



Article

Massive Loss of Transcription Factors Promotes the Initial Diversification of Placental Mammals

Xin-Wei Zhao ^{1,2,*}, Jiaqi Wu ^{3,†} , Hirohisa Kishino ^{2,4}  and Ling Chen ¹

¹ Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China

² Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan

³ Department of Molecular Life Science, Tokai University School of Medicine, Isehara 259-1193, Japan

⁴ The Research Institute of Evolutionary Biology, Yoga 2-4-28, Setagaya-ku, Tokyo 158-0098, Japan

* Correspondence: zhaoxinwei90@gmail.com

† These authors contributed equally to this work.

Abstract: As one of the most successful group of organisms, mammals occupy a variety of niches on Earth as a result of macroevolution. Transcription factors (TFs), the fundamental regulators of gene expression, may also have evolved. To examine the relationship between TFs and mammalian macroevolution, we analyzed 140,821 de novo-identified TFs and their birth and death histories from 96 mammalian species. Gene tree vs. species tree reconciliation revealed that placental mammals experienced an upsurge in TF losses around 100 million years ago (Mya) and also near the Cretaceous–Paleogene boundary (K–Pg boundary, 66 Mya). Early Euarchontoglires, Laurasiatheria and marsupials appeared between 100 and 95 Mya and underwent initial diversification. The K–Pg boundary was associated with the massive extinction of dinosaurs, which lead to adaptive radiation of mammals. Surprisingly, TF loss decelerated, rather than accelerated, molecular evolutionary rates of their target genes. As the rate of molecular evolution is affected by the mutation rate, the proportion of neutral mutations and the population size, the decrease in molecular evolution may reflect increased functional constraints to survive target genes.

Keywords: transcription factor; macroevolution; mammal; gene loss; molecular evolution



Citation: Zhao, X.-W.; Wu, J.; Kishino, H.; Chen, L. Massive Loss of Transcription Factors Promotes the Initial Diversification of Placental Mammals. *Int. J. Mol. Sci.* **2022**, *23*, 9720. <https://doi.org/10.3390/ijms23179720>

Academic Editor: Petros Kolovos

Received: 25 July 2022

Accepted: 23 August 2022

Published: 26 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gene expression patterns vary among species—even those which are closely related and share highly similar genomic sequences. These differences are believed to be important factors in organism evolution, being a major source of among species phenotypic variation [1]. Transcription factors (TFs) are sequence-specific DNA-binding trans-regulatory proteins of gene expression that perform an initial step of DNA decoding [2]. Moreover, TFs have many other important functions in eukaryotes [3,4]. As revealed by laboratory experiments, the ability of TFs to drive phenotypic changes has long been known. For example, Homobox (HOX) TF genes play a key role in proper body pattern formation [5], while Sex Determining Region Y (SRY) TF gene, is important for sex determination [6,7]. In addition, the mutations of TF genes have many unexpected consequences, such as the formation of induced pluripotent stem (iPS) cells [8] and cancers [9,10].

Simplification and complication are both critical aspects of macroevolution. Simplification, that is, the reduction in biological complexity, has received less scientific attention than complexity. Examples of simplification-driven biodiversification across the tree of life include early evolutionary histories of metazoans, fungi and eukaryotes [11]. In contrast to nonadaptive simplification, such as drift, which can lead to the accumulation of slightly deleterious mutations in bacteria [12], adaptive genome reduction may also explain some important stages of eukaryotic evolution, such as the simplification of animal metabolism [13]. The ‘less-is-more principle’ suggests that the loss of gene function is a common evolutionary response of populations undergoing an environmental shift and,

consequently, a change in the pattern of selective pressures [14]. In this regard, TF patterns may naturally evolve along with organismal macroevolution.

Available studies show that some of important components of metazoan and embryonic–plant TF kits were present much earlier in their respective single-cell ancestors [15]. Given that the origin and expansion of TFs occurred long before the big bang of speciation, macroevolution has likely been driven by a more direct factor, which is TF loss. For example, the convergent simplification of adaptin complexes in flagellar apparatus diversification caused major diversification in eukaryotes [11]. Another example is functional and phenotypic diversification of the animal mouths by the loss of signaling factors complex of the Wnt family, which evolved 650 million years ago (Mya) in the early multicellular animals [16].

So far, the exact mechanism of TFs, which work during the fast adaptation of mammals to changing environments via macroevolution, is poorly understood. Advances in comparative genomics have clearly shown that the exclusive use of genes as evolutionary units is an oversimplification of real evolutionary relationships [17,18]. Orthologous groups (OGs) refer to sets of genes that have descended from a single ancestral gene in a given ancestral species or taxonomic level. In this research, we use orthologous groups of TFs on the mammal level, rather than TF families to category TFs and construct phylogenetic trees. The OG category approach maintains the evolutionary events in mammalian history and avoids the difficulty of constructing trees for large TF families such as C2H2. By reconciling OG trees of TFs to mammal species tree, the gain and loss events of TFs can be detected. However, for trait values, only current species are available and certain. Here, we use the presence and absence of TFs on terminal branches to detect the possible association with trait divergence. In this study, we show the pattern of TF loss and address the role of TFs in the macroevolutionary process of mammals by describing the correlation between TF loss, target gene (TG) and molecular evolutionary rate, and also between TF loss and species traits.

2. Results

2.1. The Analysis of Evolutionary History of Transcription Factors (TFs) in Mammalians

To analyze the evolutionary history of TF birth and death events, we de novo identified 140,821 TFs from 96 mammalian species [19] and grouped them into 1651 orthologous groups, each having a single origin at the root of mammals (Figure 1a). The history of TF duplication and loss in each orthologous group was estimated by reconciling the gene tree with the species tree (Figure 1b).

The most recent common ancestor of extant mammals is dated back to approximately 170 million years ago (Mya) (Figure 1c). After splitting from Marsupialia, Placentalia diverged into Afrotheria, Xenarthra and Boreoeutheria about 100 Mya. We repeated the analysis by successively assuming Atlantogenata, Exafroplacentalia and Epitheria topologies at the same threshold level. The events detected under the three different topological arrangements were very similar.

Between 100 and 95 Mya, mammals, especially placental mammals underwent initial diversification, the average TF loss rate reached its first peak of 182.4/My (Figure 1c,d). The second peak of TF loss occurred near the Cretaceous–Paleogene (K–Pg) boundary, at 66 Mya. The TF loss rate was as high as 64.5/My within a 4-Mya window of the K–Pg boundary.

As shown in Figure 1d, the most rapid TF gains mainly occurred during the early stages of mammalian macroevolution; this was especially the case for Boreoeutheria, with 1480 TF gain events at corresponding rates of 161.5/My. After this period, TF loss dominated. We detected 1967 TF loss events occurring at a rate of 284.7/My along the ancestral branch of Euarchontoglires. In addition, 2879 TF loss events at a rate of 345.0/My were inferred along the ancestral branch of the Laurasiatherian lineage, while 3819 taking place at a rate of 209.3/My were uncovered on the ancestral branch of marsupials. TF loss also occurred rapidly on subsequent branches, especially where common ancestors of different mammalian orders started to differentiate. For example, the estimated TF loss rates on Chiroptera, Perissodactyla

and Carnivora ancestral branches were 279.6, 115.8 and 313.2/My, respectively. During this time period, the number of species increased significantly. Highly diversified lineages would be expected to contain more highly diversified TFs as well, but we found the opposite to be true. The TF pool became simplified during mammalian species diversification, indicating that TF loss may be related to the early diversification of mammals.

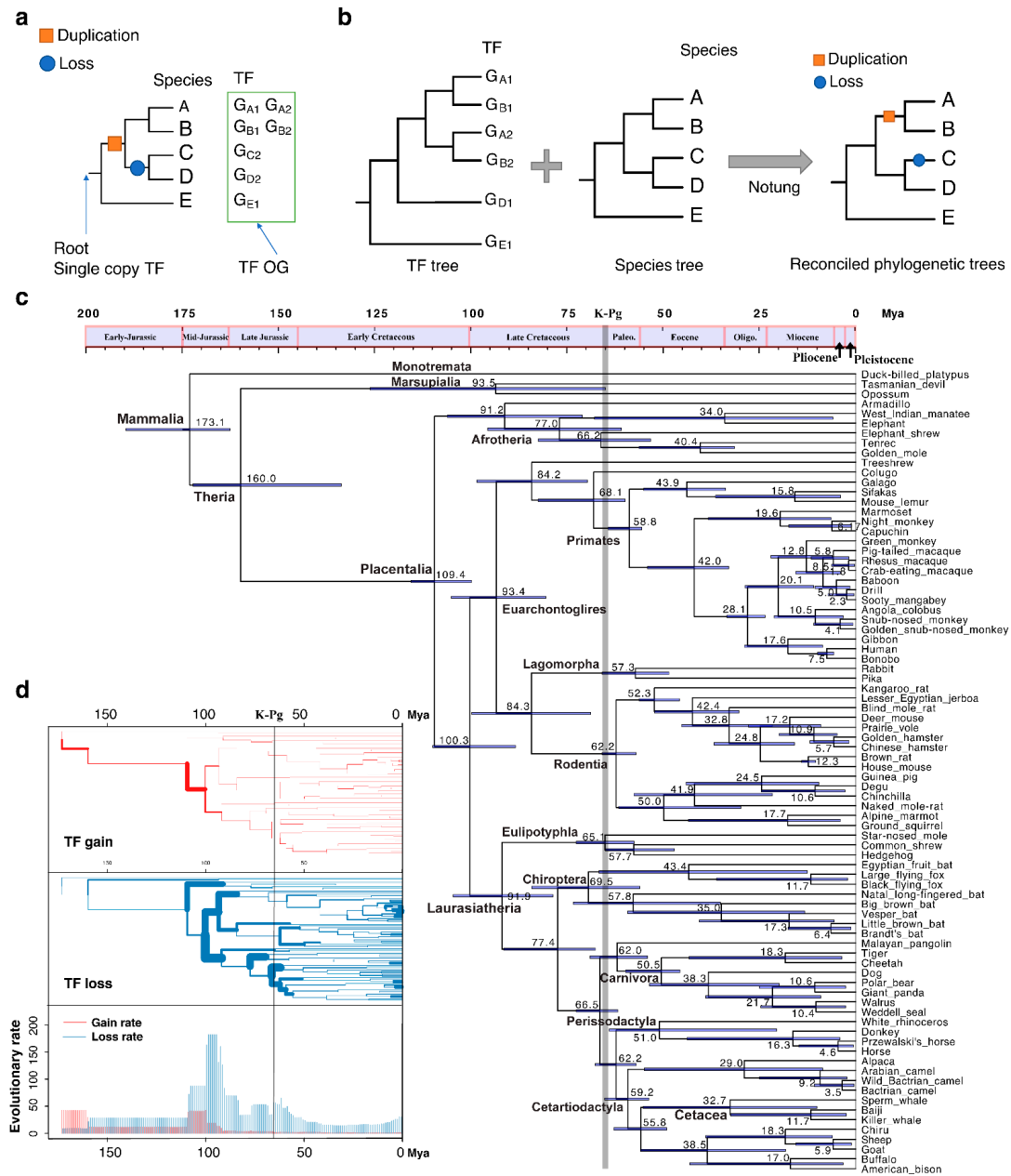


Figure 1. Simplification of transcription factors (TFs) during mammalian evolution using an effective population size of 10^6 . (a) A simple example of TF orthologous group (OG). The orange square indicates a duplication event, and the blue circle represents a loss. The genes in the green box belong to the same OG and all originated from the single-copy gene G at the root. (b) Inference of duplication and loss by TF-tree–species-tree reconciliation. G, TF genes. A–E, species. (c) Time tree of 82 mammals based on protein sequences. Numbers at internal nodes are estimated divergence times (Mya, million years ago), with 95% credible intervals indicated by horizontal bars spanning nodes. (d) Atlas of TF gain and loss during mammalian history. The width of each branch is proportional to the rate of TF gain (or loss). The bar plots are mean evolutionary rate (mean TF gain or loss events of each million years).

For species radiations, the short branches and incomplete lineage sorting poses many challenges in resolving the tree of life (reviewed by [20]). The Nothung software which we used in this study, enabled the estimation of the inconsistency between gene tree and species tree, excluding the incomplete lineage sorting and short branches due to ancestral polymorphism [21]. For this purpose, four levels of effective population size (10^4 , 10^5 , 10^6 and 10^7) were applied as a threshold when considering incomplete lineage sorting and short branches (Figure 2). Although the number of events varied extensively among the four population-sized thresholds, the overall trend was quite similar in regard to TFs. As a representative example of the overall trend, but not exact numbers, the results obtained using an effective population size of 10^6 are shown in Figure 1.

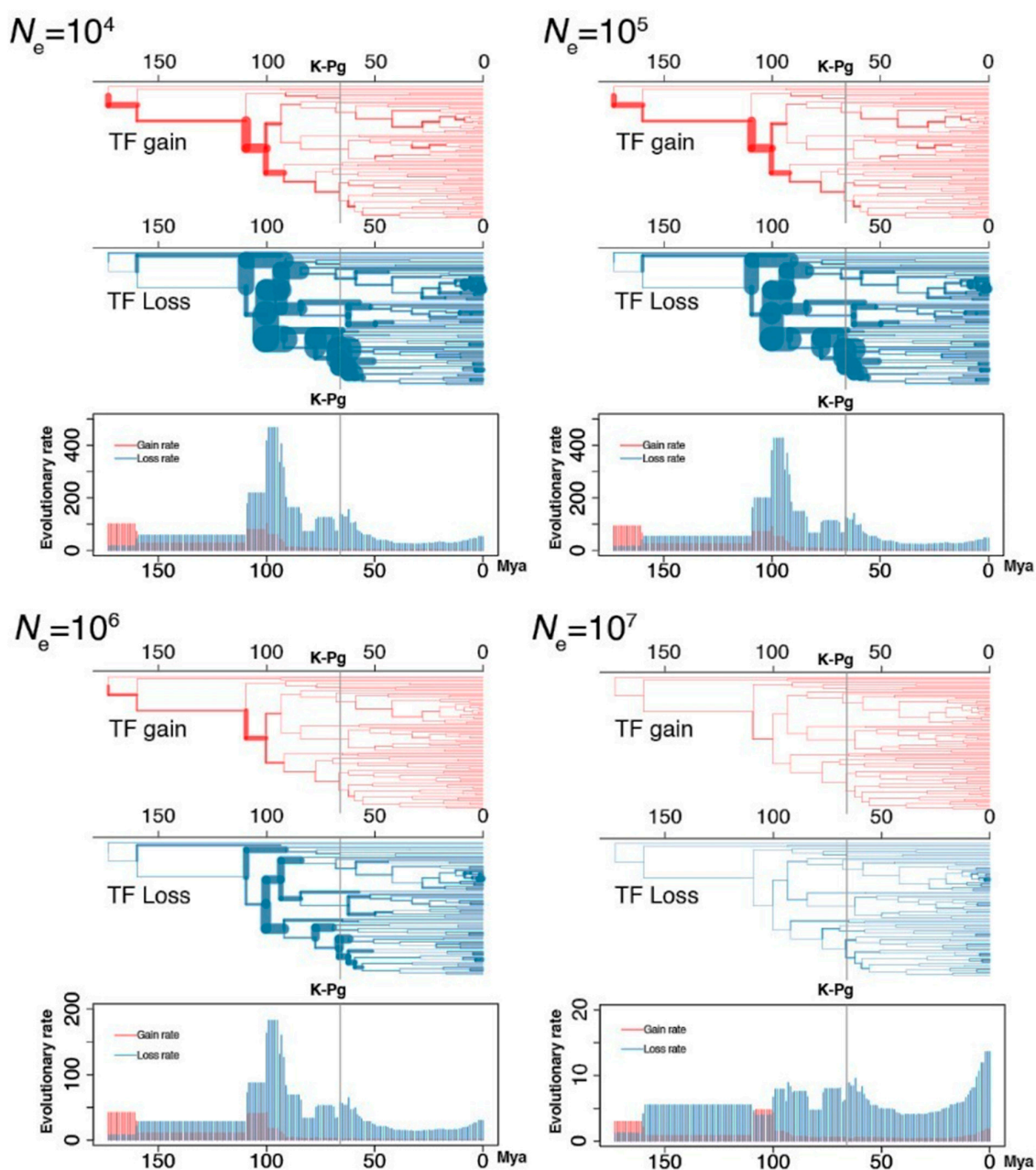


Figure 2. Transcription factor (TF) events inferred using four different effective population sizes (N_e)— 10^4 , 10^5 , 10^6 and 10^7 —as a threshold when considering incomplete lineage sorting and short branches. TF losses and gains are represented by blue and red bars, respectively. The width of each branch is proportional to the rate of TF gain (or loss). The bar plots are mean evolutionary rate (mean TF gain or loss events of each million years).

To avoid the possible variation due to the quality of the mammal root species, we traced 982 ancient mammalian TF orthologous groups, which originated before mammal common ancestor by inclusion an additional 11 outgroups (four birds, four reptiles, two amphibians and one coelacanth) and 82 mammalian species (Figures 3a and S1). From approximately 170 to 100 Mya, the rate of ancient mammal TF gain was slightly higher than the rate of loss, and then after this period the loss of TF started to dominate then loss dominated (Figure 3b).

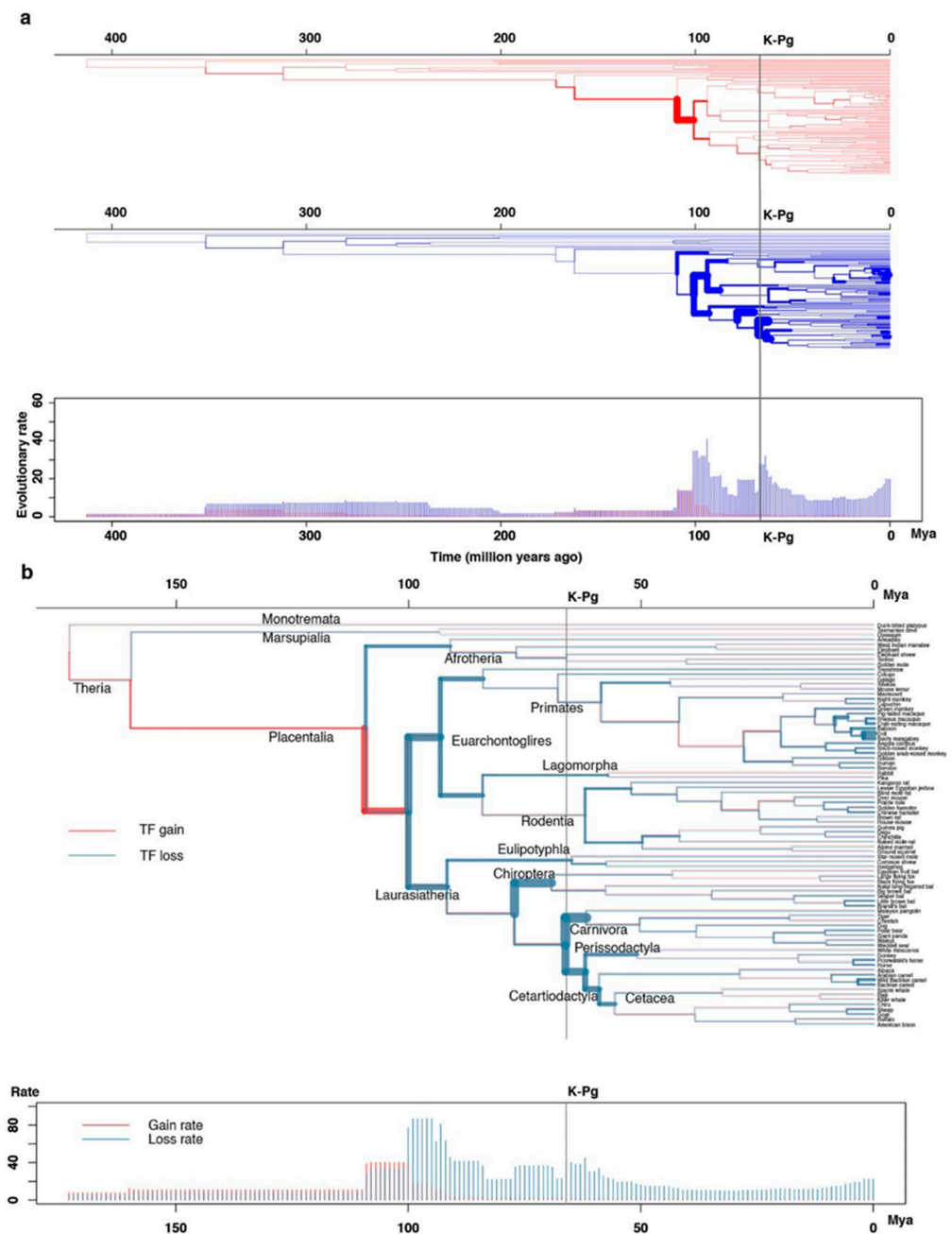


Figure 3. Atlas of ancient transcription factor (TF) gain and loss. (a) TF events in Atlantogenata inferred using an effective population size of 10^6 with outgroup species included. TF losses and gains are represented by blue and red bars, respectively. The width of each branch is proportional to the rate of TF gain (or loss). Divergence times of outgroups (Aves, reptiles, amphibians, etc., see Figure S1) originated from Timetree [22]. (b) Atlas of ancient TF gain and loss of 982 orthologous groups of TFs during mammalian history. The width of each branch is proportional to the rate of TF gain (or loss). Time scale is million years ago (Mya).

2.2. Association with Trait Values

To examine the effect of TF loss on life history traits, we regressed four binary traits—sociality, diurnality, reproductive seasonality and insectivory—on the states (presence/absence) of 1-to-1 orthologous TFs via lasso logistic regression (Figure 4, Table S1).

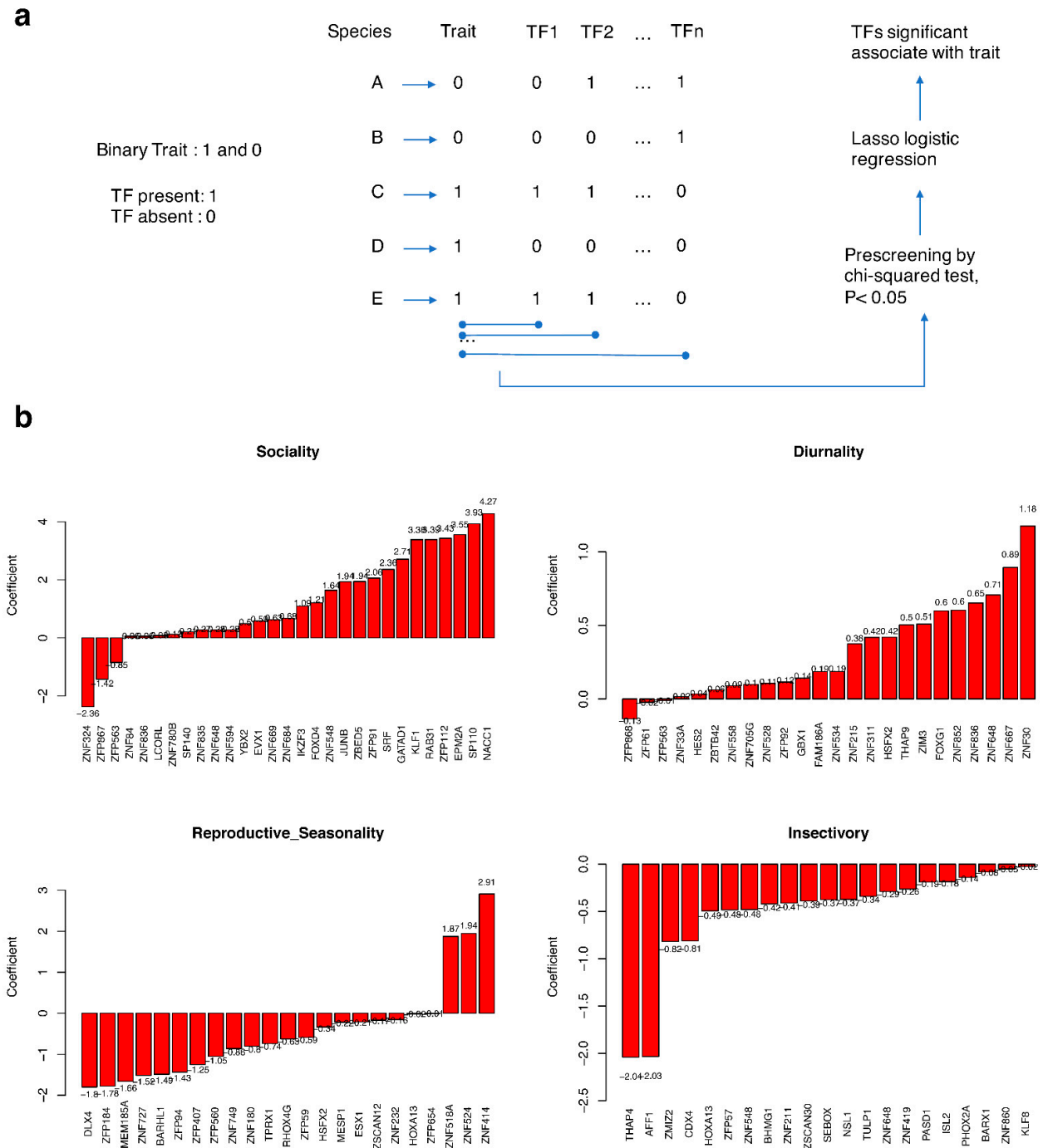


Figure 4. Association of TF loss with mammalian traits. (a) The pipeline used to detect an association between TF presence/absence and life history traits by lasso logistic regression. (b) TFs significantly associated with traits. The x-axis shows TF genes significantly related to the indicated trait, and y-axis values are lasso logistic regression coefficients pre-screened with a chi-square test ($p < 0.05$).

TFs, especially members of the zinc-finger protein with Krüppel-associated box domain (KRAB-ZNF) family, are prominent candidates for a role in mammalian speciation [23]. Consistent with this idea, ZNFs or KRAB-ZNFs comprised half of the TFs significantly associated with the four life history traits (Table S1).

As for sociality, 29 associated TF genes had non-zero correlation coefficient values. This result indicates that the loss of these TFs is possibly linked to the sociality trait. Highly social species have an increased risk of parasitic infection, which necessitates a heavy investment in immune functions [24]. Thus, among genes associated with sociality, half were related to immunity-associated phenotypes (Table S1). Sociality has different effects on shape and behavior, including the relative size of the brain and the prevalence of infanticide (reviewed by [25]). De novo mutations in *NACC1* and disruption in *EPM2A* gene can lead to cerebral atrophy and infantile epilepsy [26,27]. Epilepsy can have significant negative consequences on a child's social development [28]. The ancestor of placental mammals experienced rapid changes in these traits around the K-Pg boundary period [29]. Consequently, TF loss is considered to help the preservation of solitariness in extant solitary mammals that is also reflected in our results (Table 1).

Table 1. TF loss in *NACC1*, *SP110* and *EPM2A* associated with solitary.

Common_Names	Sociality	NACC1	SP110	EPM2A
Cheetah	solitary	0	1	1
Weddell_seal	solitary	0	1	1
Tiger	solitary	0	1	1
Duck-billed_platypus	solitary	1	0	1
Lesser_Egyptian_jerboa	solitary	1	0	1
Opossum	solitary	1	0	1
Galago	solitary	1	0	1
Orangutan	solitary	1	0	1
Tasmanian_devil	solitary	1	0	1
West_Indian_manatee	solitary	1	0	1
Hedgehog	solitary	1	1	0
Pika	solitary	1	1	0
Deer_mouse	solitary	1	1	0
Other 14 species	solitary	1	1	1
Elephant_shrew	social	1	0	1
Bottlenosed_dolphin	social	1	0	1
Other 67 species	social	1	1	1

TF absence and presence are represented by 0 and 1, respectively.

In total, 24 TF genes were significantly associated with diurnality. Starvation and a cold environment have caused the transition from nocturnal to diurnal habits in mammals, and metabolic balance has been identified as a potential common factor affecting circadian rhythm organization [30]. Functions of TFs associated with diurnality shed light on their adaptation. Overexpressed *ZNF667* (*MIPU1*) decreases oxLDL-induced cholesterol accumulation [31]. *ZNF648* is associated with high-density lipoprotein cholesterol levels [32]. Knockout of the *ZFP92* gene increases the amount of total body fat [33]. *GBX1*-knockout mice exhibit reduced thermosensory functions [34]. Changes in the metabolic balance of fat and thermosensory function might affect changes in tolerance to starvation and cold that likely promoted the adaptation to a nocturnal lifestyle.

The type of food ingested by mammals is also evolutionarily very important. Among genes associated with the insectivory trait, the knockout of *SEBOX*, *PHOX2A* and *KLF8* genes in mice has been found to increase the levels of circulating total protein or improve glucose tolerance [33,35]. Knockout of *NSL1* and *ZNF648* leads to increased circulating cholesterol levels [32,33]. According to various knockout data [33,36], *AFF1*, *NSL1* and *TULP1* are related to vision. Thus, TF loss likely played an important role in the divergence between insectivory (poor eyesight) and non-insectivorous (good eyesight) mammals.

In terms of reproductive seasonality, 23 TFs are significantly associated with this trait. More than half of the reproductive seasonality-associated TFs were ZNFs. KRAB-ZNFs play a major role in the recognition or transcriptional silencing of transposable elements, and transposon-mediated rewiring of gene regulatory networks that contributed to the evolution of pregnancy in mammals [37]. With regard to the phenotype associated with these TF genes (Table S1), 13 out of 23 are related to fetal viability or formation. Seasonally breeding mammals with these TF losses most likely have better survival chances for fit to specific niches.

2.3. The Effect of TF Loss on Its Current Surviving Target Genes

To quantify the effect of TF loss on transcription factor's target genes (TF-TGs), we examined TG evolutionary rates based on 9396 TF-TG interactions by literature curations from the TTRUST2 database [38], and 1,342,129 TF-TG interactions based on Chip-seq data from the hTFtarget database [39] (Figure 5). Figure 5b,c, which compare the evolutionary rates of 1-to-1 orthologous TGs with TFs and the evolutionary rates of TGs without TFs, clearly show the decelerating effect of TF loss on TGs over the long term (paired *t*-test, *p*-value = 0.017 for TF-TG interactions by literature curations; *p*-value < 2.2×10^{-16} for TF-TG interactions based on Chip-seq data). If it is the decelerate evolutionary rate of TG which drives the TF loss, we will see many low-rate TGs with TFs, or many points on the left side of Figure 5b,c, but they do not show this pattern. Because the rate of molecular evolution of a gene is negatively correlated with the strength of a functional constraint [40], the decelerated gene tends to be conserved. As the target genes covers more than fifteen thousand genes, it suggests that currently surviving genes have been "protected" by TF loss over a long time period, especially during macroevolution.

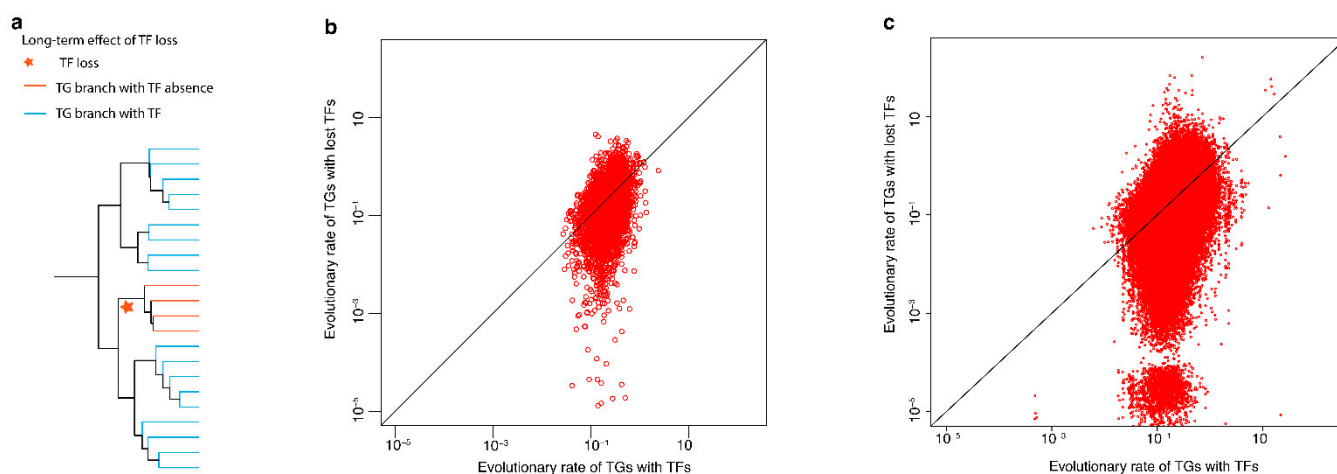


Figure 5. Deceleratory effect of the loss of transcription factors (TFs) on molecular evolutionary rates of their target genes (TGs). With decreased molecular evolutionary rate, current genes survive till present. (a) Comparison of the evolutionary rates of TGs in species in the subtree of the TF event (colored in red) with the rates of TGs in other species. (b,c) Long-term effect of TF loss on the evolutionary rate of TGs, TF-TG interactions by (b) literature curations or (c) based on Chip-seq data. *x*-axis: evolutionary rate of TGs still controlled by TFs; *y*-axis: evolutionary rate of TGs with lost TFs. TF-TG interactions by literature curations from TTRUST2 database; TF-TG interactions based on Chip-seq data from the hTFtarget database.

In this study, we have uncovered three main findings. First, the number of TFs increased greatly during the early stages of mammalian species formation; later, however, TF losses predominated during the course of macroevolution. Second, TFs significantly associated with traits change. Third, in the face of serious environmental changes, such as the K-Pg boundary, the loss of such TFs rewired the regulatory network, set functional

constraints on TGs and facilitated organismal survival. For TFs, duplication provides possibility, but loss produces inter-species differentiation.

3. Discussion

Our study showed that the main branches of mammals, especially placental mammals, Afrotheria, Xenarthra and Boreoeutheria experienced a large number of TF loss events at approximately 100 Mya, which corresponds to the first peak of TF loss in mammalian evolutionary history (Figure 1). The likely place of origin or current extant range of Afrotheria, Xenarthra and Boreoeutheria clades is Africa, South America and the northern supercontinent of Laurasia, respectively [41,42]. This distribution indicates that the living environment of these organisms has likely changed compared with their common ancestor; in other words, their niches and eco-traits may have changed. In contrast, a significant correlation exists between traits and the loss of some of the TFs (Figure 4). This correlation suggests that TF loss may be beneficial to the survival of species undergoing niche and environmental changes and can therefore become fixed in the survivors. At the second peak of TF loss, which took place near the K–Pg boundary (66 Mya), clusters of loss events are especially apparent on three branches leading to Carnivora, Perissodactyla and Cetartiodactyla. Members of Carnivora are basically carnivores, while Perissodactyla mainly comprises herbivores and Cetartiodactyla contains both herbivores and omnivores. The feeding preferences of the three taxa have thus clearly and rapidly diverged from those of their common ancestor, and the massive loss of TFs in this transition may also have been beneficial to species survival. This phenomenon suggests that TF loss may play an important role in species adaptation to environmental change and macroevolution.

The loss of genes is generally thought to lead to non-functionalization [43]. These genes are usually considered to be redundant parts that can be replaced by other genes or provide materials for evolution. While TF loss leads to different results. A positive correlation exists between the loss of TFs and mammalian ancestral traits (Figure 4). Losses of TFs decelerate the molecular evolution of TGs (Figure 5). The loss of TFs in mammals enhances functional constraints on their TGs. TF loss may keep target genes with fitness privileged.

Mutations tend to have functional loss than gain of function [44]. For TFs, this may explain why TF loss can be a rapid response to change in the environment rather than TF gain. Research on bacterial populations under different conditions has also shown that adaptive loss-of-function by the mutations of enzymatic and regulatory functions has an important role in their adaptation to new environments [45]. Possibly the “Ratchet effect” on regulatory system redundancy to improve fitness during environmental challenges. According to our previous research, TFs are located on a more peripheral state on the gene interaction network than other functional genes; thus, it is more dispensable and milder for species to endure changes [19]. This may provide clues on why TF loss takes place prior to other functional gene variations. Facing environmental changes, TF loss may increase fitness by adjusting more privileged traits directly and preserve fitness genes by functional constraints.

In this study, although TF loss is generally predicted to be one of the factors that may affect its target genes, macroevolution and traits, the opposite direction is also possible. However, if the loss of TF is driven by the low evolutionary rates of its target genes, differentiation and trait changes, it could also demonstrate that the loss of TFs could privilege species that survive these changes.

Many TFs have been lost during mammalian history, especially around the K–Pg boundary or afterwards. The massive extinction of dinosaurs as predators may have reduced functional constraints on relevant transcription machineries in our ancestors. As a result, TFs were lost unless new species shifted their habitats into recently opened niches and developed new lifestyles adapted to the new environments. However, TF loss has enhanced functional constraints on any TF TGs surviving to the present day. Mammals that lost these TFs may have increased their survival rate by retaining the ancestral traits in an unchanged niche. When an old niche was occupied, a less-adapted species with a

less-altered or insufficiently altered transcriptional machinery may have needed to take risks to adapt to a new niche and change traits for survival. To adapt to new environment, a new set of TF loss occurred. With different TF loss patterns, species in old niches and in new niches owned different traits. Their phenotypes became different from their ancestor or each other in a relatively short period of time. During macroevolution periods, ancestor species in different niches may keep different TF profiles and gene profiles due to TF losses, thus promoting biodiversity.

4. Materials and Methods

4.1. Mammalian Transcription Factors (TFs)

Based on our previous research, we obtained 140,821 TF proteins, nearly all of which were specific for mammals [19]. To assess the accuracy of TF annotations, we compared our list of 1625 human TFs (extracted from the 140,821 TFs in 96 mammalian genomes shown in Data Availability) with two well-known TF databases, AnimalTFDB3 [45] and HumanTFs [4] (Figure S2). The number of human TFs in these two databases is similar to that of our list: 1639 in humanTFs and 1665 in AnimalTFDB3. A total of 1402 TFs are listed in all three databases, while 82 are unique to HumanTFs, 123 are only found in AnimalTFDB3, and 140 are restricted to our human TF list. All three databases are based on similar DBD and HMMER [46] pipelines.

AnimalTFDB3 contains 125,135 TFs from 97 genomes ranging from *Caenorhabditis elegans* to mammalian species such as humans, whereas our database contains 140,821 TFs and focuses on 96 mammalian species. The latter database may thus provide better insights into the evolutionary history of TFs in mammals. To avoid the limitations of mammalian species, we further annotated our database with information on orthologous groups from OrthoDB [47]. This way, we can trace TFs in mammalian species back to those of bacterial species.

To further confirm events occurring in the common ancestor of mammals, we conducted a hidden Markov model (HMM) search on eleven outgroup species (four birds, four reptiles, two amphibians and one coelacanth) and collected all genes predicted to encode TFs.

4.2. Multiple Alignment of Mammalian Orthologs

We downloaded 96 complete mammalian genomes from GenBank [48] and used a custom Perl script to extract protein-coding sequences of each species. A gene pool of 21,350 mammalian genes was constructed based on NCBI genomic annotations. Using the results of the HMM search and annotations in OrthoDB, we assigned TFs to 2880 mammalian orthologous groups. We confirmed the homology of outgroup TFs and mammal-TFs by a local blast search (blastp, [49]) and assigned outgroup TFs to mammalian orthologous groups.

We generated a multi-sequence file of all genes in the gene pool for each of the orthologous groups. Codon-level alignments were performed in Prank v.170427 [50]. Sites with less than 70% coverage across all species, as well as sequences with less than 30% coverage among gene loci, were removed from the alignment.

4.3. Inference of Gene Trees

We estimated the maximum likelihood tree for each gene using the IQ-TREE software [51], which automatically performed model selection and determined the best data partitions. The best evolutionary model for each gene was independently selected based on the Bayesian information criterion and used for the inference of the nucleotide tree. All gene trees were calculated using 1000 bootstrap replicates.

4.4. Species Tree Inference and Divergence Time Estimation

We selected 823 genes that were single copies in mammals according to Wu et al. [29] and present in all 96 mammals. The 823 genes were used to infer a species tree using the

coalescent-based Njst method [52]. The topology of the inferred species tree was consistent with that of Tarver et al., 2016 [53], who placed treeshrew (*Tupaia chinensis*) as the root lineage of Glires. The phylogenetic position of treeshrew is not yet resolved; however, several researchers consider treeshrew to be the root lineage of Euarchonta not Glires [54]. As an alternative species tree in this study, we consequently used a tree in which the position of treeshrew was fixed at the root of Euarchonta rather than Glires. Following the same method as Wu et al., 2017 [29], we estimated the divergence times of 96 mammals based on the inferred branch effect (genomic rate \times genomic time) and fossil calibrations.

4.5. Reconstructing the History of TF Duplication and Loss

A history of gene duplication and loss was estimated by reconciling gene trees with the species tree and plotted by R. To focus on these events during mammalian evolution, TFs were categorized into orthologous groups (OrthoDB, <https://www.orthodb.org/>, accessed on 18 December 2019), each having a single ancestor at the root of mammals. Out of 140,821 TFs from 96 mammalian species, OrthoDB included nearly 120,000 TFs from 82 species. These TFs were assigned to 2880 orthologous groups. A phylogenetic tree was constructed for each of the 1651 orthologous groups containing more than three members. These 1651 OGs cover over 97% TFs of 82 mammal species. To infer the history of TF duplication and loss, each of the 1651 phylogenetic trees was reconciled with the species tree according to the criterion of parsimony using NOTUNG 2.9 [21]. To take into account inconsistency due to incomplete lineage sorting, we established a branch length threshold in terms of $N\mu$. Branches with lengths under the threshold were treated as weak branches and rearranged by NOTUNG. For each orthologous group tree, the value of μ corresponded to the average path length from root to terminals divided by the calculated age of the root. We assumed an average generation length of 10 years and considered four possible values for effective population size, namely, 10^4 , 10^5 , 10^6 and 10^7 , and three tree topologies (Figure S3). The events of the orthologous groups were aggregated to obtain the numbers of births and losses of TFs along each branch of the species tree. In this paper, we present the results of the analysis based on an effective population size of 10^6 . Other effective population size threshold settings gave similar results. The platypus genome [55] is not updated on the current version of OrthoDB. To confirm that the inclusion of this species did not lead to biased results, we added 11 more outgroups with 82 mammalian species, including 1 fish, 2 amphibians, 4 reptiles and 4 birds, for use in tree construction. Using the 11 outgroups, 982 TF orthologous groups were traced and analyzed by the above-mentioned method. Divergence times of the outgroups were taken from the Timetree [22] database.

4.6. The Effect of TF Loss on TG Molecular Evolutionary Rates

To examine the long-term effects of TF loss, we compared the mean rates along the terminal branches of a TG tree between species that possess the TF and species that do not. Rates of molecular evolution were estimated as the ratios of branch lengths to the relevant time periods in the species tree. In total, 9396 TF-TG interactions by literature curations were downloaded from the TTRUST2 database [38] and 1,342,129 TF-TG interactions based on Chip-seq data were downloaded from the hTFtarget database [39].

4.7. Association Study of Ecotype Traits and TF Presence/Absence

We manually collected the ecotype traits of 96 mammalian species from the Animal Diversity Web (<http://animaldiversity.org/>, accessed on 15 August 2019). The presence and absence of TF orthologs in each species were recorded as 1 and 0, respectively, based on mammal-terminal single-copy orthologs. The ecotypes were binary recorded as 1 and 0. We identified significant associations using the lasso logistic regression procedure as implemented in the glmnet package [56] in R.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23179720/s1> [32,57–110].

Author Contributions: X.-W.Z., J.W. and H.K. designed the research. X.-W.Z., J.W. and H.K. developed the methods and carried out the analysis. X.-W.Z., J.W., H.K. and L.C. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a Grant-in-Aid for Scientific Research (B) (19H04070) and a Grant-in-Aid for JSPS Research Fellows (18F18385) from the Japan Society for the Promotion of Science, a University of Tokyo Grant for PhD Research, and funding from the China Scholarship Council (CSC).

Data Availability Statement: TF data, program codes, phylogenetic trees and other data related to our research: (<https://drive.google.com/drive/folders/19jvuchCudXDZyrcJza4CpsNf22enYkHd?usp=sharing>) (accessed on 18 August 2022).

Acknowledgments: We express our deep gratitude to Ze Zhang and Wei Sun from Chongqing University and Yonezawa Takahiro from Tokyo University of Agriculture for their valuable comments and suggestions on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Erwin, D.H.; Davidson, E.H. The evolution of hierarchical gene regulatory networks. *Nat. Rev. Genet.* **2009**, *10*, 141–148. [[CrossRef](#)] [[PubMed](#)]
2. Cramer, P. Organization and regulation of gene transcription. *Nature* **2019**, *573*, 45–54. [[CrossRef](#)] [[PubMed](#)]
3. Spitz, F.; Furlong, E.E.M. Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.* **2012**, *13*, 613–626. [[CrossRef](#)] [[PubMed](#)]
4. Lambert, S.A.; Jolma, A.; Campitelli, L.F.; Das, P.K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T.R.; Weirauch, M.T. The human transcription factors. *Cell* **2018**, *175*, 598–599. [[CrossRef](#)]
5. Krumlauf, R. Hox Genes in Vertebrate Development. *Cell* **1994**, *78*, 191–201. [[CrossRef](#)]
6. Jager, R.J.; Anvret, M.; Hall, K.; Scherer, G. A Human Xy Female with a Frame Shift Mutation in the Candidate Testis-Determining Gene Sry. *Nature* **1990**, *348*, 452–454. [[CrossRef](#)]
7. Koopman, P.; Gubbay, J.; Vivian, N.; Goodfellow, P.; Lovellbadge, R. Male Development of Chromosomally Female Mice Transgenic for Sry. *Nature* **1991**, *351*, 117–121. [[CrossRef](#)]
8. Takahashi, K.; Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126*, 663–676. [[CrossRef](#)]
9. Batlle, E.; Sancho, E.; Franci, C.; Dominguez, D.; Monfar, M.; Baulida, J.; de Herreros, A.G. The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* **2000**, *2*, 84–89. [[CrossRef](#)]
10. Tomlins, S.A.; Rhodes, D.R.; Perner, S.; Dhanasekaran, S.M.; Mehra, R.; Sun, X.W.; Varambally, S.; Cao, X.H.; Tchinda, J.; Kuefer, R.; et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **2005**, *310*, 644–648. [[CrossRef](#)]
11. O'Malley, M.A.; Wideman, J.G.; Ruiz-Trillo, I. Losing Complexity: The Role of Simplification in Macroevolution. *Trends Ecol. Evol.* **2016**, *31*, 608–621. [[CrossRef](#)] [[PubMed](#)]
12. Kuo, C.H.; Ochman, H. Deletional bias across the three domains of life. *Genome Biol. Evol.* **2009**, *1*, 145–152. [[CrossRef](#)] [[PubMed](#)]
13. Zmasek, C.M.; Godzik, A. Strong functional patterns in the evolution of eukaryotic genomes revealed by the reconstruction of ancestral protein domain repertoires. *Genome Biol.* **2011**, *12*, R4. [[CrossRef](#)] [[PubMed](#)]
14. Olson, M.V. When less is more: Gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* **1999**, *64*, 18–23. [[CrossRef](#)]
15. de Mendoza, A.; Sebe-Pedros, A. Origin and evolution of eukaryotic transcription factors. *Curr. Opin. Genet. Dev.* **2019**, *58–59*, 25–32. [[CrossRef](#)]
16. Somorjai, I.M.L.; Marti-Solans, J.; Diaz-Gracia, M.; Nishida, H.; Imai, K.S.; Escrava, H.; Canestro, C.; Albalat, R. Wnt evolution and function shuffling in liberal and conservative chordate genomes. *Genome Biol.* **2018**, *19*, 98. [[CrossRef](#)]
17. Gabaldon, T.; Koonin, E.V. Functional and evolutionary implications of gene orthology. *Nat. Rev. Genet.* **2013**, *14*, 360–366. [[CrossRef](#)]
18. Tatusov, R.L.; Koonin, E.V.; Lipman, D.J. A genomic perspective on protein families. *Science* **1997**, *278*, 631–637. [[CrossRef](#)]
19. Zhao, X.-W.; Kishino, H. Multiple Isolated Transcription Factors Act as Switches and Contribute to Species Uniqueness. *Genes* **2020**, *11*, 1148. [[CrossRef](#)]
20. Kapli, P.; Yang, Z.; Telford, M.J. Phylogenetic tree building in the genomic age. *Nat. Rev. Genet.* **2020**, *21*, 428–444. [[CrossRef](#)]
21. Chen, K.; Durand, D.; Farach-Colton, M. NOTUNG: A program for dating gene duplications and optimizing gene family trees. *J. Comput. Biol.* **2000**, *7*, 429–447. [[CrossRef](#)] [[PubMed](#)]
22. Kumar, S.; Stecher, G.; Suleski, M.; Hedges, S.B. TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **2017**, *34*, 1812–1819. [[CrossRef](#)] [[PubMed](#)]
23. Nowick, K.; Carneiro, M.; Faria, R. A prominent role of KRAB-ZNF transcription factors in mammalian speciation? *Trends Genet.* **2013**, *29*, 130–139. [[CrossRef](#)] [[PubMed](#)]

24. Moller, A.P.; Merino, S.; Brown, C.R.; Robertson, R.J. Immune defense and host sociality: A comparative study of swallows and martins. *Am. Nat.* **2001**, *158*, 136–145. [[CrossRef](#)] [[PubMed](#)]
25. Silk, J.B. The adaptive value of sociality in mammalian groups. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2007**, *362*, 539–559. [[CrossRef](#)]
26. Schoch, K.; Meng, L.; Szelinger, S.; Bearden, D.R.; Stray-Pedersen, A.; Busk, O.L.; Stong, N.; Liston, E.; Cohn, R.D.; Scaglia, F.; et al. A Recurrent De Novo Variant in NACC1 Causes a Syndrome Characterized by Infantile Epilepsy, Cataracts, and Profound Developmental Delay. *Am. J. Hum. Genet.* **2017**, *100*, 343–351. [[CrossRef](#)]
27. Ganesh, S.; Delgado-Escueta, A.V.; Sakamoto, T.; Avila, M.R.; Machado-Salas, J.; Hoshii, Y.; Akagi, T.; Gomi, H.; Suzuki, T.; Amano, K.; et al. Targeted disruption of the Epm2a gene causes formation of Lafora inclusion bodies, neurodegeneration, ataxia, myoclonus epilepsy and impaired behavioral response in mice. *Hum. Mol. Genet.* **2002**, *11*, 1251–1262. [[CrossRef](#)]
28. Gilliam, F. Social Cognition and Epilepsy: Understanding the Neurobiology of Empathy and Emotion. *Epilepsy Curr.* **2015**, *15*, 118–119. [[CrossRef](#)]
29. Wu, J.; Yonezawa, T.; Kishino, H. Rates of Molecular Evolution Suggest Natural History of Life History Traits and a Post-K-Pg Nocturnal Bottleneck of Placentals. *Curr. Biol.* **2017**, *27*, 3025–3033. [[CrossRef](#)]
30. van der Vinne, V.; Riede, S.J.; Gorter, J.A.; Eijer, W.G.; Sellix, M.T.; Menaker, M.; Daan, S.; Pilorz, V.; Hut, R.A. Cold and hunger induce diurnality in a nocturnal mammal. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15256–15260. [[CrossRef](#)]
31. Leib, A.; Cieza, A.; Tschiesner, U. Perspective of physicians within a multidisciplinary team: Content validation of the comprehensive ICF core set for head and neck cancer. *Head Neck* **2012**, *34*, 956–966. [[CrossRef](#)] [[PubMed](#)]
32. Beaney, K.E.; Cooper, J.A.; McLachlan, S.; Wannamethee, S.G.; Jefferis, B.J.; Whincup, P.; Ben-Shlomo, Y.; Price, J.F.; Kumari, M.; Wong, A.; et al. Variant rs10911021 that associates with coronary heart disease in type 2 diabetes, is associated with lower concentrations of circulating HDL cholesterol and large HDL particles but not with amino acids. *Cardiovasc. Diabetol.* **2016**, *15*, 115. [[CrossRef](#)]
33. Smith, C.L.; Blake, J.A.; Kadin, J.A.; Richardson, J.E.; Bult, C.J.; Mouse Genome Database, G. Mouse Genome Database (MGD)-2018: Knowledgebase for the laboratory mouse. *Nucleic Acids Res.* **2018**, *46*, D836–D842. [[CrossRef](#)] [[PubMed](#)]
34. Meziane, H.; Fraulob, V.; Riet, F.; Krezel, W.; Selloum, M.; Geffarth, M.; Acampora, D.; Herault, Y.; Simeone, A.; Brand, M.; et al. The homeodomain factor Gbx1 is required for locomotion and cell specification in the dorsal spinal cord. *PeerJ* **2013**, *1*, e142. [[CrossRef](#)] [[PubMed](#)]
35. Clement, S.; Krause, U.; Desmedt, F.; Tanti, J.F.; Behrends, J.; Pesesse, X.; Sasaki, T.; Penninger, J.; Doherty, M.; Malaisse, W.; et al. The lipid phosphatase SHIP2 controls insulin sensitivity. *Nature* **2001**, *409*, 92–97. [[CrossRef](#)]
36. Isaacs, A.M.; Oliver, P.L.; Jones, E.L.; Jeans, A.; Potter, A.; Hovik, B.H.; Nolan, P.M.; Vizor, L.; Glenister, P.; Simon, A.K.; et al. A mutation in Af4 is predicted to cause cerebellar ataxia and cataracts in the robotic mouse. *J. Neurosci.* **2003**, *23*, 1631–1637. [[CrossRef](#)]
37. Lynch, M. The evolution of genetic networks by non-adaptive processes. *Nat. Rev. Genet.* **2007**, *8*, 803–813. [[CrossRef](#)]
38. Han, H.; Cho, J.W.; Lee, S.; Yun, A.; Kim, H.; Bae, D.; Yang, S.; Kim, C.Y.; Lee, M.; Kim, E.; et al. TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **2018**, *46*, D380–D386. [[CrossRef](#)]
39. Zhang, Q.; Liu, W.; Zhang, H.-M.; Xie, G.-Y.; Miao, Y.-R.; Xia, M.; Guo, A.-Y. hTFtarget: A comprehensive database for regulations of human transcription factors and their targets. *Genom. Proteom. Bioinform.* **2020**, *18*, 120–128. [[CrossRef](#)]
40. Kimura, M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* **1977**, *267*, 275–276. [[CrossRef](#)]
41. Waddell, P.J.; Okada, N.; Hasegawa, M. Towards resolving the interordinal relationships of placental mammals. *Syst. Biol.* **1999**, *48*, 1–5. [[CrossRef](#)] [[PubMed](#)]
42. O’Leary, M.A.; Bloch, J.I.; Flynn, J.J.; Gaudin, T.J.; Giallombardo, A.; Giannini, N.P.; Goldberg, S.L.; Kraatz, B.P.; Luo, Z.X.; Meng, J.; et al. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science* **2013**, *339*, 662–667. [[CrossRef](#)] [[PubMed](#)]
43. Albalat, R.; Canestro, C. Evolution by gene loss. *Nat. Rev. Genet.* **2016**, *17*, 379–391. [[CrossRef](#)]
44. Hottes, A.K.; Freddolino, P.L.; Khare, A.; Donnell, Z.N.; Liu, J.C.; Tavazoie, S. Bacterial adaptation through loss of function. *PLoS Genet.* **2013**, *9*, e1003617. [[CrossRef](#)] [[PubMed](#)]
45. Hu, H.; Miao, Y.R.; Jia, L.H.; Yu, Q.Y.; Zhang, Q.; Guo, A.Y. AnimalTFDB 3.0: A comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res.* **2018**, *47*, D33–D38. [[CrossRef](#)]
46. Eddy, S.R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **2011**, *7*, e1002195. [[CrossRef](#)]
47. Kriventseva, E.V.; Kuznetsov, D.; Tegenfeldt, F.; Manni, M.; Dias, R.; Simao, F.A.; Zdobnov, E.M. OrthoDB v10: Sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* **2019**, *47*, D807–D811. [[CrossRef](#)]
48. Benson, D.A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2012**, *41*, D36–D42. [[CrossRef](#)]
49. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *15*, 421. [[CrossRef](#)]
50. Löytynoja, A. Phylogeny-aware alignment with PRANK. In *Multiple Sequence Alignment Methods*; Humana Press: Totowa, NJ, USA, 2014; pp. 155–170.

51. Nguyen, L.-T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)]
52. Liu, L.; Yu, L. Estimating species trees from unrooted gene trees. *Syst. Biol.* **2011**, *60*, 661–667. [[CrossRef](#)] [[PubMed](#)]
53. Tarver, J.E.; Dos Reis, M.; Mirarab, S.; Moran, R.J.; Parker, S.; O'Reilly, J.E.; King, B.L.; O'Connell, M.J.; Asher, R.J.; Warnow, T. The interrelationships of placental mammals and the limits of phylogenetic inference. *Genome Biol. Evol.* **2016**, *8*, 330–344. [[CrossRef](#)] [[PubMed](#)]
54. Mason, V.C.; Li, G.; Minx, P.; Schmitz, J.; Churakov, G.; Doronina, L.; Melin, A.D.; Dominy, N.J.; Lim, N.T.; Springer, M.S. Genomic analysis reveals hidden biodiversity within colugos, the sister group to primates. *Sci. Adv.* **2016**, *2*, e1600633. [[CrossRef](#)] [[PubMed](#)]
55. Zhou, Y.; Shearwin-Whyatt, L.; Li, J.; Song, Z.; Hayakawa, T.; Stevens, D.; Fenelon, J.C.; Peel, E.; Cheng, Y.; Pajpach, F. Platypus and echidna genomes reveal mammalian biology and evolution. *Nature* **2021**, *592*, 756–762. [[CrossRef](#)]
56. Friedman, J.; Hastie, T.; Tibshirani, R.; Narasimhan, B.; Tay, K.; Simon, N.; Qian, J. Package 'Glmnet'. CRAN R Repository. 2021. Available online: <https://cran.r-project.org/web/packages/glmnet/index.html> (accessed on 1 July 2022).
57. Melo, J.B.; Estevinho, A.; Saraiva, J.; Ramos, L.; Carreira, I.M. Cutis Aplasia as a clinical hallmark for the syndrome associated with 19q13.11 deletion: The possible role for UBA2 gene. *Mol. Cytogenet.* **2015**, *8*, 21. [[CrossRef](#)]
58. Simpson, J.C.; Joggerst, B.; Laketa, V.; Verissimo, F.; Cetin, C.; Erfle, H.; Bexiga, M.G.; Singan, V.R.; Heriche, J.K.; Neumann, B.; et al. Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nat. Cell Biol.* **2012**, *14*, 764–774. [[CrossRef](#)] [[PubMed](#)]
59. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T.I.; Nudel, R.; Lieder, I.; Mazor, Y.; et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinform.* **2016**, *54*, 1.30.1–1.30.33. [[CrossRef](#)]
60. Boggio, E.M.; Pancrazi, L.; Gennaro, M.; Lo Rizzo, C.; Mari, F.; Meloni, I.; Ariani, F.; Panighini, A.; Novelli, E.; Biagioni, M.; et al. Visual impairment in FOXP1-mutated individuals and mice. *Neuroscience* **2016**, *324*, 496–508. [[CrossRef](#)]
61. Hurov, K.E.; Cotta-Ramusino, C.; Elledge, S.J. A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. *Genes Dev.* **2010**, *24*, 1939–1950. [[CrossRef](#)]
62. Majumdar, S.; Singh, A.; Rio, D.C. The human THAP9 gene encodes an active P-element DNA transposase. *Science* **2013**, *339*, 446–448. [[CrossRef](#)]
63. Alders, M.; Ryan, A.; Hodges, M.; Blik, J.; Feinberg, A.P.; Privitera, O.; Westerveld, A.; Little, P.F.; Mannens, M. Disruption of a novel imprinted zinc-finger gene, ZNF215, in Beckwith-Wiedemann syndrome. *Am. J. Hum. Genet.* **2000**, *66*, 1473–1484. [[CrossRef](#)] [[PubMed](#)]
64. Yasukochi, Y.; Sakuma, J.; Takeuchi, I.; Kato, K.; Oguri, M.; Fujimaki, T.; Horibe, H.; Yamada, Y. Six novel susceptibility loci for coronary artery disease and cerebral infarction identified by longitudinal exome-wide association studies in a Japanese population. *Biomed. Rep.* **2018**, *9*, 123–134. [[CrossRef](#)] [[PubMed](#)]
65. Liu, L.; Bu, H.; Yang, Y.; Tan, Z.; Zhang, F.; Hu, S.; Zhao, T. A Targeted, Next-Generation Genetic Sequencing Study on Tetralogy of Fallot, Combined With Cleft Lip and Palate. *J. Craniofac. Surg.* **2017**, *28*, e351–e355. [[CrossRef](#)] [[PubMed](#)]
66. Patel, N.; Smith, L.L.; Faqeih, E.; Mohamed, J.; Gupta, V.A.; Alkuraya, F.S. ZBTB42 mutation defines a novel lethal congenital contracture syndrome (LCCS6). *Hum. Mol. Genet.* **2014**, *23*, 6584–6593. [[CrossRef](#)] [[PubMed](#)]
67. Sharma, S.; Quintana, A.; Findlay, G.M.; Mettlen, M.; Baust, B.; Jain, M.; Nilsson, R.; Rao, A.; Hogan, P.G. An siRNA screen for NFAT activation identifies septins as coordinators of store-operated Ca²⁺ entry. *Nature* **2013**, *499*, 238–242. [[CrossRef](#)]
68. Sivan, G.; Martin, S.E.; Myers, T.G.; Buehler, E.; Szymczyk, K.H.; Ormanoglu, P.; Moss, B. Human genome-wide RNAi screen reveals a role for nuclear pore proteins in poxvirus morphogenesis. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 3519–3524. [[CrossRef](#)]
69. Jeannot, E.; Harle, A.; Holmes, A.; Sastre-Garau, X. Nuclear factor IX is a recurrent target for HPV16 insertions in anal carcinomas. *Genes Chromosomes Cancer* **2018**, *57*, 638–644. [[CrossRef](#)] [[PubMed](#)]
70. Raghuram, V.; Weber, S.; Raber, J.; Chen, D.H.; Bird, T.D.; Maylie, J.; Adelman, J.P. Assessment of mutations in KCNN2 and ZNF135 to patient neurological symptoms. *Neuroreport* **2017**, *28*, 375–379. [[CrossRef](#)]
71. Wu, H.; Wang, Y.; Zhang, Y.; Yang, M.; Lv, J.; Liu, J.; Zhang, Y. TALE nickase-mediated SP110 knockin endows cattle with increased resistance to tuberculosis. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1530–E1539. [[CrossRef](#)]
72. Ren, M.; Qin, H.; Ren, R.; Cowell, J.K. Ponatinib suppresses the development of myeloid and lymphoid malignancies associated with FGFR1 abnormalities. *Leukemia* **2013**, *27*, 32–40. [[CrossRef](#)]
73. Grismayer, B.; Solch, S.; Seubert, B.; Kirchner, T.; Schafer, S.; Baretton, G.; Schmitt, M.; Luther, T.; Kruger, A.; Kotsch, M.; et al. Rab31 expression levels modulate tumor-relevant characteristics of breast cancer cells. *Mol. Cancer* **2012**, *11*, 62. [[CrossRef](#)] [[PubMed](#)]
74. Siatecka, M.; Bieker, J.J. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood* **2011**, *118*, 2044–2054. [[CrossRef](#)]
75. Lomniczi, A.; Wright, H.; Castellano, J.M.; Matagne, V.; Toro, C.A.; Ramaswamy, S.; Plant, T.M.; Ojeda, S.R. Epigenetic regulation of puberty via Zinc finger protein-mediated transcriptional repression. *Nat. Commun.* **2015**, *6*, 10195. [[CrossRef](#)]
76. Fleige, A.; Alberti, S.; Grobe, L.; Frischmann, U.; Geffers, R.; Muller, W.; Nordheim, A.; Schippers, A. Serum response factor contributes selectively to lymphocyte development. *J. Biol. Chem.* **2007**, *282*, 24320–24328. [[CrossRef](#)] [[PubMed](#)]
77. Unoki, M.; Okutsu, J.; Nakamura, Y. Identification of a novel human gene, ZFP91, involved in acute myelogenous leukemia. *Int. J. Oncol.* **2003**, *22*, 1217–1223. [[CrossRef](#)] [[PubMed](#)]

78. Ooi, Y.S.; Stiles, K.M.; Liu, C.Y.; Taylor, G.M.; Kielian, M. Genome-wide RNAi screen identifies novel host proteins required for alphavirus entry. *PLoS Pathog.* **2013**, *9*, e1003835. [[CrossRef](#)]
79. Passegue, E.; Jochum, W.; Schorpp-Kistner, M.; Mohle-Steinlein, U.; Wagner, E.F. Chronic myeloid leukemia with increased granulocyte progenitors in mice lacking junB expression in the myeloid lineage. *Cell* **2001**, *104*, 21–32. [[CrossRef](#)]
80. Minoretti, P.; Arra, M.; Emanuele, E.; Olivieri, V.; Aldeghi, A.; Politi, P.; Martinelli, V.; Pesenti, S.; Falcone, C. A W148R mutation in the human FOXD4 gene segregating with dilated cardiomyopathy, obsessive-compulsive disorder, and suicidality. *Int. J. Mol. Med.* **2007**, *19*, 369–372. [[CrossRef](#)] [[PubMed](#)]
81. Kronke, J.; Udeshi, N.D.; Narla, A.; Grauman, P.; Hurst, S.N.; McConkey, M.; Svinkina, T.; Heckl, D.; Comer, E.; Li, X.; et al. Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **2014**, *343*, 301–305. [[CrossRef](#)]
82. Bastian, H.; Gruss, P. A murine even-skipped homologue, *Evx 1*, is expressed during early embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* **1990**, *9*, 1839–1852. [[CrossRef](#)]
83. Hammoud, S.; Emery, B.R.; Dunn, D.; Weiss, R.B.; Carrell, D.T. Sequence alterations in the YBX2 gene are associated with male factor infertility. *Fertil Steril* **2009**, *91*, 1090–1095. [[CrossRef](#)] [[PubMed](#)]
84. Lee, J.; Park, H.S.; Kim, H.H.; Yun, Y.J.; Lee, D.R.; Lee, S. Functional polymorphism in H2BFWT-5'UTR is associated with susceptibility to male infertility. *J. Cell Mol. Med.* **2009**, *13*, 1942–1951. [[CrossRef](#)] [[PubMed](#)]
85. Konig, R.; Zhou, Y.; Elleder, D.; Diamond, T.L.; Bonamy, G.M.; Ireland, J.T.; Chiang, C.Y.; Tu, B.P.; De Jesus, P.D.; Lilley, C.E.; et al. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* **2008**, *135*, 49–60. [[CrossRef](#)] [[PubMed](#)]
86. Madani, N.; Millette, R.; Platt, E.J.; Marin, M.; Kozak, S.L.; Bloch, D.B.; Kabat, D. Implication of the lymphocyte-specific nuclear body protein Sp140 in an innate response to human immunodeficiency virus type 1. *J. Virol.* **2002**, *76*, 11133–11138. [[CrossRef](#)]
87. Mesuraca, M.; Galasso, O.; Guido, L.; Chiarella, E.; Scicchitano, S.; Vatrinet, R.; Morrone, G.; Bond, H.M.; Gasparini, G. Expression profiling and functional implications of a set of zinc finger proteins, ZNF423, ZNF470, ZNF521, and ZNF780B, in primary osteoarthritic articular chondrocytes. *Mediators Inflamm.* **2014**, *2014*, 318793. [[CrossRef](#)]
88. Metzger, J.; Schrimpf, R.; Philipp, U.; Distl, O. Expression levels of LCORL are associated with body size in horses. *PLoS ONE* **2013**, *8*, e56497. [[CrossRef](#)]
89. Li, P.; Guo, H.; Zhou, G.; Shi, H.; Li, Z.; Guan, X.; Deng, Z.; Li, S.; Zhou, S.; Wang, Y.; et al. Increased ZNF84 expression in cervical cancer. *Arch. Gynecol. Obstet.* **2018**, *297*, 1525–1532. [[CrossRef](#)]
90. Chittoor, G.; Haack, K.; Mehta, N.R.; Laston, S.; Cole, S.A.; Comuzzie, A.G.; Butte, N.F.; Voruganti, V.S. Genetic variation underlying renal uric acid excretion in Hispanic children: The Viva La Familia Study. *BMC Med. Genet.* **2017**, *18*, 6. [[CrossRef](#)]
91. Yamaguchi-Kabata, Y.; Morihara, T.; Ohara, T.; Ninomiya, T.; Takahashi, A.; Akatsu, H.; Hashizume, Y.; Hayashi, N.; Shigemizu, D.; Boroevich, K.A.; et al. Integrated analysis of human genetic association study and mouse transcriptome suggests LBH and SHF genes as novel susceptible genes for amyloid-beta accumulation in Alzheimer's disease. *Hum. Genet.* **2018**, *137*, 521–533. [[CrossRef](#)]
92. Wang, Y.; Yabuuchi, A.; McKinney-Freeman, S.; Ducharme, D.M.; Ray, M.K.; Chawengsaksophak, K.; Archer, T.K.; Daley, G.Q. *Cdx* gene deficiency compromises embryonic hematopoiesis in the mouse. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7756–7761. [[CrossRef](#)]
93. Mortlock, D.P.; Innis, J.W. Mutation of HOXA13 in hand-foot-genital syndrome. *Nat. Genet.* **1997**, *15*, 179–180. [[CrossRef](#)]
94. Mackay, D.J.; Callaway, J.L.; Marks, S.M.; White, H.E.; Acerini, C.L.; Boonen, S.E.; Dayanikli, P.; Firth, H.V.; Goodship, J.A.; Haemers, A.P.; et al. Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. *Nat. Genet.* **2008**, *40*, 949–951. [[CrossRef](#)] [[PubMed](#)]
95. Zhang, E.E.; Liu, A.C.; Hirota, T.; Miraglia, L.J.; Welch, G.; Pongsawakul, P.Y.; Liu, X.; Atwood, A.; Huss, J.W., 3rd; Janes, J.; et al. A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell* **2009**, *139*, 199–210. [[CrossRef](#)] [[PubMed](#)]
96. Crujeiras, A.B.; Morcillo, S.; Diaz-Lagares, A.; Sandoval, J.; Castellano-Castillo, D.; Torres, E.; Hervas, D.; Moran, S.; Esteller, M.; Macias-Gonzalez, M.; et al. Identification of an episinature of human colorectal cancer associated with obesity by genome-wide DNA methylation analysis. *Int. J. Obes.* **2019**, *43*, 176–188. [[CrossRef](#)]
97. Banerjee, P.; Kleyn, P.W.; Knowles, J.A.; Lewis, C.A.; Ross, B.M.; Parano, E.; Kovats, S.G.; Lee, J.J.; Penchaszadeh, G.K.; Ott, J.; et al. TULP1 mutation in two extended Dominican kindreds with autosomal recessive retinitis pigmentosa. *Nat. Genet.* **1998**, *18*, 177–179. [[CrossRef](#)]
98. Broen, K.; Levenga, H.; Vos, J.; van Bergen, K.; Fredrix, H.; Greupink-Draaisma, A.; Kester, M.; Falkenburg, J.H.; de Mulder, P.; de Witte, T.; et al. A polymorphism in the splice donor site of ZNF419 results in the novel renal cell carcinoma-associated minor histocompatibility antigen ZAPHIR. *PLoS ONE* **2011**, *6*, e21699. [[CrossRef](#)]
99. Michael, A.K.; Harvey, S.L.; Sammons, P.J.; Anderson, A.P.; Kopalle, H.M.; Banham, A.H.; Partch, C.L. Cancer/Testis Antigen PASD1 Silences the Circadian Clock. *Mol. Cell* **2015**, *58*, 743–754. [[CrossRef](#)] [[PubMed](#)]
100. Edqvist, P.H.; Myers, S.M.; Hallbook, F. Early identification of retinal subtypes in the developing, pre-laminated chick retina using the transcription factors Prox1, Lim1, Ap2alpha, Pax6, Isl1, Isl2, Lim3 and Chx10. *Eur. J. Histochem.* **2006**, *50*, 147–154.
101. Miletich, I.; Yu, W.Y.; Zhang, R.; Yang, K.; Caixeta de Andrade, S.; Pereira, S.F.; Ohazama, A.; Mock, O.B.; Buchner, G.; Sealby, J.; et al. Developmental stalling and organ-autonomous regulation of morphogenesis. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19270–19275. [[CrossRef](#)]

102. Murthi, P.; Said, J.M.; Doherty, V.L.; Donath, S.; Nowell, C.J.; Brennecke, S.P.; Kalionis, B. Homeobox gene DLX4 expression is increased in idiopathic human fetal growth restriction. *Mol. Hum. Reprod.* **2006**, *12*, 763–769. [[CrossRef](#)]
103. Li, S.; Qiu, F.; Xu, A.; Price, S.M.; Xiang, M. Barhl1 regulates migration and survival of cerebellar granule cells by controlling expression of the neurotrophin-3 gene. *J. Neurosci.* **2004**, *24*, 3104–3114. [[CrossRef](#)] [[PubMed](#)]
104. Song, O.R.; Queval, C.J.; Iantomasi, R.; Delorme, V.; Marion, S.; Veyron-Churlet, R.; Werkmeister, E.; Popoff, M.; Ricard, I.; Jouny, S.; et al. ArfGAP1 restricts Mycobacterium tuberculosis entry by controlling the actin cytoskeleton. *EMBO Rep.* **2018**, *19*, 29–42. [[CrossRef](#)] [[PubMed](#)]
105. Warner, N.; Burberry, A.; Pliakas, M.; McDonald, C.; Nunez, G. A genome-wide small interfering RNA (siRNA) screen reveals nuclear factor-kappaB (NF-kappaB)-independent regulators of NOD2-induced interleukin-8 (IL-8) secretion. *J. Biol. Chem.* **2014**, *289*, 28213–28224. [[CrossRef](#)] [[PubMed](#)]
106. Jackson, M.; Watt, A.J.; Gautier, P.; Gilchrist, D.; Driehaus, J.; Graham, G.J.; Keebler, J.; Prugnolle, F.; Awadalla, P.; Forrester, L.M. A murine specific expansion of the RhoX cluster involved in embryonic stem cell biology is under natural selection. *BMC Genom.* **2006**, *7*, 212. [[CrossRef](#)]
107. Passananti, C.; Corbi, N.; Paggi, M.G.; Russo, M.A.; Perez, M.; Cotelli, F.; Stefanini, M.; Amati, P. The product of Zfp59 (Mfg2), a mouse gene expressed at the spermatid stage of spermatogenesis, accumulates in spermatozoa nuclei. *Cell Growth Differ.* **1995**, *6*, 1037–1044.
108. Saga, Y. Genetic rescue of segmentation defect in MesP2-deficient mice by MesP1 gene replacement. *Mech. Dev.* **1998**, *75*, 53–66. [[CrossRef](#)]
109. Li, Y.; Behringer, R.R. Esx1 is an X-chromosome-imprinted regulator of placental development and fetal growth. *Nat. Genet.* **1998**, *20*, 309–311. [[CrossRef](#)]
110. Kranz, D.; Boutros, M. A synthetic lethal screen identifies FAT1 as an antagonist of caspase-8 in extrinsic apoptosis. *EMBO J.* **2014**, *33*, 181–197. [[CrossRef](#)]