roe, K.M. Roskin, E.M. Rubin, A.G. Rust, R. Santos, V. Sapojnikov, B. Schultz, J. Schultz, M.S. Schwartz, S. Schwartz, C. Scott, S. Seaman, S. Searle, T. Sharpe, A. Sheridan, R. Shownkeen, S. Sims, J.B. Singer, G. Slater, A. Smit, D.R. Smith, B. Spencer, A. Stabenau, N. Stange-Thomann, C. Sugnet, M. Suyama, G. Tesler, J. Thompson, D. Torrents, E. Trevaskis, J. Tromp, C. Ucla, A. Ureta-Vidal, J.P. Vinson, A.C. Von Niederhauser, C.M. Wade, M. Wall, R.J. Weber, R.B. Weiss, M.C. Wendt, A.P. West, K. Wetterstrand, R. Wheeler, S. Whelan, J. Wierzbowski, D. Willey, S. Williams, R.K. Wilson, E. Winter, K.C. Worley, D. Wyman, S. Yang, S.P. Yang, E.M. Zdobnov, M.C. Zody, E.S. Lander, Initial sequencing and comparative analysis of the mouse genome, *Nature* 420 (2002) 520–562.
- [59] T.S. Mikkelsen, M.J. Wakefield, B. Aken, C.T. Amemiya, J.L. Chang, S. Duke, M. Garber, A.J. Gentles, L. Goodstadt, A. Heger, J. Jurka, M. Kamal, E. Mauceli, S.M. Searle, T. Sharpe, M.L. Baker, M.A. Batzer, P.V. Benos, K. Belov, M. Clamp, A. Cook, J. Cuff, R. Das, L. Davidow, J.E. Deakin, M.J. Fazzari, J.L. Glass, M. Grabherr, J.M. Greally, W. Gu, T.A. Hore, G.A. Huttley, M. Kleber, R.L. Jirtle, E. Koina, J.T. Lee, S. Mahony, M.A. Marra, R.D. Miller, R.D. Nicholls, M. Oda, A.T. Papenfuss, Z.E. Parra, D.D. Pollock, D.A. Ray, J.E. Schein, T.P. Speed, K. Thompson, J.L. VandeBerg, C.M. Wade, J.A. Walker, P.D. Waters, C. Webber, J.R. Weidman, X. Xie, M.C. Zody, J.A. Graves, C.P. Ponting, M. Breen, P.B. Samollow, E.S. Lander, K. Lindblad-Toh, Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences, *Nature* 447 (2007) 167–177.
- [60] M. Kasahara, K. Naruse, S. Sasaki, Y. Nakatani, W. Qu, B. Ahsan, T. Yamada, Y. Nagayasu, K. Doi, Y. Kasai, T. Jindo, D. Kobayashi, A. Shimada, A. Toyoda, Y. Kuroki, A. Fujiyama, T. Sasaki, A. Shimizu, S. Asakawa, N. Shimizu, S. Hashimoto, J. Yang, Y. Lee, K. Matsushima, S. Sugano, M. Sakaizumi, T. Narita, K. Ohishi, S. Haga, F. Ohta, H. Nomoto, K. Nogata, T. Morishita, T. Endo, I.T. Shin, H. Takeda, S. Morishita, Y. Kohara, The medaka draft genome and insights into vertebrate genome evolution, *Nature* 447 (2007) 714–719.
- [61] Initial sequence of the chimpanzee genome and comparison with the human genome, *Nature* 437 (2005) 69–87.
- [62] R.A. Gibbs, G.M. Weinstock, M.L. Metzker, D.M. Muzny, E.J. Sodergren, S. Scherer, G. Scott, D. Steffen, K.C. Worley, P.E. Burch, G. Okwuonu, S. Hines, L. Lewis, C. DeRamo, O. Delgado, S. Dugan-Rocha, G. Miner, M. Morgan, A. Hawes, R. Gill, R.A. Celera, M.D. Holt, P.G. Adams, H. Amanatides, M. Baden-Tillson, S. Barnstead, C.A. Chin, S. Evans, C. Ferriera, A. Fosler, Z. Glodet, D. Gu, C.L. Jennings, T. Kraft, C.M. Nguyen, C. Pfannkoch, G.G. Sitter, J.C. Sutton, T. Venter, D. Woodage, H.M. Smith, E. Lee, P. Gustafson, A. Cahill, L. Kana, K. Doucette-Stamm, K. Weinstock, R.B. Fecthel, D.M. Weiss, E.D. Dunn, R.W. Green, G.G. Blakesley, P.J. Bouffard, K. De Jong, B. Osoegawa, M. Zhu, J. Marra, I. Schein, C. Bosdet, S. Fjell, M. Jones, C. Krzywinski, A. Mathewson, N. Siddiqui, J. Wye, S. McPherson, C.M. Zhao, J. Fraser, S. Shetty, K. Shatsman, J. Geer, S. Chen, W.C. Abramzon, P.H. Nierman, R. Havlak, K.J. Chen, A. Durbin, Y. Egan, X.Z. Ren, B. Song, Y. Li, X. Liu, S. Qin, A.J. Cawley, L.M. Cooney, K. D'Souza, J.Q. Martin, M.L. Wu, A.R. Gonzalez-Garay, K.J. Jackson, M.P. Kalafus, A. McLeod, D. Milosavljevic, A. Virk, D.A. Volkov, Z. Wheeler, J.A. Zhang, E.E. Bailey, E. Eichler, E. Tuzun, E. Birney, A. Mongin, C. Ureta-Vidal, E. Woodwark, P. Zdobnov, M. Bork, D. Suyama, M. Torrents, B.J. Alexandersson, J.M. Trask, H. Young, H. Huang, H. Wang, S. Xing, D. Daniels, J. Getz, K. Schmidt, U. Stevens, J. Vitti, F. Wingrove, M. Camara, J.F. Mar Alba, R. Abril, A. Guigo, I. Smit, E.M. Dubchak, O. Rubin, A. Couronne, N. Poliakov, D. Hubner, C. Ganten, O. Goesele, T. Hummel, Y.A. Kreitler, J. Lee, H. Monti, H. Schulz, H. Zimdahl, H. Himmelbauer, H.J. Lehrach, S. Jacob, J. Bromberg, M.I. Gullings-Handley, A.E. Jensen-Seaman, J. Kwitek, D. Lazar, P.J. Pasko, S. Tonellato, C.P. Twigger, J.M. Ponting, S. Duarte, L. Rice, S.A. Goodstadt, R.D. Beaton, E.E. Emes, C. Winter, P. Webber, G. Brandt, M. Nyakatura, F. Adetobi, L. Chiaromonte, P. Eltnitski, R.C. Eswara, M. Hardison, D. Hou, K. Kolbe, W. Makova, A. Miller, C. Nekrutko, S. Riemer, J. Schwartz, S. Taylor, Y. Yang, K. Zhang, T.D. Lindpaintner, M. Andrews, M. Caccamo, L. Clamp, V. Clarke, R. Curwen, E. Durbin, S.M. Eyraas, G.M. Searle, S. Cooper, M. Batzoglu, A. Brudno, E.A. Sidow, B.A. Stone, G. Payseur, C. Bourque, X.S. Lopez-Otin, K. Puentes, S. Chakrabarti, C. Chatterji, L. Dewey, N. Pachter, V.B. Bray, A. Yap, G. Caspi, P.A. Tesler, D. Pevzner, K.M. Haussler, R. Roskin, H. Baertsch, T.S. Clawson, A.S. Furey, D. Hinrichs, W.J. Karolchik,

- K.R. Kent, H. Rosenbloom, M. Trumbower, D.N. Weirauch, P.D. Cooper, B. Stenson, M. Ma, M. Brent, D. Arumugam, R.R. Shteynberg, M.S. Copley, H. Taylor, U. Riethman, J. Mudunuri, M. Peterson, A. Guyer, S. Felsenfeld, S. Old, F. Mockrin, Collins, Genome sequence of the Brown Norway rat yields insights into mammalian evolution, *Nature* 428 (2004) 493–521.
- [63] W.C. Warren, D.F. Clayton, H. Ellegren, A.P. Arnold, L.W. Hillier, A. Kunstner, S. Searle, S. White, A.J. Vilella, S. Fairley, A. Heger, L. Kong, C.P. Ponting, E.D. Jarvis, C.V. Mello, P. Minx, P. Lovell, T.A. Velho, M. Ferris, C.N. Balakrishnan, S. Sinha, C. Blatti, S.E. London, Y. Li, Y.C. Lin, J. George, J. Sweedler, B. Southey, P. Gunaratne, M. Watson, K. Nam, N. Backstrom, L. Smeds, B. Nabholz, Y. Itoh, O. Whitney, A.R. Pfening, J. Howard, M. Volker, B.M. Skinner, D.K. Griffin, L. Ye, W.M. McLaren, P. Flicek, V. Quesada, G. Velasco, C. Lopez-Otin, X.S. Puente, T. Olender, D. Lancet, A.F. Smit, R. Hubley, M.K. Konkel, J.A. Walker, M.A. Batzer, W. Gu, D.D. Pollock, L. Chen, Z. Cheng, E.E. Eichler, J. Stapley, J. Slate, R. Ekblom, T. Birkhead, T. Burke, D. Burt, C. Scharff, I. Adam, H. Richard, M. Sultan, A. Soldatov, H. Lehrach, S.V. Edwards, S.P. Yang, X. Li, T. Graves, L. Fulton, J. Nelson, A. Chinwalla, S. Hou, E.R. Mardis, R.K. Wilson, The genome of a songbird, *Nature* 464 (2010) 757–762.