bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Comparison of developmental genome expression in rodent molars reveals extensive developmental system drift.

Marie Sémon[1]*, Laurent Guéguen[2], Coraline Petit[1], Carine Rey[1,2], Anne Lambert, Manon Peltier, Sophie Pantalacci[1]*

[1] *UnivLyon, ENS de Lyon, Univ Claude Bernard, CNRS UMR 5239, INSERM U1210, Laboratoire de Biologie et Modélisation de la Cellule, 15 parvis Descartes, F-69007, Lyon, France* [2] *Laboratoire de Biométrie et Biologie Évolutive (LBBE), Université de Lyon, Université Lyon 1, CNRS, Villeurbanne, France* *Equal contribution, and corresponding authors

Abstract

In evolution, it is widely believed that phenotypic changes root in developmental changes and phenotype conservation, in developmental conservation. Seeming phenotype conservation may however hide evolutionary changes in the underlying developmental mechanisms by which a trait is produced. This cryptic evolution is also called Developmental System Drift, and the extent of this phenomenon unclear. We used a well-characterized evo-devo model system, rodent molars, to test the correlation between phenotypic and developmental evolution. Between mouse and hamster, the morphology of the lower molars has much less diverged than the morphology of the upper molars. Is development accordingly more conserved? We compared molar crown formation with a standard approach, and with a tight transcriptome time-series to get a quantitative molecular profiling of developmental states. Our data identify common trends in the development of all molars. Upper and lower molars have their specificities since the early steps of morphogenesis, at the levels of the pattern of cusp formation, cellular composition and biased gene expression. The extent of difference in lower vs. upper molar development within one species does correlate with the extent of difference in final morphology. However, the specificity of lower vs. upper molar development is drowned among the rapid evolution of development, which is highly species-specific in term of expression levels and temporal profiles. Divergence in developmental systems is almost as high for lower as it is for upper molar, despite much lesser morphological changes in lower molar crown. Hence, our results point an extensive drift in this developmental system. Because serial organs are largely sharing gene networks, it supports previous theoretical work that suggest a causal link between pleiotropy and DSD.

Introduction

Comparing embryonic development to understand the evolution of forms is at the roots of evolutionary biology. A generally accepted idea is that the variation in development between species should be proportional to their variation in adult shape. When a phenotype differs between species, it is tacitly assumed that the development differs. When

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the phenotype is similar in different species, then it is tacitly assumed that they have identical underpinnings, and little attention is given to the comparison of their development. But even when a phenotype is similar in different species, this seeming conservation may hide evolutionary changes in the underlying developmental mechanisms by which that trait is produced. This cryptic evolution is also called Developmental System Drift (DSD, [11, 12, 32, 17]). An emblematic example is the development of nematodes vulva, where quantitative changes of the gene signalling network in different nematode species have no impact on the patterning of epidermal cells [17]. Another example is the segmentation in larval development of different dipteran insects [36]. At a first glance, DSD is a less intuitive concept than developmental conservation. It may be useful to remember that natural selection mainly acts on the final product of development. Then, if divergent developmental paths are used to reach the same final phenotype, this should not make a big difference. Further taking into account that genomes are constantly mutating, DSD appear as a likely alternative to developmental conservation.

Because empirical evidence is still scarce, the extent of developmental conservation, and reciprocally, the prevalence of DSD is still unknown in most systems. Furthermore, how developmental gene expression is globally modified in the case of DSD remains unexplored. The factors that favour or prevent DSD are also largely unknown: for instance, theoretical work suggests that pleiotropy, which characterizes developmental genes, may be a factor facilitating DSD [23].

Rodent first molar is an appropriate model for asking whether variation in development mirrors variation in crown morphologies or, on the contrary, whether DSD is prevalent. Basic steps of tooth development have been described a long time ago, and lower molar morphogenesis is well known in mouse ([4, 16, 28] Figure 1-A): Initially, the jaw homeotic-code is transferred to the molars presumptive field (upper: Dlx1,2; lower: Dlx1,2,5,6; molar: Barx1). Later in development, dental epithelium grows and folds to first embed the dental mesenchyme and then form the cusps. Cusps initiate in a sequential manner in the growing tooth germ, and are marked by a signaling center called secondary enamel knot (SEK; Figure 1-A). This involves the iterative use of a regulatory network with key signalling pathways which is thought to be evolutionary conserved [16]. Importantly, crown shape is determined by morphodynamic mechanisms, which is a mode of development where patterning, morphogenesis and growth occur concomitantly. Morphodynamic development confer particular variational properties [29]. For example, the genotype-phenotype relationship is far from straightforward: close genotypes may underlie distant phenotypes, and reciprocally.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

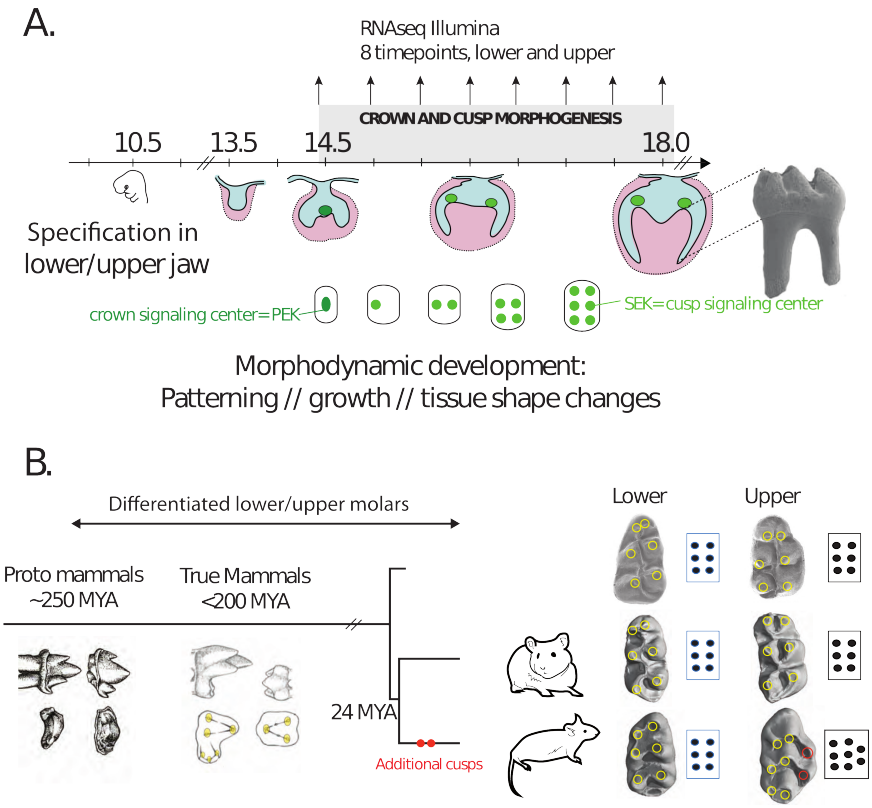

Figure 1. Main characteristics of rodent molar crown evolution and development. A- Molar crown morphogenesis involves morphodynamic mechanisms, where patterning and growth are concomitant. The dental epithelium (blue) grows and folds to first embed the dental mesenchyme and later form the cusps. Cusps initiate in a sequential manner and are marked by a signaling center called secondary enamel knot (SEK). B- Morphological evolution of lower/upper molars. Upper and lower molars displayed different morphologies since their origins in early mammals. In the mouse lineage, the upper molar was drastically transformed with the acquisition of a third cusp row (red). In contrast, the lower molar displayed much more limited morphological changes. In hamster, both the lower and upper molar kept a dental plan similar to the ancestral plan.

The evolution of molar morphology is extremely well documented. Thanks to their abundant fossil record, teeth have had a central role in documenting the evolution of vertebrates since the 19th century. In mammals particularly, molar crown shape provides characters for both fine and large-scale phylogenetic reconstructions. It can also directly reflect adaptations such as changes in diet, with typical features being recurrently found in carnivorous or herbivorous species [16]. More recently, the molar has become a forefront model in evo-devo, with the development of an *in silico* model of tooth development linking variation in development with variation in morphology [31]. Within rodents, many changes have been reported, with alternating periods of stasis, gradual changes and more rapid morphological evolution [25, 24]. One particularly drastic change occurred in the mouse lineage: the crown of the upper molar was drastically transformed with the acquisition of a third cusp row, making a pattern of cusp number and arrangement which is called the "murine dental plan" (Figure 1-B [18]). In contrast, the lower molar displayed much more limited morphological changes. In ham-

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ster, both the lower and upper molar kept a dental plan which is similar to the ancestral plan. Hence, in mouse and hamsters, the crown shape in lower molars has been much less modified in evolution than in upper molars. Is development accordingly more conserved in lower molars?

We quantify similarities in development during the period of crown morphogenesis, and interpret them in relation with variation in adult morphology. We combine a descriptive approach with a cusp marker gene followed in large embryonic series, and a comparative transcriptomics approach on a tightly sampled RNAseq dataset. The extent of difference in lower versus upper molar development within one species does correlate with the extent of difference in final morphology. However, the specificity of lower versus upper molar development is drowned among the rapid evolution of development, which is highly species-specific in term of expression levels and temporal profiles. Divergence in developmental systems is almost as high for lower as it is for upper molar, despite much lesser morphological changes in lower molar crown. Hence our results are fully consistent with an extensive developmental system drift in this system.

Results

Patterns and timing of cusp formation is overall conserved in all types of tooth, with specificities attributable to history and morphological differences

To measure the variation in hamster and mouse crown molar development, we started by investigating the timing and pattern of cusp patterning. For this purpose, we followed the spatial pattern and the temporal dynamics of cusp formation with series of fixed embryos hybridized against the cusp marker *Fgf4* and dated according to embryo body weight. We collected 153 lower (and 141 upper) molars in mouse, and 86 lower (87 lower) molars in hamster, which makes a dataset amenable to quantitative analysis. We then built an in house method, based on Markov models, to infer from these data the amount of time spent in each stereotyped stage (see methods). The rationale underlying this method is very simple: under the hypothesis that embryos are sampled uniformly during development, long-lasting developmental stages should be sampled more often than ephemeral stages. An interesting feature of this model is to handle statistical comparisons of the relative duration of the different stages, within and between species and tooth types.

Cusp formation exhibits clear spatial and temporal similarities in all types of tooth (Figure 2). The first cusp patterned is bucco-central in all cases (paracone for lower mo-

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

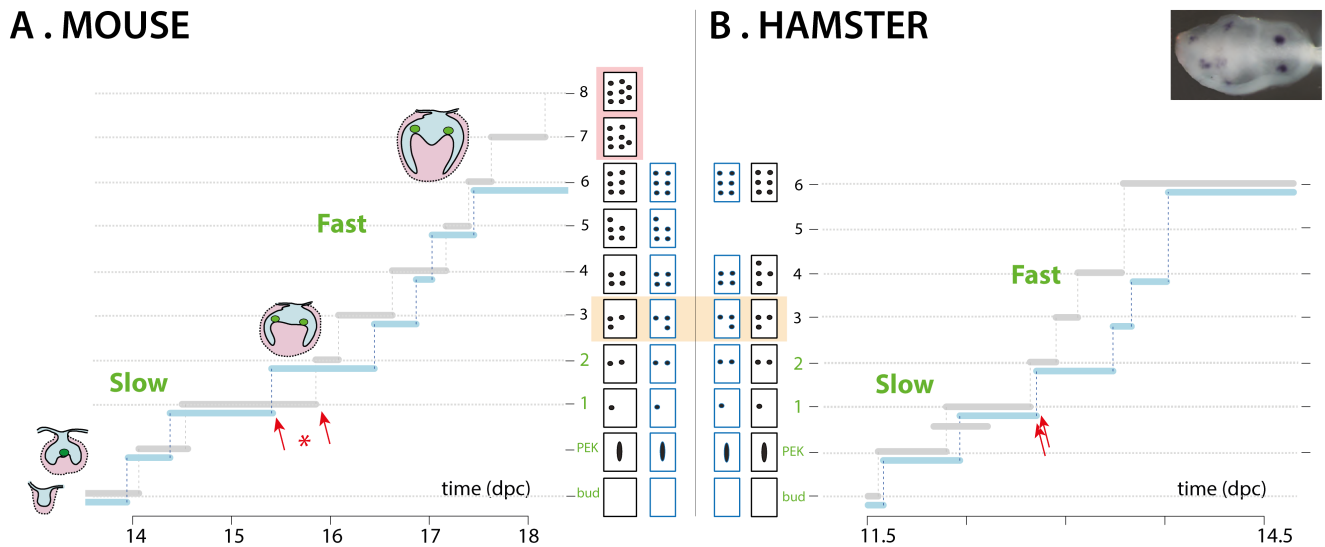

Figure 2. Conservation and divergence in dynamics and pattern of cusp formation in mouse (left) and hamster (right). Pattern and dynamics of cusp formation was followed in series of fixed embryos hybridized against the cusp marker Fgf4 and dated according to embryo body weight. We built an in house method, based on Markov models, to infer from these data the amount of time spent in each stereotyped stage. In mouse (left) the best Markov model fits cusp patterning with a succession of stages with 3 different durations represented by the length of blue and grey bars for lower and upper molars respectively. In hamster, the best models have 2 different durations. Globally the sequences are resembling in both species, with for the lower molar, two long first stages ($1 >>> 2 >>> 3 > 4 > 5 > 6$) and for the upper molar, an extra long duration of the 1 cusp stage, but a short 2-cusp stage ($1 >>>> 2 > 3 > 4 > 5 > 6(> 7 > 8)$.

lar, protoconid for upper molar), followed by a second cusp at its lingual side. The third cusp is posterior is all cases, with later addition to the anterior cusps. The similarities extend also to the rates of cusp addition. The best Markov models fit cusp patterning with 3 different rates in mouse and hamster (LRT to compare model with 3 to model with 2 rates, upper mouse: p-value = 1.710^{-9}; lower mouse : p-value = 3.510^{-5}, upper hamster p-value = 0.008, lower hamster p-value = 0.05). In all cases, the rate of cusp acquisition is biphasic, early cusps being added slowly, and then more rapidly in late crown morphogenesis.

Interestingly, lower and upper molar present distinguishing features, the most notable being the third cusp patterned buccally in upper molars, and lingually in lower molars. Our preliminary results suggest that this is also true in gerbils and acomys (data not shown). Hence, the patterning of the 3 first cusps show a deep conservation (lingual patterning of the second cusp) and a feature distinguishing lower and upper molar (third cusp). Altogether, the conservation of this specific patterning of is possibly reminiscent of early mammalian molar evolution. A further specific characteristic of upper molars is the duration of the 2-cusps stage which lasts 3 to 7 times more in lower by comparison to upper molars (Markov model estimates).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Because two supplementary cusps add to the upper molar dental plan in the mouse lineage, we searched for individual characteristics in mouse upper tooth development. These additional cusps form last, which is conform to expectations of the model of terminal addition where most recent evolutionary features should be added in the later steps of development (see discussion). The spatial pattern of the upper crowns at 4-cusp stage differs between mouse and hamster. Unexpectedly, earlier stages of upper molars development also display specificities, with a significant delay in the patterning of the second cusp (stage 1 is twice longer in mouse compared to hamster, as estimated from the Markov models). Conversely, lower molar development is very similar in mouse and hamster, which mirrors the fact that their dental plan is similar.

Overall, cusp formation exhibit clear similarities in all teeth, with specificities for lower versus upper molar, possibly reminiscent of early mammal molar evolution. As expected from morphology, lower molar development is very similar in mouse and hamster. Mouse additional cusps form last, according to their phylogenetic history (recapitulation [1]). Clear differences are also seen earlier in upper molars (stage 1 and 4). Finally, this morphological analysis of cusp patterning shows clear specificities of upper and lower tooth and comparatively much less characters distinguishing the species. Is this pattern attributable to evolutionarily conserved underlying gene regulatory networks? Temporal genome expression is now fully recognized to register the course of development [2] and transcriptomes are widely used to quantify differences between species and organs e.g. [27, 38, 19, 14, 6, 5]. Next we constructed a transcriptome dataset to compare upper and lower molar development in mouse and hamster.

Lower and upper molars in mouse and hamster largely share a common dynamics of genome expression but also exhibit distinguishing features

We prepared RNA-seq libraries from upper and lower whole tooth germs in mouse and hamster, corresponding to 8 development stages roughly equally spaced during the period of crown development. We quantified expression levels of 10,516 protein-coding genes in mouse and hamster, for which the sequence of the transcript was made rigorously similar, despite the fact that no genome sequence is available in hamster (see methods). In a first step, we conducted the analyses separately for each species. Visually inspecting the projection of the transcriptome samples into a space of reduced dimension has become a classical first step in comparative transcriptomics [21, 2, 5]. PCA plots in each species show that the main source of variation reflects a temporal dynamics which is common to both lower and upper molar (31.7% in hamster, 29.2%

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

in mouse; Figure 3).

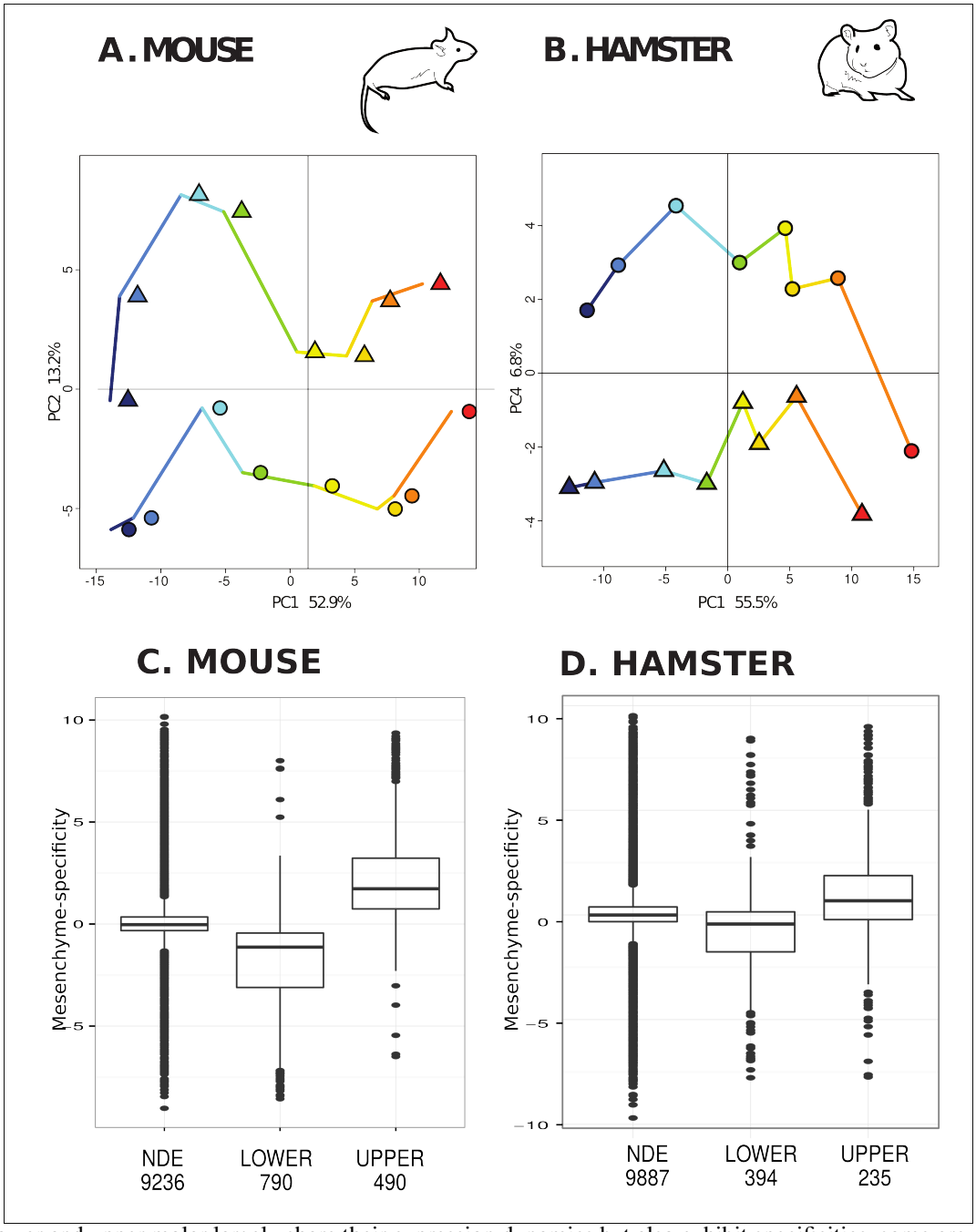

Figure 3. Lower and upper molar largely share their expression dynamics but also exhibit specificities, some are conserved in mouse (A) and hamster (B). (Top) Map of 16 transcriptomes from upper (triangles) and lower (circles) molar germs, at 8 stages of development, on the two first principal components of a PCA analysis (orthologous sets of 10,516 genes in each species). PCA1 orders samples with time, hence represents shared temporal variation. PCA2 separates upper and lower molar germs. The percentage of variation explained by each axis is indicated. (Bottom) Upper genes are more mesenchyme specific and lower genes are more epithelium specific in mouse tooth germs, but not in hamster. Upper and lower genes were defined based on classical test for differential expression between upper and lower molar, taking developmental stage into account (adjusted p-value < 0.05). The number of genes in each category is indicated. Mesenchyme specificity was estimated as the ratio of mesenchyme and epithelium expression levels from pure tissue transcriptomes (log2 fold change).

A clear organ identity is also visible from the second axis of the PCA plot, which

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

is separating upper and lower samples and represents 11.6% of the total variation for hamster, and 16.9% for mouse. The ordinations of the genes on the main PCA axes, which were obtained independently from separate analysis in mouse and hamster, are nonetheless significantly correlated (measurement of the global correlation by coinertia analysis, p-value $< 10^{-3}$). The correlation between mouse and hamster are R = 0.39 for PCA1 (p-value $< 10^{-16}$) and R =0.14 for PCA2 (p-value $< 10^{-16}$). This is congruent with an overall correlation of the ratios of upper/lower expression between species (Pearson R = 0.31, p-value $< 10^{-16}$). An examination of the lists of differentially expressed genes between lower and upper tooth germs reveals a moderate but significant overlap between mouse and hamster: we observed about 3.4-5.2 times enrichment between the lists of DE genes depending on the threshold used (respectively adjusted p-value = 0.1 and 0.05; Chi-squared test p-value $< 10^{-16}$).

A temporal analysis of the main clusters of coexpression also reveals a visible tooth type "identity", with clusters in the lower molars being smoother than clusters in the upper molar (Figure 4; see later analysis of the conservation of these clusters). In a previous analysis in mouse, we had shown that this corresponds to processes of morphogenesis being transiently more active in upper than in lower molars. Here we confirm, using GO analyses, that early tooth germs likely contain less mitotic cells and more migrating cells in upper molar than in lower molar. This pattern is common to mouse and hamster. The timing of this process corresponds to a developmental stage where the transcriptome is most different between lower and upper molars (test of the polynomial term of the linear model, p-value = 0.005 see supplementary Figure 7). Most interestingly, this timing also matches the "3-cusp" developmental stage, where the difference in the arrangement of the third cusp patterned is the main specificity distinguishing lower from upper molars (Figure 2).

The second PCA axis, which separates lower and upper samples, is 45% stronger in mouse than in hamster (16.9% and 11.6%), as would be expected from the difference in adult morphologies. This is congruent with a between-group multivariate analysis (37% stronger, p-value = 0.0045 see Sup figure 7) and with differential analysis data, which detects at least 3 times more genes differentially expressed between upper and lower samples in mouse than in hamster (4.1 more for p-adjust=0.05, 3.70 for p-adjust=0.1). Our previous analysis of mouse data (Pantalacci et al, submitted) has shown that this second axis is partly explained by a difference in the relative mesenchyme proportion between lower and upper molars. This was based on marker genes and pure-tissue microarray data. To explore this further in mouse and compare with the patterns observed in hamster, we prepared RNAseq libraries from pure epithelium and mesenchymal cells

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

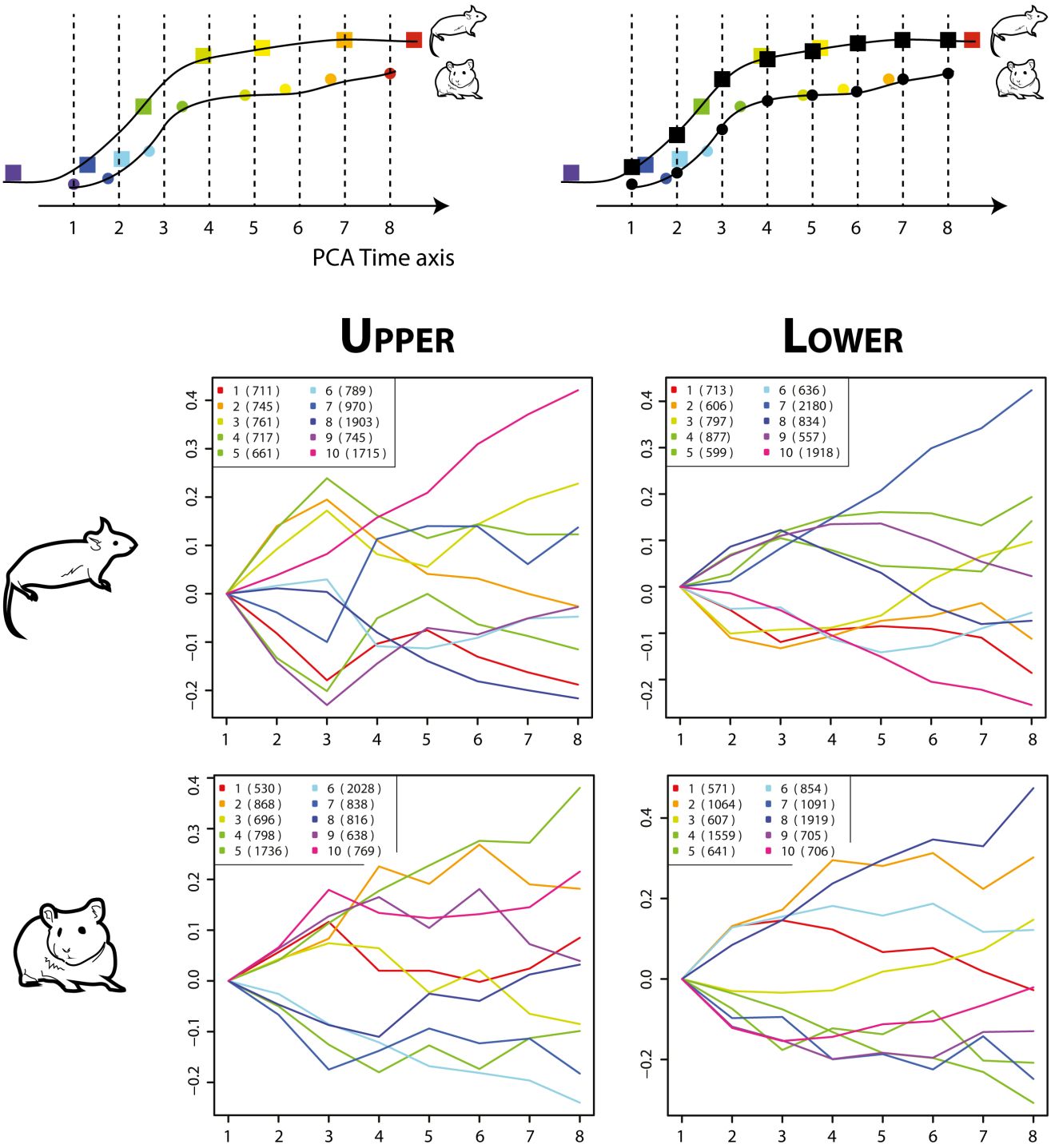

Figure 4. Ten main temporal profiles of expression determined independently for the lower (right) and upper (left) molar, in mouse (top) and in hamster (bottom). The second axis of PCA with all samples (Fig 3) was taken as a global index for developmental time. This PCA2 was split in equally sized 'time' windows. The timepoints were made comparable between species by estimating expression levels in each window. 10 clusters were obtained by K-mean clustering (see methods), and associated number of genes are indicated. Cluster numbering and colours are arbitrary and do not indicate a correspondence between the four conditions. Each profile represents the median of relative expression level (normalized with expression level at the first stage) for all genes associated to a particular cluster.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

in both species (see methods). In mouse, this new pure-tissue data confirm our previous observations (Figure 3] and Sup figure 8): Using deconvolutions, we confirm that the relative proportion of mesenchymal cells is greater in upper than in lower tooth germs, the average of the 8 time points being respectively 53.1% and 44.0% for upper and lower molar in mouse (Wilcoxon test p-value = 0.000622). This quantification is based on 6 markers for enamel knot known from previous literature, 333 epithelial markers and 655 mesenchyme markers based on our pure-tissue transcriptomes (log2 ratio > 3 and adjusted p-value < 0.1), but the result is very robust to changes in marker lists. We computed an index of mesenchyme-specificity which is simply the log-ratio of mesenchyme and epithelium expression levels. This index is significantly correlated with the second PCA axis in mouse (Spearman Rho = 0.42, p-value $< 10^{-16}$) and to the average log ratio of upper and lower expression levels in mouse (Spearman Rho = 0.58, p-value $< 10^{-16}$). Of note, the nature of differentially expressed genes is impacted by this difference in tissue proportions. This is illustrated Figure 3, which shows that genes with a significant bias towards the upper molar are more mesenchyme-specific than genes biased towards the lower molar (Wilcoxon test, p-value $< 10^{-16}$). The median "upper" genes is 3.5 times more expressed in mesenchyme (median log2 fold change of 490 UPPER genes = 1.72) than in epithelium, while the median "lower" DE genes is twice more expressed in epithelium (median of 790 LOWER genes = -1.13). In mouse, the relative proportion of epithelial and mesenchymal cells has a pervasive effect on the difference between upper and lower transcriptomes.

The results of symmetrical analyses conducted in parallel in hamster stand in contrast with these observations, although the index of mesenchyme specificity are extremely well correlated between the two species (R = 0.799, p-value $< 10^{-16}$). The difference in mesenchyme tissue proportion between upper and lower molar is very moderate in hamster, which can be seen in mesenchyme proportions estimated from deconvolutions (respectively 49.9% and 49.2%, Wilcoxon test p-value=0.96, predicted using 6 enamel knot markers, 393 epithelial markers and 568 mesenchyme markers from pure tissue transcriptomes such as log2 ratio > 3 and adjusted p-value ¡ 0.1). This is also visible from the weaker correlation between PCA2 and the index of mesenchymal-specificity in hamster (Spearman Rho = 0.14, p-value $< 10^{-16}$). Mesenchyme-specificity is also much less correlated to the ratio of upper and lower expression levels (Rho = 0.19, p-value $< 10^{-16}$). In consequence, differentially expressed genes biased towards the upper molar are much less mesenchyme-specific in hamster than in mouse (median = 1.6 for 235 UPPER genes, Wilcoxon test, p-value $< 10^{-16}$, Figure 3).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In conclusion, these results on temporal genome expression mirror our findings on the dynamics and pattern of cusp formation, with a shared developmental program, lower/upper specificities found in both species and a supplementary difference between lower and upper tooth germs in mouse which is congruent with their final morphologies. The patterns and timing of cusp formation does not display any obvious species specificity. What is the situation for genome expression?

Expression levels are highly species-specific, including for functionally relevant genes.

The PCA plot of a pooled dataset including all samples from mouse and hamster displays a stricking difference between species (Figure 5-A). Species effect explains 53% of variance, whereas developmental timing drops to 11% and the type of teeth only 4%. RNA-seq analyses are highly sensitive to technical artifacts, such as sequence composition and batch effects. To take care of the first issue, we carefully annotated orthologs based on phylogenies and cropped the sequences of the reference transcriptomes before measuring expression levels, so that they become fully comparable between species (see methods). A controversy related to the second issue erupted recently, when the Mouse ENCODE Consortium reported that comparative gene expression data from human and mouse tend to cluster more by species rather than by tissue on a PCA plot [20]. This analysis was challenged when a confounding sequencing batch effect was removed [10]. To prevent similar batch effects, we analysed an homogeneous part of the dataset (8 stages in mouse and in hamster), for which library preparation and sequencing were performed synchronously in mouse and hamster. This does not change the results (Sup figure 9).

This major species effect is also confirmed by other methods, such as "between" multivariate analyses that quantify the relative effects of the species (54.3% on the homogenized dataset) and of the tooth (2.33%). Using differential expression analysis, we find 8,387 genes that vary significantly according to the species, which is much more than the 130 genes that vary according to the upper/lower difference (DESeq2 analysis taking development stage into account, adjusted p-value <0.1).

Hence, when the dataset is taken globally, the variation in expression between species exceeds largely the variation in expression between organs. In pioneering comparative transcriptomics studies [37] similar patterns were interpreted as the product of rapid evolution of gene expression, and considered as an evidence for its neutral evolution.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

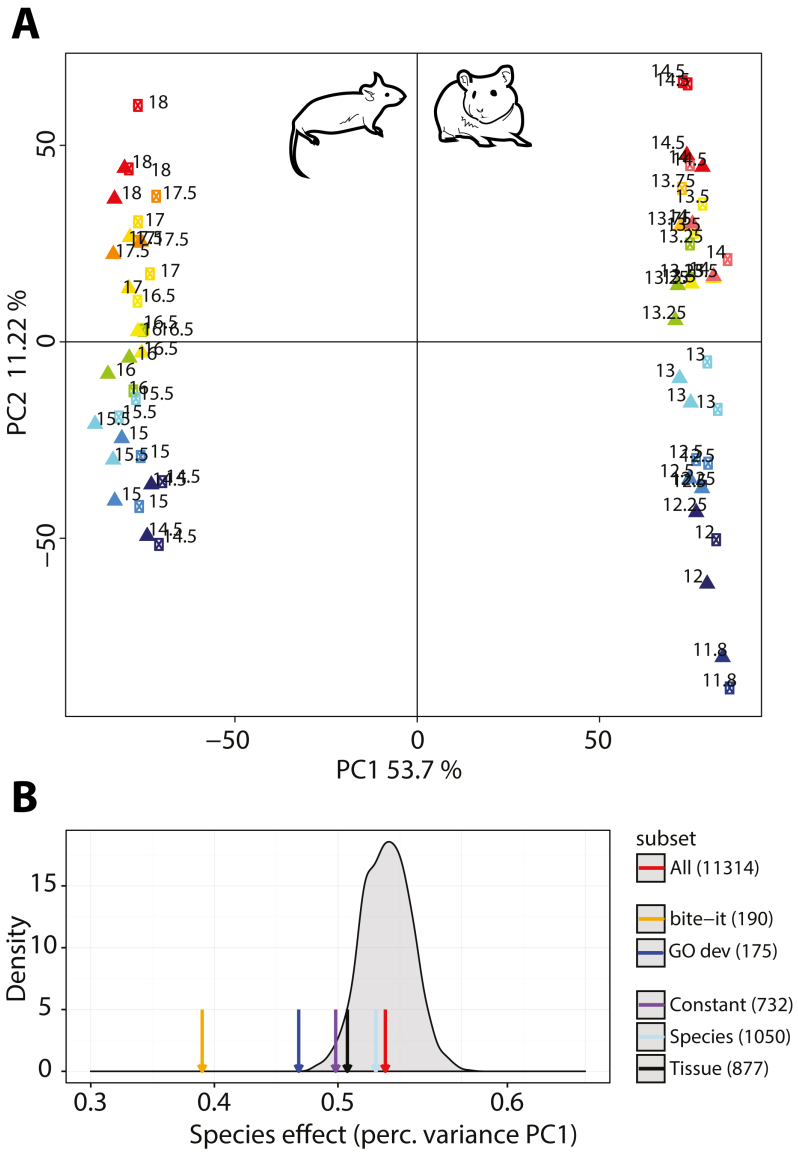

Figure 5. Species effect is the major source of variation when comparing all four transcriptomic series. A- Principal component analysis with 64 samples from both species and tooth types. Expression levels are so divergent between species that tooth type differences (circles and triangles) are negligible in comparison. PCA1 separates mouse samples (circled in black) and hamster samples (circled in white). The color gradient indicates the time (in days) which runs from 11.8 days to 14.5 days in hamster, and from 14.5 days to 18 days in mouse. PCA2 sorts the samples with developmental age. B- Distribution of the species effect for the whole genome and for specific subsets of genes. In grey: distribution of %variation associated to PCA1 for random samplings (200 genes) in the whole genome. Red arrow: value for the whole genome. Other arrows show the %variation associated to PCA1 for specific subsets, genes with acknowledged function in tooth/organ development (bite-it database, and "development" GO category), and genes with different evolutionary rates in adult organs (computed using the dataset from [5]).

According to this interpretation, the changes in expression level that are observed in our dataset may be relatively neutral, without much impact on molar development. However, most comparative transcriptomic studies [7, 5, 3, 22] now demonstrate an overall conservation of expression levels between species, at least for protein-coding genes, which is compatible with strong constraints on the evolution of expression.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Hence, we reasoned that the strong species effect that is observed in our data may be carried by a proportion of the genome only, while the rest is evolving under strong selective constraints. We randomly sampled subsets of 200 genes in the genome, and for each set of genes independently, we conducted a PCA analysis and measured the percentage of variance associated to PCA1 as an index of the species effect. The resulting distribution shows that species effect represents at least half of the total variation in the transcriptomes in all cases, and never less than 40%.

Then, we considered that the species effect should be even more pronounced than the genome average for genes known to be species-specific, and much smaller for genes known to play an important role in tooth development. From a large published dataset providing expression levels in adult organs across a wide range of species, we defined three sets of genes : genes for which expression level varies significantly according to the species (N=1,050), according to the tissue (N=877) and genes that do not vary (N=732 "Constant" genes; see methods). We then measured the species effect in our dataset for these three sets of genes. As expected, the species effect is greater for the set of "species" genes than for "constant" and "tissue" genes (Figure 5-B). But importantly, even genes whose expression level evolve slowly in adult organs, for which the evolution of expression level is unlikely to be neutral, also show a strong species effect (>50%). We then reasoned that changes in the expression levels of many genes may not have much consequence on tooth development. We then measured the species effect on the group of genes annotated to "development" function in Gene Ontology (N = 175 genes), and finally, on a set of genes functionally relevant for tooth development (190 genes from the bite-it database). Even on these sets of genes which are much likely to be important for tooth development, there is a very strong effect (40-46%). In conclusion for this part, expression levels are so divergent between species that tooth type differences are negligible in comparison.

Not only expression levels, but also temporal dynamics is highly species-specific, including for functionally relevant genes.

Decades of studies in evolution and development have highlighted that temporal profiles of gene expression, and not just their expression levels, is instrumental for correct organ development. Hence we compared the profiles of expression in hamster and mouse lower and upper molars. Embryo were sampled so that developmental stages are as comparable as possible between species, and we further refined staging based on transcriptome developmental clock (see methods). Then, we build the ten main expression

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

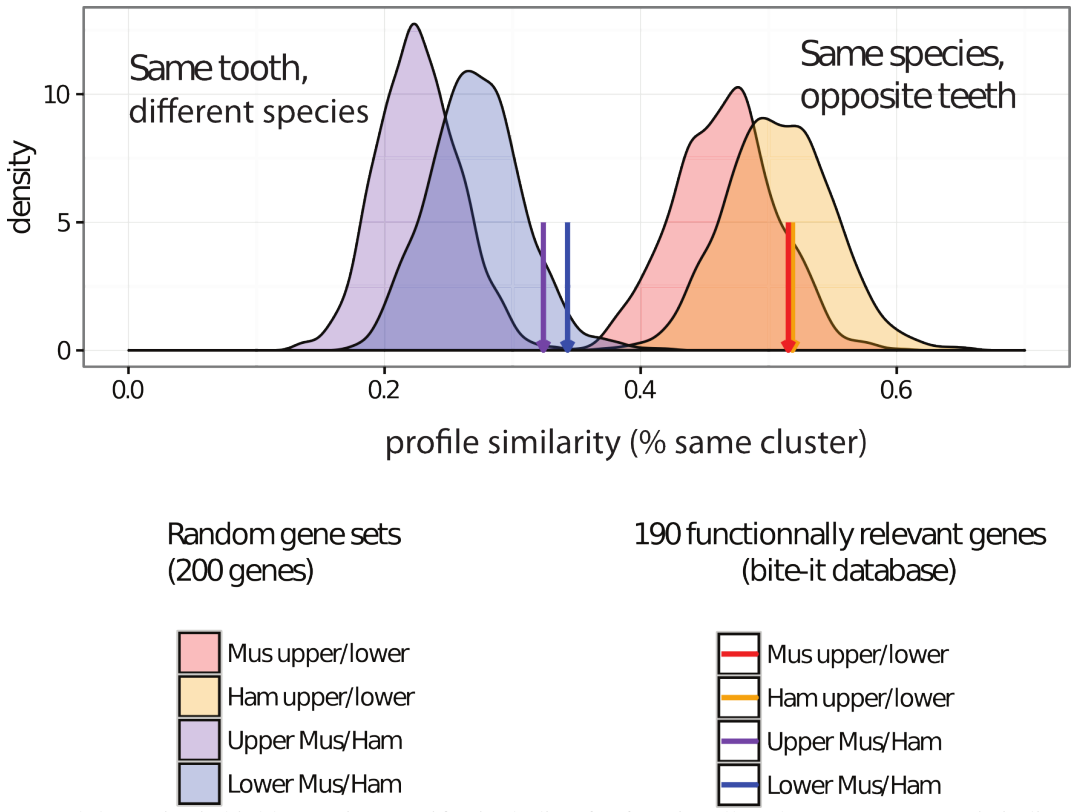

Figure 6. Temporal dynamics is highly species-specific, including for functionally relevant genes. X-axis indicates the similarity of expression profiles between pairs of conditions, as measured by the proportion of genes that adopt the same broad temporal profile of expression. Histograms represent the distribution obtained by randomly sampling sets of 200 genes, and the color indicates the type of comparison. The arrows represent the profile similarities obtained for functionaly relevant genes (190 genes from bite-it database).

profiles independently for each tooth (by clustering analysis, Figure 4) and computed the resemblance between the dynamics of transcriptome in the four conditions. Let's consider we are comparing upper molars in mouse and hamster: For each gene, we took its gene expression profile in the upper molar (VUp), and calculated the correlation of this vector VUp with the 10 main expression clusters drawn from lower molar expression. This way, we obtain the cluster with which the vector VUp is best associated. By construction the expression profile of the same gene in the lower molar (VLo) is associated with one of the ten main expression clusters. Are VUp and VLo associated with the same cluster of expression? Through this simple reasoning, we build an index of similarity between temporal profiles. We sampled randomly sets of 200 genes, to build a distribution of the resemblance in temporal profiles. Lower and upper molar of the same species have twice more resembling temporal profiles than lower molars or upper molars of different species (Figure 6). This is again a strong indicator that profiles of expression are more specific to the species than to the type of tooth. The same effect is seen with a set of genes functionally relevant for tooth development (190 genes from bite-it database). Gene expression temporal dynamics is poorly conserved for a given

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

tooth type, whereas there are extensive similarities between the two teeth of a given species. This is best explained by Developmental Systems Drift (see discussion).

Furthermore, among the 30% of genes whose profile is similar in the same tooth of different species, 75% have in fact a similar profile in all four teeth. This leaves only 7% of genes with a conserved lower or upper specific profile. Then, a minor proportion of the genes display a profile of expression which distinguishes lower and upper molars. Lower molars (6 cusps in mouse and hamster) have slightly more resembling temporal profiles than upper molars (6 cusps in hamster, 8 cusps in mouse: the average profile similarity as measured Figure 6 is 20% higher). Hence, expression dynamics is slightly more conserved in lower molars than in upper molars. This is coherent with adult morphologies, but unexpectedly small : From the adult morphologies, much sharper difference could have been expected.

Discussion

Like in many other developmental systems, global properties of molar development are conserved for very long periods of time. For instance, playing with levels of signaling molecules, mouse teeth have been engineered in vitro to generate tooth shapes which are widespread in the mammalian fossil record [13]. Moreover we know that molar development in different mammal species use similar set of key signaling pathways and transcription factors [15]. Then tooth development was expected to be strongly conserved, especially in closely related species like mouse and hamster (estimated divergence time = 24 MYA [18]). Here, as a read-through of the development process we quantify the similarities in genome expression during molar crown morphogenesis in mouse and hamster. We find that expression profiles are twice more similar within species than within tooth type. This is also true for developmental genes, even for genes known to function in tooth development. This is even more pronounced for expression levels. How to interpret this observation? We did expect a difference in upper molar developmental processes, because the crown morphology has been drastically transformed in mouse. Seeing that the development was nearly as different between lower molars, is much more astonishing. Such rapid evolutionary changes in the developmental programs of both molars, without relation to the final morphology, is best explained by Developmental Systems Drift. Below we hypothesise that the molars, and serial organs in general, may be particularly prone to DSD.

Lower and upper molars morphological evolution was described up to their origins in early mammals. They already displayed different morphologies at the time [34]. Hence, it would be far-stretched to pretend that they have ever shared the same developmental

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

mechanism. But it is both expected [35, 34]) and observed (Pantalacci et al, submitted) that similar gene regulatory network are directing the development of upper and lower first molars. Indeed, different types of teeth in mammals, but also fore- and hindlimbs, plant leaves and insect appendages, are serial organs, that is, multiple realisations of a very similar developmental program, which share the same genes except 1-3 "selector" genes (homeotic genes). This causes extreme pleiotropic effects. A mutation at one gene can affect the development of all iterations of the same serial organ [23]. In accordance with this idea, upper and lower molar shape covary in wild rodent populations/species [26]. Intuitively, one may think that such gene pleiotropy will induce a strong conservation of the development (Fisher's models) by the effect of a strong purifying selection. But at the same time, the idea that pleiotropy would strongly constrain the evolution of these organs is contradicted by the fact that the morphology of different types of serial organs is extremely variable in the same organism and in evolution. Counter-intuitively, pleiotropy may induce extensive developmental system drift, according to a model proposed by Pavlicev and Wagner [23]. This model, termed "selection, pleiotropy and compensation model" (SPC model), is in two steps. One genetic change which has an strong adaptive effect on one organ is fixed (Step 1). Because of pleiotropy, it is associated with antagonist (deleterious) effects on other organs. Hence, this initial adaptive change is followed by subsequent selection in favor of supplementary changes to compensate these pleiotropic effects (Step2). Importantly, although such compensatory changes are selected for by stabilizing selection, they amplify the divergence in development between species, producing extensive developmental systems drift. Serial organ pleiotropy is very high, thus this hypothetic mechanism could work strongly and serial organ be prone to DSD. In our specific case, accommodation of lower molar for drastic upper morphological change in mouse may have further amplified DSD. DSD may facilitate differential morphological evolution of serial appendages, because it is enabling one appendage to cope with extensive changes in the other.

Furthermore, DSD may be facilitated by because molars have a peculiar mode of development. During tooth development, signalling and tissue growth are intermingled, in what is called a morphodynamic mechanism. It has been proposed, based on modelling and experimental data, that Genotype-Phenotype relationship is complex in these cases, with several genotypes producing the same phenotype [30, 31]. In other words, different developmental trajectories can lead to the same phenotype. DSD may be facilitated in morphodynamic systems like rodent molar [33]. We note that many serial organs develop through morphodynamic mechanisms (arthropod appendages, limb, teeth, leaves. . .). Hence the model proposed here may apply to other serial appendages

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

in general. To push the model of Pavlicev and Wagner further, we suggest that DSD and morphodynamic mechanisms may work hand in hand to alleviate the pleiotropic constraint. This may have render possible the emergence of generalized pleiotropy in developmental systems, while preserving their evolvability.

Methods (preliminary draft)

Mice breeding and embryo sampling

CD1 (ICR) adult mice and AURA adult hamsters were purchased from Charles River and Janvier Laboratories, respectively (France). Females were mated overnight (night at 20:00 pm) and the noon after morning detection of a vaginal plug was indicated as embryonic day (ED) 0.5. Other breeding pairs were kept in a light-dark reversed cycle (night at 12:00 pm), so that the next day at 16:00 pm was considered as ED1.0. Pregnant females were killed by cervical dislocation (mice) or medicamentous overdose (hamster) and embryos were harvested on cooled Hank's or DMEM advanced medium, weighted as described in [73] and immediately decapitated. This study was performed in strict accordance with the European guidelines 2010/63/UE and was approved by the Animal Experimentation Ethics Committee CECCAPP (Lyon, France; reference ENS_2012_045).

Epithelium dissociations and in situ hybridizations

Complete or hemi mandibles and maxillae were dissected in Hank's medium and treated with Dispase (Roche) 100% at 37°for 1h30 to 2h20 depending on embryonic stage. Epithelium was carefully removed and fixed overnight in PFA 4% at 4°. DIG RNA antisense Fgf4 probe were prepared from plasmids described elsewhere [74]. In situ hybridization was done according to a standard protocol. Photographs were taken on a Leica M205FA stereomicroscope with a Leica DFC450 digital camera (Wetzlar, Germany) or on a Zeiss LUMAR stereomicroscope with a CCD CoolSNAP camera (PLATIM, Lyon).

RNA-seq sample preparation

In mouse, 16 samples were prepared for upper and lower whole tooth germ RNA-seq analysis. This adds to the 16 samples described previously (Pantalacci et al., submitted). We obtained a total of 32 samples representing 8 stages (ED14.5, 15.0, 15.5, 16.0, 16.5, 17.0, 17.5, 18.0), coming from 8 individuals. In hamster, 32 samples were prepared, representing roughly 8 stages (from ED12.5 to 14.5), coming from uppers and lower first molars of 16 individuals.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Each sample contained 2 tooth germs, the left and right first molars (M1) of the same male individual, and for a given stage, the upper and lower samples came from the same individual. The heads of harvested embryo were kept for a minimal amount of time in cooled PBS (small scale) or advanced DMEM medium (large scale). The M1 lower and upper germs were dissected under a stereomicroscope and stored in 200ul of RNA later (SIGMA). Similarly dissected tooth germs from the same litter and same weight were fixed overnight in PFA 4% for immunolocalization and 3D reconstruction (see later). Another embryo of the same litter and same weight was processed as indicated above for Fgf4 in situ hybridization. For epithelium and mesenchyme specific samples, tooth germs were dissected as above, then the two components were separated following a dispase treatment (15 minutes, 37°, ROCHE) and immediately stored in RNA later (SIGMA).

Total RNA was prepared using the RNeasy micro kit from QIAGEN following lysis with a Precellys homogenizer. RNA integrity was controlled on a Bioanalyzer (Agilent Technologies, a RIN of 10 was reached for all samples used in this study). PolyA+ libraries of the large-scale dataset were prepared with the Truseq V2 kit (Illumina), starting with 150ng total mRNA and reducing the amplification step to only 12 cycles and sequenced (100 bp paired-end sequences) on an Illumina Hi-seq2000sequencer at the GENOSCOPE (Evry, France). Raw data is available in the gene expression omnibus repository (XX).

Expression levels estimation using RNA-seq and differential expression analysis.

Comparable sequences for 10,516 pairs of orthologous genes were extracted by an automated and accurate pipeline (Amalgam, GitHub `https://github.com/CarineRey/apytram`) which combines data from RNA-Seq and existing multi-species gene family alignments. In Amalgam, reconstructed sequences are first attributed at the good gene family and then aligned in the corresponding existing multi-species gene family alignment. Then an accurate gene tree is inferred to use phylogenetic proximity between sequences as gene annotations. This allows taking into account complete genomic knowing of this clade (duplication, and losses) and not of an unique reference species. Reads were mapped to theses sequences using Bowtie2 (version 2.0.2) with standard settings. We obtained a median of XX million reads that mapped uniquely.

Raw counts for the published datasets by Brawand et al [5] were obtained, and anal-

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ysed by DEseq2 to obtain the lists of genes differentially expressed with species or tissue factor. A subset of the whole dataset (primates and rodents only) was used in this analysis.

Two main analysis used in the text include differential gene expression analysis (DE) and multivariate analysis. Differential expression analysis was performed using DEseq package (DESeq2, version 1.6.3 [21]), with settings indicated in the main text. Multivariate analyses (principal component analyses, between-class analysis and coinertia analysis) were performed using ade4 package (ade4_1.6-2 [8]).

Gene Ontology analysis

Gene ontology (GO) analysis were performed and visualized with GORILLA and GOStats (version 2.34.0, with `Mmusculus.UCSC.mm10.ensGene_3.1.2` and GO.db_3.1.2), using the full list of genes expressed in the corresponding dataset as a background.

Clustering of upper and lower time-series and estimation of similarity between temporal profiles in pairs of conditions.

First, to allow the comparison of temporal profiles between different species and different tooth types, we constructed a dataset with comparable timepoints. For this, we considered the second principal component (PCA2) as an index of developmental timing. We divided the range of coordinates on PCA2 in 8 windows of equal size, and for each gene and each window we extrapolated the expression level that is expected at each window boundary (taking the median expression levels of the samples that belong to this window do not change qualitatively the results). This is schematized on the top of the figure 4.

Then, for each gene and at each stage of development, we took the level of expression divided by the initial level of expression (i.e. level at stage ED 14.5). To classify the genes we first generated theoretical profiles representing all possible combinations of transitions between stages (expression level being either increasing, decreasing or flat between the 8 consecutive stages), making up a total of 2,187 profiles. We then correlated each real expression profile to each theoretical profile using Spearman correlations. We clustered the obtained correlation matrix using K-means (choosing 10 clusters, and checking that the obtained profiles are repeatable over several iterations of the method). The median of the expression profiles of these 10 clusters are drawn on figure 4. Similarities between the different time-series was computed as follows, taking

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the example of lower/upper comparison in mouse, but same reasoning was applied to all comparisons: We computed the number of genes which lower timeprofile is robustly assigned to one of the 10 main "lower" clusters (Pearson R>0.7). We counted, the number of genes whose timeprofile in upper molar associates best (and with R>0.7) to the same lower cluster: such genes have similar timeprofile during lower and upper molar development. The ratio of this number to the total of genes is an index of profile similarity between lower and upper molars un mouse.

Estimating tissue proportions from RNAseq data : deconvolutions

Performing gene expression deconvolution, is estimating cell type proportions (and/or cell-specific gene expression signatures) from global expression data in heterogeneous samples. Here, we used the R package CellMix ($CellMix_1$.6.2, [9]), to estimate the proportions of the three tissue compartments in our tooth germ transcriptomes. CellMix implements in particular the method Digital Sorting Algorithm (DSA) which performs complete gene expression deconvolution using a set of marker genes only.

We needed therefore to define 3 sets of markers, as specific as possible for each tissue compartment (epithelium, mesenchyme, enamel knot). We used 6 enamel knot markers with a large pattern of expression obtained from bite-it database (Fgf4, Slit1, Wnt3, Wnt10a, Dkk4 and Fgf20). Epithelium and mesenchyme markers were defined separately for mouse and hamster, based on our pure-tissue RNAseq data. We selected genes for with significant difference in expression levels between mesenchyme and epithelium (adjusted p-value < 0.1) with a log fold change greater than 5. The result does not change depending on the threshold used (3x up to 7x). The threshold used was set up to guarantee enough markers relatively equilibrated among the 3 tissues, and yet that are specific enough for the tissue under study. To get an idea of the sensitivity of the results to the choice of the set of markers, for each analysis, we computed 500 bootstrapped values by randomly resampling half of the epithelial and mesenchymal markers and recomputing deconvolutions on these 500 random subsets. The average of these resamplings are indicated in the barplots (Sup figure 8).

Modeling the timing of cusp patterning

We modeled the timing of cusp patterning as a continuous time Markovian process, with successive states the numbers of cusps, from 0 to the final number. Each state N has a specific rate of exit q_N, to the next state $N+1$. It means that sojourn time in N follows an exponential distribution of mean $\frac{1}{q_N}$. The final state is the absorbing state. From this,

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

we build the generator Q of the Markovian process of succession between states. The probability to observe a number n of cusps at time t is the value of transition probability matrix e^{Qt} at coordinates (0,n). If all states have the same mean sojourn time, all parameters q_N are equal, and the model has only one parameter. On the contrary, when all q_N are independent, there are as many parameters as the final number of cusps.

Concretely, the input data corresponds to N=141/153 teeth sampled for upper/lower molar in mouse, and 87/86 teeth sampled for upper/lower molar in hamster. On each tooth, patterned cusps were counted on Fgf4 in situ hybridizations of the dissociated epithelium. The development time was estimated in each sample through the weight of the embryo (in milligrams). In hamster, the relationship is the following:

$$Dev.Time.Hamster = \frac{weight^{0.23} + 5.3791}{0.7426}$$

The best transformation was obtained by a boxcox procedure, but is essentially similar to squared root of the weight. In mouse, the relation is the following :

$$Dev.Time.Mouse = \frac{weight^{0.23} + 2.4141}{0.4192}$$

Statistical analysis

Statistical analysis was performed using R (version 3.2.2, 2015-08-14), with ggplot2 for graphics ($ggplot2_1$.0.0). The data and R scripts, which permit to reproduce the figures and tests presented here, are provided in supplementary material XX.

References

[1] Arhat Abzhanov. Von Baer's law for the ages: Lost and found principles of developmental evolution. *Trends in Genetics*, 29(12):712–722, dec 2013.

[2] Leon Anavy, Michal Levin, Sally Khair, Nagayasu Nakanishi, Selene L Fernandez-Valverde, Bernard M Degnan, and Itai Yanai. BLIND ordering of large-scale transcriptomic developmental timecourses. *Development (Cambridge, England)*, 141(5):1161–6, mar 2014.

[3] Nuno L. Barbosa-Morais, Manuel Irimia, Qun Pan, Hui Y. Xiong, Serge Gueroussov, Leo J. Lee, Valentina Slobodeniuc, Claudia Kutter, Stephen Watt, Recep Çolak, TaeHyung Kim, Christine M. Misquitta-Ali, Michael D. Wilson, Philip M. Kim, Duncan T. Odom, Brendan J. Frey, and Benjamin J. Blencowe. The Evolutionary Landscape of Alternative Splicing in Vertebrate Species. *Science*, 338(6114), 2012.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[4] Marianna Bei. Molecular Genetics of Tooth Development. *Curr Opin Genet Dev*, 19(5):504–510, 2009.

[5] David Brawand, Magali Soumillon, Anamaria Necsulea, Philippe Julien, Gábor Csárdi, Patrick Harrigan, Manuela Weier, Angélica Liechti, Ayinuer Aximu-Petri, Martin Kircher, Frank W Albert, Ulrich Zeller, Philipp Khaitovich, Frank Grützner, Sven Bergmann, Rasmus Nielsen, Svante Pääbo, and Henrik Kaessmann. The evolution of gene expression levels in mammalian organs. *Nature*, 478(7369):343–8, oct 2011.

[6] Alessandra Breschi, Sarah Djebali, Jesse Gillis, Dmitri D. Pervouchine, Alex Dobin, Carrie A. Davis, Thomas R. Gingeras, and Roderic Guigó. Gene-specific patterns of expression variation across organs and species. *Genome Biology*, 17(1):151, dec 2016.

[7] Esther T Chan, Gerald T Quon, Gordon Chua, Tomas Babak, Miles Trochesset, Ralph A Zirngibl, Jane Aubin, Michael J H Ratcliffe, Andrew Wilde, Michael Brudno, Quaid D Morris, and Timothy R Hughes. Conservation of core gene expression in vertebrate tissues. *Journal of biology*, 8(3):33, 2009.

[8] Stéphane Dray and A. B. Dufour. The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*, 22(4):1 – 20, 2007.

[9] Renaud Gaujoux and Cathal Seoighe. CellMix: A comprehensive toolbox for gene expression deconvolution. *Bioinformatics*, 29(17):2211–2212, 2013.

[10] Yoav Gilad, Orna Mizrahi-Man, Yoav Gilad, and Orna Mizrahi-Man. A reanalysis of mouse ENCODE comparative gene expression data. *F1000Research*, 4, may 2015.

[11] Eric S. Haag. The Same but Different: Worms Reveal the Pervasiveness of Developmental System Drift. *PLoS Genetics*, 10(2):1–3, 2014.

[12] Eric S. Haag, Helen Chamberlin, Avril Coghlan, David H A Fitch, Andrew D. Peters, and Hinrich Schulenburg. Caenorhabditis evolution: if they all look alike, you aren't looking hard enough, mar 2007.

[13] Enni Harjunmaa, Kerstin Seidel, Teemu Häkkinen, Elodie Renvoisé, Ian J Corfe, Aki Kallonen, Zhao-Qun Zhang, Alistair R Evans, Marja L Mikkola, Isaac Salazar-Ciudad, Ophir D Klein, and Jukka Jernvall. Replaying evolutionary transitions from the dental fossil record. *Nature*, 512(7512):44–8, aug 2014.

[14] Jennifer W Israel, Megan L Martik, Maria Byrne, Elizabeth C Raff, Rudolf A Raff, David R McClay, and Gregory A Wray. Comparative Developmental Transcriptomics Reveals Rewiring of a Highly Conserved Gene Regulatory Network during

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

a Major Life History Switch in the Sea Urchin Genus Heliocidaris. *PLoS biology*, 14(3):e1002391, 2016.

[15] Jukka Jernvall and Isaac Salazar-Ciudad. The Economy of Tinkering Mammalian Teeth. In *Tinkering: The Microevolution of Development*, pages 207–224. wiley, jun 2007.

[16] Jukka Jernvall, Irma Thesleff, Y. Ahn, B. W. Sanderson, O. D. Klein, and Et Al. Tooth shape formation and tooth renewal: evolving with the same signals. *Development (Cambridge, England)*, 139(19):3487–97, 2012.

[17] Karin Kiontke, Antoine Barrière, Irina Kolotuev, Benjamin Podbilewicz, Ralf Sommer, David H A Fitch, and Marie Anne Félix. Trends, Stasis, and Drift in the Evolution of Nematode Vulva Development. *Current Biology*, 17(22):1925–1937, nov 2007.

[18] Vincent Lazzari, Cyril Charles, Paul Tafforeau, Monique Vianey-Liaud, Jean Pierre Aguilar, Jean Jacques Jaeger, Jacques Michaux, and Laurent Viriot. Mosaic convergence of rodent dentitions. *PLoS ONE*, 3(10), oct 2008.

[19] Michal Levin, Leon Anavy, Alison G Cole, Eitan Winter, Natalia Mostov, Sally Khair, Naftalie Senderovich, Ekaterina Kovalev, David H Silver, Martin Feder, Selene L Fernandez-Valverde, Nagayasu Nakanishi, David Simmons, Oleg Simakov, Tomas Larsson, Shang-Yun Liu, Ayelet Jerafi-Vider, Karina Yaniv, Joseph F Ryan, Mark Q Martindale, Jochen C Rink, Detlev Arendt, Sandie M Degnan, Bernard M Degnan, Tamar Hashimshony, and Itai Yanai. The mid-developmental transition and the evolution of animal body plans. *Nature*, 531(7596):637–41, mar 2016.

[20] Gong-Hua Lin, Kun Wang, Xiao-Gong Deng, Eviatar Nevo, Fang Zhao, Jian-Ping Su, Song-Chang Guo, Tong-Zuo Zhang, and Huabin Zhao. Transcriptome sequencing and phylogenomic resolution within Spalacidae (Rodentia). *BMC genomics*, 15:32, jan 2014.

[21] Michael I Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12):550, 2014.

[22] Jason Merkin, Caitlin Russell, Ping Chen, and Christopher B Burge. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science (New York, N.Y.)*, 338(6114):1593–9, dec 2012.

[23] Mihaela Pavlicev, Günter P Wagner, T.F. Mackay, et Al., F.W. Stearns, D.J. Kliebenstein, G.P. Wagner, J. Zhang, A.E. Tsong, et Al., R.R. Ackermann, J.M. Cheverud, N.M. Young, et Al., N.M. Young, B. Hallgrimsson, R. Berg, A.R. Palmer, T.A.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Williams, L.M. Nagy, J.M. Cheverud, et Al., M. Pavlicev, et Al., L.J. Leamy, et Al., M. Pavlicev, et Al., M. Pavlicev, et Al., L.F. Delph, et Al., A. Monteiro, et Al., H.S. Kim, J.C. Fay, H.S. Kim, et Al., S. Wright, R.A. Fisher, M. Baatz, G.P. Wagner, R.E. Lenski, S.J. Schrag, et Al., C. Chevillon, et Al., S. Maisnier-Patin, D.I. Andersson, A. Bossi, B. Lehner, Z. Wang, et Al., G.P. Wagner, et Al., A.Y. Albert, et Al., J.P. Kenney-Hunt, J.M. Cheverud, D.L. Stern, V. Orgogozo, J.R. True, E.S. Haag, E.S. Haag, N.A. Johnson, A.H. Porter, G.P. Wagner, T.F. Hansen, et Al., T.F. Hansen, D. Houle, K.R. Takahasi, et Al., M.Z. Ludwig, et Al., M.Z. Ludwig, et Al., M.T. Weirauch, T.R. Hughes, B.Z. He, et Al., C.R. Landry, et Al., A.S. Kondrashov, S.A. Shabalina, R.J. Kulathinal, R.S. Singh, D. Kuo, et Al., D.S. Falconer, T.M. Williams, S.B. Carroll, O. Pellmyr, H.W. Krenn, M.A. Bell, et Al., L.V. Roth, M.D. Shapiro, et Al., M.C. Brandley, et Al., T. Kohlsdorf, et Al., C.D. Siler, R.M. Brown, T. Kohlsdorf, G.P. Wagner, Q. Zhong, et Al., M. Dreze, et Al., B. Charloteaux, et Al., R.D. Barrett, et Al., G.P. Wagner, et Al., R.E. Lenski, T. Flatt, T.J. Kawecki, K.M. Pepin, et Al., S. Li, et Al., and N. Blow. A model of developmental evolution: selection, pleiotropy and compensation. *Trends in ecology & evolution*, 27(6):316–22, jun 2012.

[24] Sabrina Renaud, Mouloud Benammi, and Jean-Jacques Jaeger. Morphological evolution of the murine rodent Paraethomys in response to climatic variations (Mio-Pleistocene of North Africa). *Paleobiology*, 25(3):369–382, 1999.

[25] Sabrina Renaud, Jacques Michaux, Jean-Jacques Jaeger, and Jean-Christophe Auffray. Fourier analysis applied to Stephanomys (Rodentia, Muridae) molars: nonprogressive evolutionary pattern in a gradual lineage. *Paleobiology*, 22(2):255–265, 1996.

[26] Sabrina Renaud, Sophie Pantalacci, Jean Pierre Quéré, Vincent Laudet, and Jean Christophe Auffray. Developmental constraints revealed by co-variation within and among molar rows in two murine rodents. *Evolution and Development*, 11(5):590–602, sep 2009.

[27] Julien Roux, Marta Rosikiewicz, and Marc Robinson-Rechavi. What to compare and how: Comparative transcriptomics for Evo-Devo. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 324(4):372–382, jun 2015.

[28] I Salazar-Ciudad and J Jernvall. A gene network model accounting for development and evolution of mammalian teeth. *Proc Natl Acad Sci U S A*, 99(12):8116–8120, 2002.

[29] Isaac Salazar-Ciudad. Tooth patterning and evolution. *Current opinion in genetics & development*, 22(6):585–92, dec 2012.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[30] Isaac Salazar-Ciudad and Jukka Jernvall. How different types of pattern formation mechanisms affect the evolution of form and development. *Evolution and Development*, 6(1):6–16, jan 2004.

[31] Isaac Salazar-Ciudad and Jukka Jernvall. A computational model of teeth and the developmental origins of morphological variation. *Nature*, 464(7288):583–6, mar 2010.

[32] John R. True and Eric S. Haag. Developmental system drift and flexibility in evolutionary trajectories, 2001.

[33] S. Urdy, N. Goudemand, and S. Pantalacci. Looking beyond the genes: The interplay between signaling pathways and mechanics in the shaping and diversification of epithelial tissues. In *Current Topics in Developmental Biology*, volume 119, pages 227–290. Academic Press Inc., 2016.

[34] L M Van Valen. Serial homology : the crests and cusps of mammalian teeth. *Acta Palaeontologica Polonica*, 38(3/4):145–158, 1994.

[35] Günter P Wagner. The developmental genetics of homology. *Nature reviews. Genetics*, 8(6):473–9, jun 2007.

[36] Karl R. Wotton, Eva Jimenez-Guri, Anton Crombach, Hilde Janssens, Anna Alcaine-Colet, Steffen Lemke, Urs Schmidt-Ott, and Johannes Jaeger. Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly megaselia abdita. *eLife*, 2015(4), jan 2015.

[37] Itai Yanai, Dan Graur, and Ron Ophir. Incongruent Expression Profiles between Human and Mouse Orthologous Genes Suggest Widespread Neutral Evolution of Transcription Control. *OMICS: A Journal of Integrative Biology*, 8(1):15–24, jan 2004.

[38] Itai Yanai, Leonid Peshkin, Paul Jorgensen, and Marc W. Kirschner. Mapping Gene Expression in Two Xenopus Species: Evolutionary Constraints and Developmental Flexibility. *Developmental Cell*, 20(4):483–496, 2011.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary figures

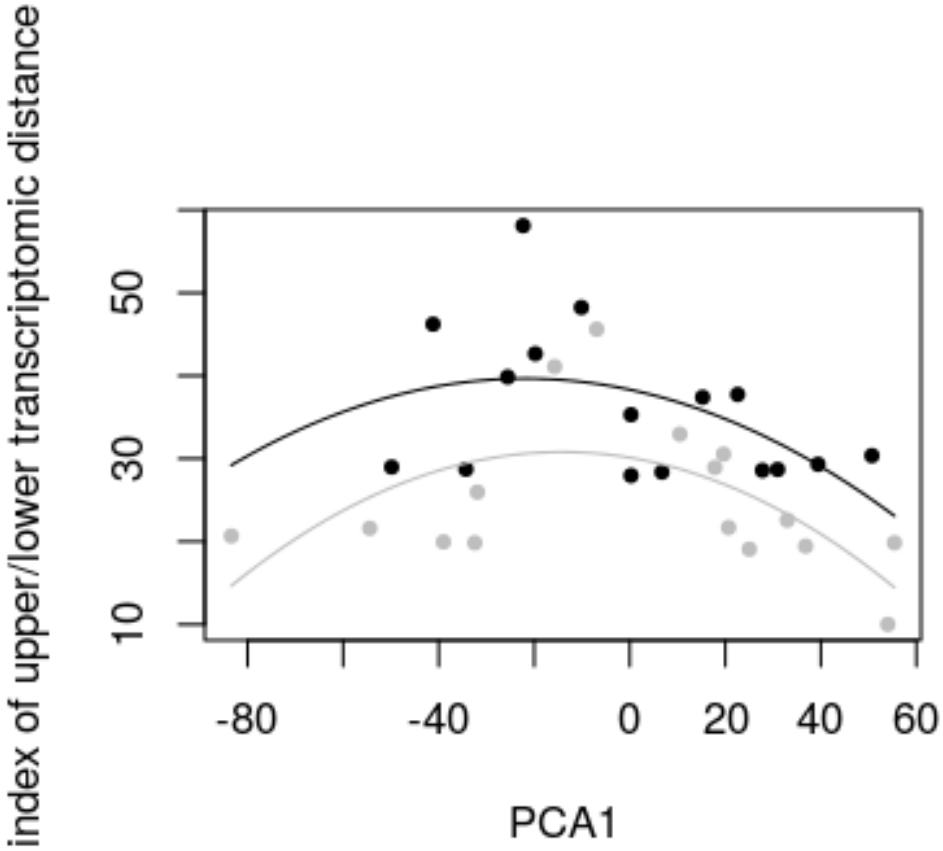

Figure 7. The distance between upper and lower transcriptomes peaks in early morphogenesis. Between multivariate analysis was used to extract an axis of variation associated to the difference between upper and lower tooth in the whole dataset. The distance between upper and lower samples from the same embryo on this axis is represented on the y-axis (in black for mouse embryos, in grey for hamster embryos). The x-axis is the second axis of the PCA (as in Figure 5-A) which is a developmental clock. The curves represent the best polynomial fit of this data (R2=0.51; quadratic term of the relation between y and x, p=0.00422). The index of upper/lower transcriptomic distance is significantly higher in mouse (p=0.0044) than in hamster.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

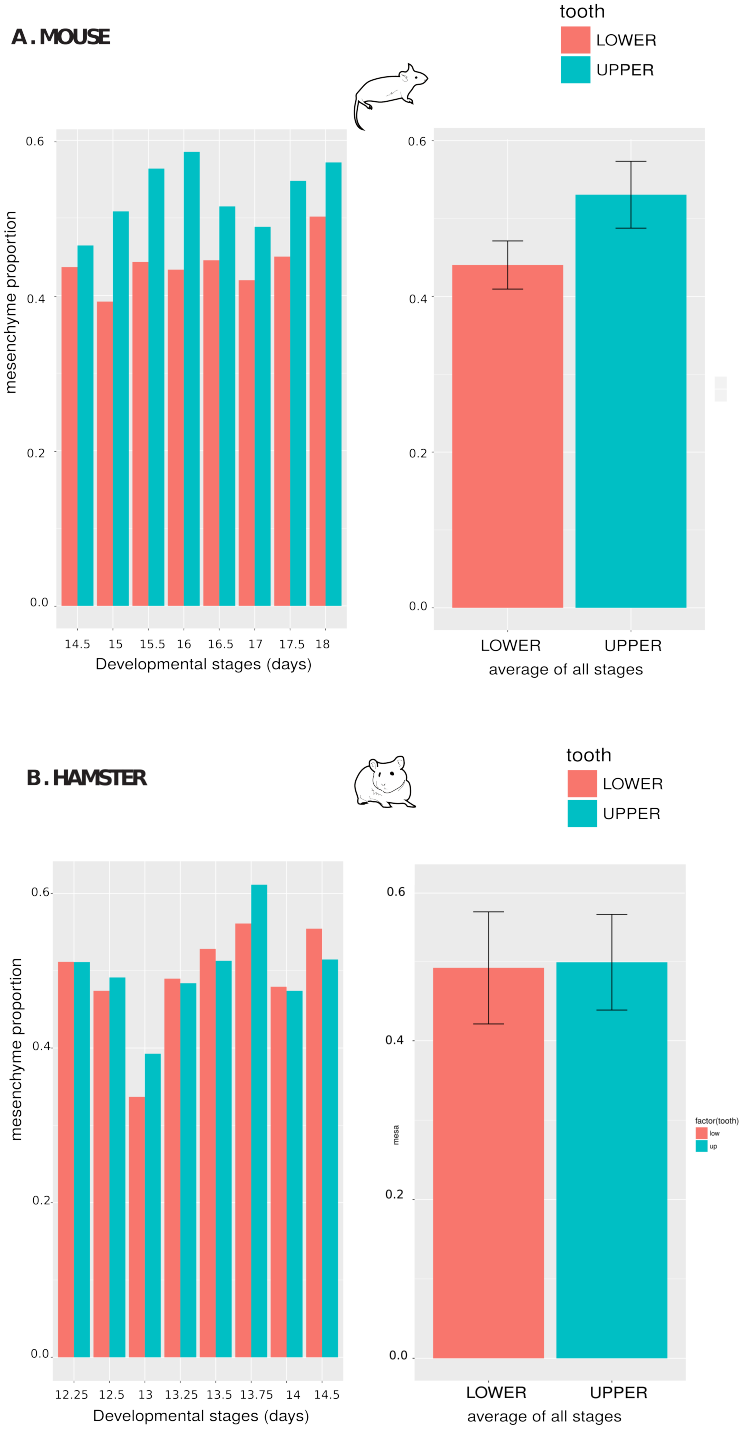

Figure 8. Mesenchyme proportions in mouse (A) and hamster (B) estimated from deconvolutions for 8 developmental stages. Markers for enamel knot were obtained from bite-it database (6 markers). Markers for mesenchyme and epithelium were obtained separately for each species from pure tissue transcriptomes. Left: The median proportion from 500 resampling of half of the markers is indicated for each time point. Right: the average of the time points is indicated.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.22.043422; this version posted April 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

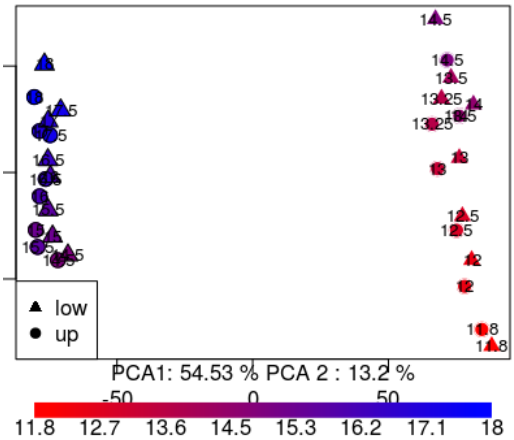

Figure 9. Principal component analysis with 32 samples from both species and tooth types. The location of the samples on the 2 first principal components is shown. All samples come from the same run of sequencing, and the libraries were prepared together. The mouse samples (left) and the hamster samples (right) are clearly separated. The upper/lower samples (respectively circles and triangles) are not separated on this map.