A Quantitative Method for Staging Mouse Embryos based on Limb Morphometry

M. Musy[1,2,3*], K. Flaherty[4*], J. Raspopovic[1,2], A. Robert-Moreno[1,2,3], J. Richtsmeier[4], J. Sharpe[1,2,3,5]

[1]*Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology Dr. Aiguader 88, Barcelona 08003, Spain*

[2]*Universitat Pompeu Fabra (UPF), Barcelona, Spain*

[3]*EMBL Barcelona, Dr. Aiguader 88, Barcelona 08003, Spain*

[4]*Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania, USA*

[5]*Institució Catalana de Recerca i Estudis Avançats (ICREA) Passeig Lluís Companys 23, 08010 Barcelona, Spain*

Abstract

To determine the developmental stage of embryonic mice we apply a geometric morphometric approach to the changing shape of the mouse limb bud as it grows from embryonic day 10 to embryonic day 15 post conception. As the ontogenetic sequence results in the *de novo* emergence of shape features not present in the early stages, we have created a standard ontogenetic trajectory for limb bud development -- a quantitative characterization of shape change during limb morphogenesis. This trajectory of form as a function of time also gives us the reverse function: the ability to infer developmental stage from form, with a typical uncertainty of 2 hours. We introduce *eMOSS* (embryonic Mouse Ontogenetic Staging System, freely available at `limbstaging.embl.es`) as a fast, reliable and convenient on-line tool for staging embryos from 2D images of their limb buds, and illustrate its use in phenotyping early limb abnormalities.

* These authors contributed equally to this work.

1 Introduction

D'Arcy Thompson illustrated through his work how geometric shapes of biological forms could be directly invoked by simple mathematical equations that might reflect the underlying growth process. He compared shapes of different species or strains, rather than shapes at different developmental time-points and noted that the study of growth simply adds the dimension of time to that of spatial form [1]. However, analysis of a developmental sequence can pose a fundamental challenge because as immature shapes progress to fully developed ones, genuinely new forms and novel features emerge.

A myriad of different developmental processes are coordinated to achieve typical morphogenesis, including: gene transcription and expression, geometric and mechanical three-dimensional organization of proteins and cells, mechano-transductive forces of and on cells and tissues, and extracellular matrix content and structure. Understanding how they fit together in space and time calls for a rigorous and quantitative standardized sequence for the typically developing wild-type morphologies, calibrated to real time.

To understand process and determine mechanism we need precise information about developmental timing. The age of an embryo is typically described using one of two metrics: harvesting age or developmental stage [2, 3]. Harvesting age, the age at sacrifice, measures the time elapsed between conception and collection of an embryo based on timed matings. It is affected by a large uncertainty because the moment of fertilization is routinely only roughly estimated by observation of the vaginal plug. In practice, harvesting age is the same for each embryo in a litter. By contrast, developmental stage - or staging age - is a relative measure of the amount of progress that an individual embryo has made along its ontogenetic trajectory, and so describes the morphological maturity of an embryo

relative to a typically developing individual. An additional source of uncertainty results from the fact that embryos in the same litter can develop at different speeds [4, 5]. The lack of synchrony between harvesting age and developmental stage represents a confounding factor when comparing and combining the results of experiments that use embryonic mice produced by timed matings. This is particularly problematic when comparisons are made between different strains of inbred mice as quantitative and qualitative morphological analyses have demonstrated that different strains of inbred mice develop at different rates [6, 7].

Harvesting age is routinely used because it is a simple metric that is rather effortless to apply in practice. Because differences in developmental progress exist among embryos of any litter, the use of harvesting age introduces a source of variation into the statistical samples used for analysis. This variation can affect our understanding of the timing and sequence of important developmental processes and events (e.g., initiation of cell differentiation, bone mineralization, or determining whether ligand expression occurs independently or in concert with other markers).

Developmental stage can be described for murine embryos through the definition of a staging system. The most widely used is the Theiler system [8] that enables researchers to stage an embryo to a particular embryonic day by observing the presence of suites of external features (e.g., embryonic day 13 is marked by ear pinna, five rows of whiskers, prominent hair follicles above the eye and in front of the ear, marked indentation of the hand plate, incipient indentation of the footplate [8]). Though adequate for many descriptive investigations, the vast majority of embryos collected on a particular day might display all of the named features. Consequently, this system is too coarse for experiments designed to query the dynamics of gene expression and tissue interactions that require temporal precision finer than an embryonic day. The staging atlases of Theiler and Kaufman [8, 9]

consider the developmental progress of multiple systems throughout all of development, but for some developmental intervals its application requires time intensive examination, dissection, and/or histological study of diverse anatomical and physiological systems.

As the principles of developmental biology are now integrated into a series of new disciplines across biology [10], a simplified method for staging embryos is needed. Such a system enables scientists using techniques from developmental biology in their studies of cell biology, genetics, immunology, or neurobiology but who may not have the expertise or knowledge to gather and combine information from multiple tissue and organ systems, to effectively assign a developmental stage to each embryo studied.

To acquire knowledge of mechanisms operating during the rapid development of murine embryos, a staging system with the following characteristics is required:

1) the ability to measure developmental stage in embryonic *hours* as opposed to embryonic *days*,

2) the ability to be easily integrated into research projects in multiple fields in developmental biology without requiring fine-scale knowledge of different anatomic structures,

3) the ability to rapidly obtain an estimation of developmental time without performing any complicated analysis or technical procedure on the embryo.

Here we propose such a staging system, called the Embryonic Mouse Ontogenetic Staging System, or eMOSS. This system uses the developing hindlimb and a method of pattern recognition that relies on the *maximum likelihood principle*, obviating the need to use standardized *a priori* landmarks or morphological features. Using eMOSS we are able to describe the ontogenetic trajectory of the two dimensional outline of the hindlimb throughout seven days of

embryological development from about 10 embryonic days to 15 embryonic days post conception.

We have chosen to focus on the hindlimb instead of the forelimb because the forelimb is more developmentally advanced than the hindlimb during the temporal interval considered, allowing the hindlimb to be used for staging to an older age. Before E10 the limb structures are not yet visible, while changes in limb shape after E15 are less marked compared to the change in absolute size. After E15, staging becomes difficult due to the complexity of the shape of the autopod and the tendency for digits to curl toward the plantar surface of the foot after fixation due to the strength of the plantar flexors relative to the dorsiflexors. The eMOSS system provides a direct and high-resolution method for rapid determination of the precise developmental stage of embryos requiring only an image of the hindlimb without invasive examination of embryo anatomy.

2 Results

We previously reported a staging system which used a subset of the current dataset, but employed a different mathematical approach for morphometric comparisons based on the curvature of the outline, rather than the 2D coordinates of the outline [6]. Here, we use a method related to shape transforms in Cartesian coordinates - we align our dataset in three dimensional space, we establish mean shapes for the 12 subsets of known harvesting time, and interpolate between them in the temporal dimension. The assumption of a normal distribution of the numerical data enables the estimation of the error on limb stage from the observed χ_v^2 (see Eq. 1 in Material and Methods). Consequently, we are able to describe the ontogenetic trajectory of the two dimensional outline of the hindlimb throughout five days of embryological development.

In order to distinguish results generated by our staging methodology from the common denotation for the age at harvest of a litter, we have adopted the

following convention from [6]: results from eMOSS are denoted in the form *mE(day):(hours)*, with *mE* standing for *morphometric Embryonic day*. This is distinguished from the age at harvest, which is denoted in the form E(day).(fraction of day elapsed since midnight). Thus, an embryo collected at 12:00PM on embryonic day 13 would have an age at harvest of E13.5, and if the eMOSS staging result is embryonic day 13 and 12 hours old, the estimated stage will be *mE13:12*. Embryos of a certain harvesting age can have a developmental stage estimated by eMOSS that corresponds directly with the harvesting age or is relatively younger or older than the harvesting age. In the present work we focus on the range of stages between *mE10:09* and *mE15:21*.

The data-set for this morphometric analysis consists of 864 images of mouse hindlimb buds ranging from E10 to E15, of which the younger half (E10 to E12) had already been used in a previous study [6], and the older half (E12 to E15) is newly-generated for the current study. Variation in developmental stage, even for embryos within the same litter, is large (up to ± 12 hours) [6], however by harvesting embryos at fixed times of the day, 12 hours apart, it was possible to obtain enough embryos to estimate a reliable average shape for each time-point. The dorsal surface of all limb buds were photographed, and the outline of the hindlimb was manually collected (see Material and Methods for more details). Then, for each of the 12-hour time-points an average limb outline was estimated by performing multiple alignments. Figure 1A shows the data for the sample of limb buds at each time-point and the average calculated shape (primary shape). The emergence of the digits (Figure 1A) can be seen at about *mE12:09*, while stages *mE11:09* and *mE10:21* show completely smooth distal outlines.

The subsequent step is the interpolation of shapes that are intermediate between the primary shapes. For each 12-hour interval we determined equivalence points between the younger and older primary shapes, and then performed an interpolation to determine the intermediate shapes. An important consideration is that equivalence points do not have to represent the same piece of physical tissue over time - i.e. they are not necessarily homologous, and do not have to be equivalent to fate-maps. The goal of the procedure is only to create the correct intermediate shapes which are defined by a continuous outline boundary, rather than by specific points. The final limb bud growth trajectory (whose creation is described in greater detail in the Methods section) is shown in Figure 1B.

From this standardized limb trajectory we built the on-line tool eMOSS (Figure 2). To use eMOSS, limb bud image files are uploaded to the system, and the user is given convenient tools to easily select a set of points along each contour to identify its shape. The system then compares the new shape to all time-points in the standard trajectory, and the result of the fitting procedure provides an estimate of developmental stage for the limb bud. The gray points on the limb outline shown at left of Figure 2 are the positions identified by the experimenter along the contour of the limb shape where the shading represents the Gaussian uncertainty in the determination. The black line in this plot corresponds to the reference shape that minimizes the log-likelihood term in Eq. 1.

An important indicator of the goodness of the fit is given by its value at the minimum. Figure 3A shows the 2-dimensional distribution of χ_v^2 for the complete data-set as a function of time. A high χ_v^2 value could indicate that the algorithm is unable to find a reference shape that can fit all the measured points because either the user did not accurately select them, or because there is genuinely no shape that

resembles the limb outline at that level of accuracy. In such cases, the user should evaluate the situation on a case by case basis, by judging for example if the fitted reference shape in the left plot still corresponds to a meaningful fit to the points (see left of Figure 2). Indeed, in cases where a very large number of points is drawn (e.g. $n > 50$), one is likely to obtain a higher χ_v^2 value because the fit will become sensitive enough to discriminate systematic differences between the staged limb and the set of reference limbs used in the database for comparison.

The absolute scale of magnification for each image can be reconstructed using meta-data contained within the picture, including the pixel dimensions. This information could in principle be used to further constrain the age determination, but in our case it can be safely disregarded if it is not known by the experimenter, which is often likely to be the case. Figure 4 shows the distribution of the dataset of 864 limbs stages for the 12 different age groups. The inset of Figure 4 shows how stages of limbs of individuals from the same harvesting age show large variation in developmental maturity.

Knowledge of which genes are being expressed by cell and tissue types during development and the duration of that expression is critical to understand and to model the genetic, cellular, and biomechanical interactions that govern rapid developmental processes. Figure 5 shows the pattern of expression of the *Osterix* (*Osx*) gene [11] in the heads of typically developing *Osx*-GFP mice at various harvesting times as demonstration of the ability of the staging system to parse through litter effects and individual variation in the rate of development and create a coherent sequence of developmental events. The top row corresponds to the ordering of these specimens according to harvesting time, showing that the individual variation in the rate of development can create a confusing temporal reconstruction of the spread of *Osx* expression (in the frontal and parietal ossification centers) which should increase in area as the embryo matures. The

bottom sequence of images, ordered according to estimates derived from the staging system, shows that eMOSS produces a sequence of staged embryos that is coherent with a localized expansion of *Osx* gene expression regardless of the age at harvest.

2.1 Experimental uncertainties and left-right asymmetry

We wished to test the distribution of errors and observe its dependence on the developmental stage. Figure 3B shows the distribution of the errors estimated by our algorithm as a function of the staging time. A clear result is a consistently larger error estimation from approximately 300 *h* (*mE12:12*) to 320 *h* (*mE13:08*). This larger uncertainty derives directly from the shapes in that age interval being more similar to each other - that makes it harder to distinguish shapes across ages. Consequently, this graph serves as an inverse measure of the rate of shape change. Examination of the shapes in Figure 1A confirms that between *mE12:09* and *mE13:09* the shape does not change rapidly.

A reliable evaluation of errors is often difficult in a fit to several unknown variables. For validation we can use the fact that the harvesting age of the left limb a_L, and right limb a_R, of the same embryo are by definition the same. To cross-check the uncertainties estimated by our algorithm, we can compare the spread of the distribution of the difference $d_{LR} = a_L - a_R$ with our calculation. Figure 3C shows the distribution of d_{LR} as a function of the average $(a_L + a_R)/2$. If σ_L and σ_R are the estimated uncertainties on a_L and a_R, then the error on the difference d_{LR} is simply given by $\sigma_d = \sqrt{\sigma_L^2 + \sigma_R^2}$. Figures 3D1 and 3D2 show the distributions of the quantity d_{LR}/σ_d in the two stage ranges below and above 290 *h* (*mE12:02*) respectively. If our estimation of the uncertainty is correct their variance must be close to one, where in fact we find a value of $\sigma = 1.1$ in the whole range. Splitting the sample in the three stage bins [249 *h*, 290 *h*], [290 *h*, 381 *h*] and [249 *h*, 381 *h*]

gives us $\sigma = 1.2$, 0.9, 1.1 respectively. Interestingly, for young stages (up to 290 h) a small yet statistically significant ($p < 10^{-6}$) difference of $d_{LR} = -0.7\,h$ is observed, meaning that the left limb bud is on average only slightly less developmentally advanced than the right limb buds in this range of stages.

2.2 Correlation between two different developmental structures: limbs and somites

Somite counting is a widely used method of measuring developmental stage, though the usefulness of somites is limited to earlier ages as they begin to disappear around E13.75. In addition to the dataset used to establish the ontogenetic trajectory of embryonic limb development we also generated an independent dataset of 50 embryos for which we measured the number of somites, in addition to staging both the forelimbs and hindlimbs. Figure 6A shows the linear relation of developmental stage with number of somites within an embryo. The blue dots with large error bars are obtained from Ref. [12]. While there is good agreement with the linear regression for embryos younger than $\approx 290\,h$ (*mE12:02*), a deviation from linearity is observed for older ages that is not observed in our measurements. In Figure 6B we also show a regression analysis of hindlimbs against forelimbs. This result allows the use of the forelimb in place of the hindlimb to stage an embryo if the embryo is younger than $\approx 330\,h$ (*mE13:18*).

2.3 Determining the onset of dysmorphogenesis in mutant phenotypes

Although eMOSS has been designed to stage wild type mice, another possible use is to apply the staging algorithm mice with hindlimb morphology that deviates from the typical developmental trajectory - not for the purpose of staging the limbs, but instead to identify and quantify the morphological deviation when it is still too subtle to be detected by visual inspection. This is relevant because complex events downstream of a genetic perturbation mean that many observed phenotypic changes can be the result of *secondary effects* of the primary etiology.

We staged embryos from a mutant named *Ssq*, in which the gene *Sonic Hedgehog* (*Shh*) is mis-regulated by insertion of a reporter construct directly into its limb-specific enhancer. The phenotype is characterized by a twisted and enlarged hind-foot morphology with polydactyly [13]. Hindlimb polydactyly was consistently more severe in homozygous mice.

In Figure 6C the limb buds from mutant (red) and wild-type (blue) mice are plotted, showing their deviation from the standard trajectory as a function of stage. While the older mutant limb buds are easily spotted by eye (upper insert), a few younger mutant limb buds are found with deviations from normality, but which look normal to the human eye (lower insert), thus showing the power of the morphometric approach to reveal subtle differences in phenotypes. In the graph, a linear regression of the abnormal points (dashed line) projects backwards to indicate that the phenotypic onset of the mutation should sit around 270 *h* (*mE11:06*).

3 Discussion

Variation in the rate of development among individuals creates a confounding factor in understanding morphogenesis that has previously been ignored due to the lack of an easy to use system that accounts for developmental variation with tight temporal precision. While previous staging systems with relatively low temporal resolution were adequate to address questions that focused mainly on obvious morphological changes, they are inadequate to assess fine scale and short duration gene and tissue interactions. Our method relies on data that describe the global shape of the object (limb) and does not rely on knowledge of the absolute scale of the object. It also goes beyond standard geometric morphometrics by incorporating the appearance of novel morphology over developmental time, thereby allowing for the dynamically changing complexity of biological shape. It tackles the challenges of a dynamically changing complexity of

shape during ontogeny by avoiding the need for identifying specific landmarks along the hindlimb.

We demonstrate that our algorithm is both stable and robust. A web-based version of the eMOSS algorithm is publicly available (at **http://limbstaging.embl.es**) [14] where users can upload images of mouse limbs, acquire the necessary input data, and obtain developmental stage estimates. This system is designed to be easily and rapidly implemented into current research protocols without requiring a significant investment in training or practice. We tried to keep the user experience of our system as simple as possible: users only need to upload a photographic picture in any common format, without any specific calibration (no need to know the pixel size or aspect ratio). Once uploaded it should take one or two minutes for the user to mark points on the outline of the limb, while the algorithm typically takes less than 5 seconds to estimate its developmental stage. While it is certainly possible to correlate the development of a specific mouse with variables related to different structures/organs than the limb, this requires the user to measure those features reliably which can be impractical and include unknown error from several sources, including instrument error and human bias.

By enabling researchers to accurately measure the difference in developmental age between specimens, eMOSS provides a remedy for the natural variability in developmental morphology between individuals.

4 Materials and Methods

4.1 Sample and outline collection

In this study we make use of a relatively large sample of C57Bl6/J mouse embryo hindlimbs harvested at fixed intervals of 12 hours during gestation. The full dataset consists of 864 images of mouse hindlimbs from embryos of known harvesting age that range from 249 hours post conception (10 days and 9 hours) to

381 hours (15 days and 21 hours). The fact that these time intervals are fixed and regular is not a necessary requirement for our method. In [6] we found that the growth trajectory of the shape of the hindlimb is similar in each of the different inbred strains, despite apparent differences in the rate of development.

Embryos were removed from their placenta and fixed in 4% PFA for 24 hours, then washed and preserved in 0.01% Sodium azide/PBS solution. Limbs were imaged using a QImaging Retiga 2000R monochromatic camera attached to a light microscope. During imaging, limbs were coated with PBS to prevent deformation due to drying. All mouse procedures employed in this research were approved by the Institutional Animal Care and Use Committee of the Pennsylvania State University (IACUC46558, IBC46590).

4.2 Staging system methodology

The morphological basis for the staging system is the two dimensional outline of the hindlimb, which is obtained by collecting 2D coordinate data from a series of points that describe the shape of the outline. This set of measurements of positions d_i are assumed to be distributed according to the normal distribution and they are compared to the expected positions μ_i based on the assumption of a specific age shape. The most probable shape is found by minimizing the quantity

$$\chi_v^2(\theta) = \frac{1}{v} \sum_{i=1}^{n} [\frac{d_i - \mu_i(\theta)}{\sigma_i}]^2, \tag{1}$$

with respect to the set of the unknown parameters indicated by the vector θ. In practice this is achieved numerically by presenting $\chi_v^2(\theta)$ to a minimization algorithm [15].

This quantity depends on the number of points and on the number of components of vector θ. In our case we want to fit two relative positions in x and y, a global scaling factor, and 3 Euler angles that represent an arbitrary rotation in space. This adds up to a 6-dimensional fit. Accounting for possible rotations in

space can help in the case that the image of the limb was not taken sufficiently orthogonal to the hand plate. To avoid the risk of introducing biases in the determination of the stage, the possible range of values for rotations is limited to ±0.25 *rad* in each of the three Euler angles.

The general strategy for constructing our reference groups is as follows:

1) The original dataset corresponding to a specific (harvesting) age group is initially aligned using the well known *Procrustes algorithm* [16]. In our case we have 12 such age groups, which range in age from 249 h (10 days and 9 h) to 381 h (15 days and 21 h) post-conception.

2) Within each age group, we iteratively align each limb shape with respect to the remaining ones minimizing the term in Eq. 1. Since each picture is taken from a different position, angle and magnification scale, the maximization of the likelihood will regard such parameters as free parameters in the fit (the x and y positions and one rotation angle). This process is repeated iteratively five times. Because the same harvesting age group can contain shapes that are at genuinely different morphological stages of development, at each iteration we discarded the 5% of shapes that deviate the most from the remaining sample of shapes.

3) After all the limb shapes corresponding to the same harvesting age group are superimposed, a convolution curve is estimated for each of the age groups (Figure 1A). The red curves in the plots are obtained by spline-interpolation in the two Cartesian coordinates separately.

4) Once our set of references has been determined for all harvesting age groups, we need to interpolate the shapes to obtain the intermediate stages to complete the developmental trajectory. A simple way to interpolate one curve into another is to associate points linearly along the two curves. This is justified when one can identify a sufficient number of landmark points to make the association. In our case this is applicable only for ages after E13 when one can visually recognize

the initial appearance of the fingers. We obtained such interpolation by splining the coordinates positions in the first 6 *principal components* as a function of time.

Figure 1B shows the result of the interpolation, where the red curves indicate the different age groups. We stress that such an interpolation between shapes is not aimed at describing the biological dynamics of tissue movement, but rather the geometrical correspondence between the outlines of two subsequent limb stages.

5 Acknowledgements

References

[1] Thompson, D. W., 1917. On Growth and Form. Cambridge University Press.

[2] Hall, B.K., and Miyake, T. (1995). How do embryos measure time. *Evolutionary Change and Heterochrony* 3 20.

[3] Klingenberg, C.P. (1998). Heterochrony and allometry: the analysis of evolutionary change in ontogeny. Biological Rev. 73, 79 123.

[4] Miyake T., Cameron A.M., and Hall B.K. (1995). Detailed staging of inbred C57BL/6 mice between Theiler's [1972] stages 18 and 21 (11-13 days of gestation) based on craniofacial development. Journal of Craniofacial Genetics and Dev. Biology 6, 1 31.

[5] Wanek N., Muneoka K., Holler–Dinsmore G., Burton R., and Bryant S.V. (1989). A staging system for mouse limb development. Journal of Experimental Zoology 249, 41 49.

[6] Boehm B., Rautschka M., Quintana L., Raspopovic J., Jan Z., and Sharpe

J. (2011). A landmark-free morphometric staging system for the mouse limb bud. Development 138, 1227 1234.

[7] Miyake, T., Cameron, A.M., and Hall, B.K. (1996). Variability of embryonic development among three inbred strains of mice. Growth, Development, and Aging: GDA 61, 141 155.

[8] Theiler K., (1989). The house mouse: atlas of embryonic development (Springer-Verlag New York).

[9] Kaufman, M.H., (1992). The Atlas of Mouse Development (Academic press London).

[10] Gilbert, Scott F. (2017). Developmental biology, the stem cell of biological disciplines. PLoS biology 15, e2003691.

[11] Rodda S.J. (2006).Distinct roles for Hedgehog and canonical Wnt signaling in specification, differentiation and maintenance of osteoblast progenitors. Development 133, 3231 3244.

[12] Richardson L., Venkataraman S., Stevenson P., Yang Y., Moss J., Graham L., Burton N., Hill B., Rao J., and Baldock R.A. (2014). EMAGE mouse embryo spatial gene expression database: 2014 update. Nucleic Acids Research 42, D835 D844. At the link:
`http://www.emouseatlas.org/emap/ema/theiler_stages/StageDefinition/stagedefinition.html`

[13] Sharpe J., Lettice L., Hecksher–Sorensen J., Fox M., Hill R., and Krumlauf R. (1999). Identification of sonic hedgehog as a candidate gene responsible for the polydactylous mouse mutant Sasquatch. Current Biology 9, 97-S1.

[14] The Embryonic Mouse Ontogenetic Staging System Online (eMOSS). Available at: http://limbstaging.embl.es

[15] Brun R. and Rademakers F., ROOT - An Object Oriented Data Analysis

Framework, Proc. AIHENP96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86.

[16] Kendall D.G. (1989). A survey of the statistical theory of shape. Statistical Science 87 99.
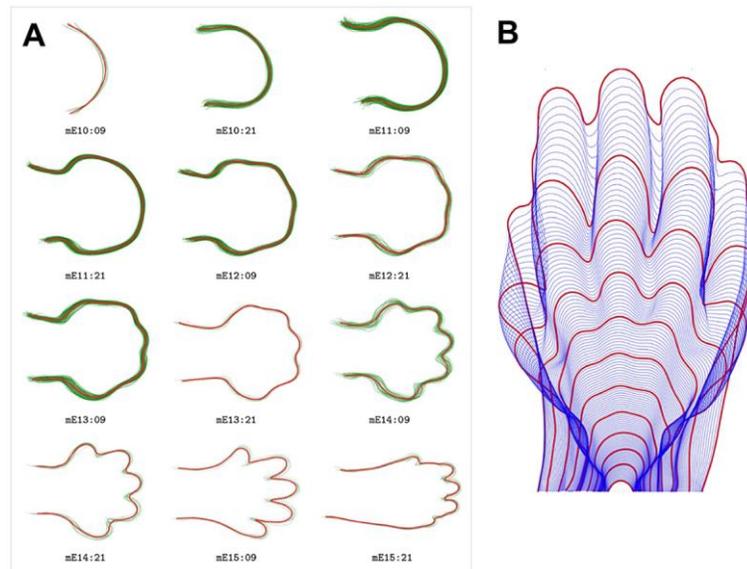
Figures

Figure 1: *The limb developmental trajectory.*

(A) Superposition of data points for the whole dataset discretized into equally spaced intervals of 12 hours, from 249 to 381 hours. The red lines represent the average shapes, for each of the harvesting age groups.

(B) Interpolation of one reference shape to the next in the time sequence. The red lines indicate the shapes that correspond to the experimental averages for each age group.
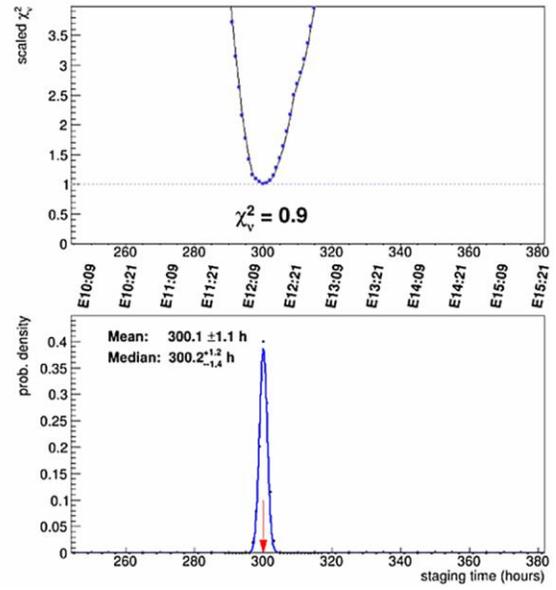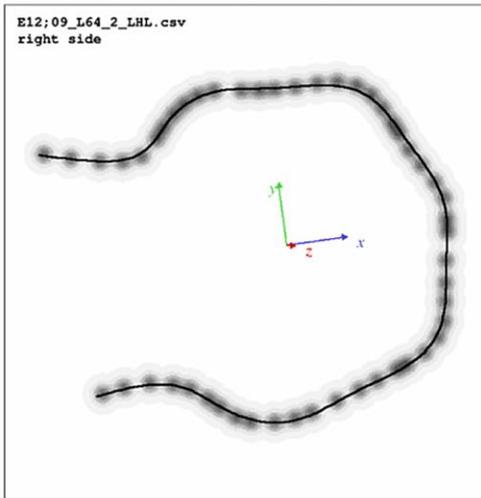
Figure 2: *The staging system.*

The result of the global fit of a limb as shown in the web application. The scaled χ_ν^2 curve (see Eq. 1 and the corresponding probability density function are shown in the right panel.
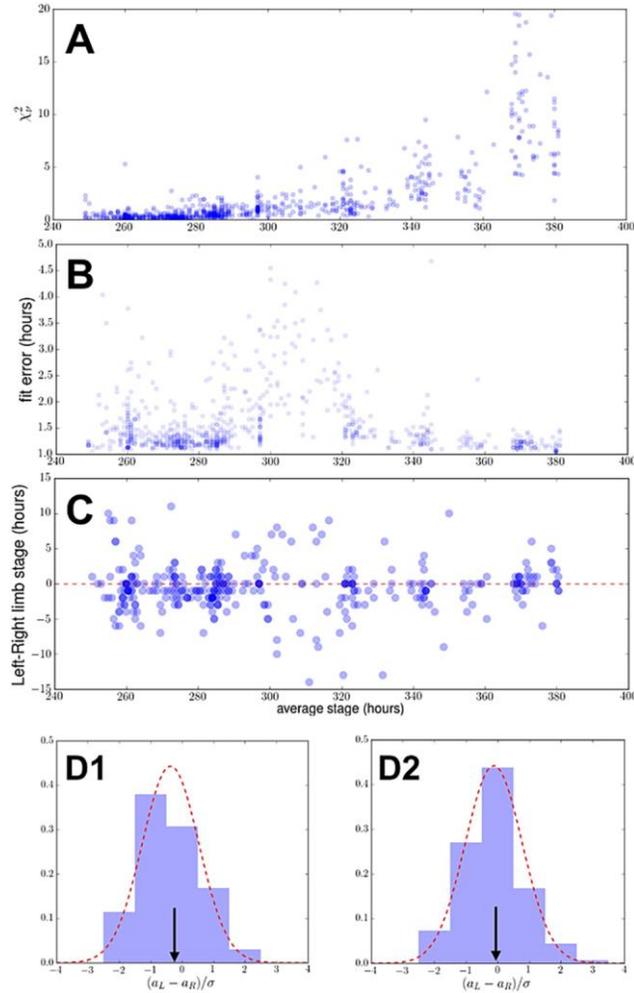
Figure 3: *Distributions of different fit variables.*

(A) Distribution of the fit score χ_ν^2, for the whole dataset of hindlimbs.

(B) Distribution of the estimated error as a function of the staging time.

(C) Distribution of differences of the stage estimates for the left and the right limbs ($d_{LR} = a_L - a_R$) as a function of the average staging time $(a_L + a_R)/2$.

(D1, D2) Distributions of d_{LR}/σ_d in the ranges below and above 290 h (*mE12:02*), where the arrow indicates the position of the Gaussian peaks (at -0.5 and -0.1 respectively).
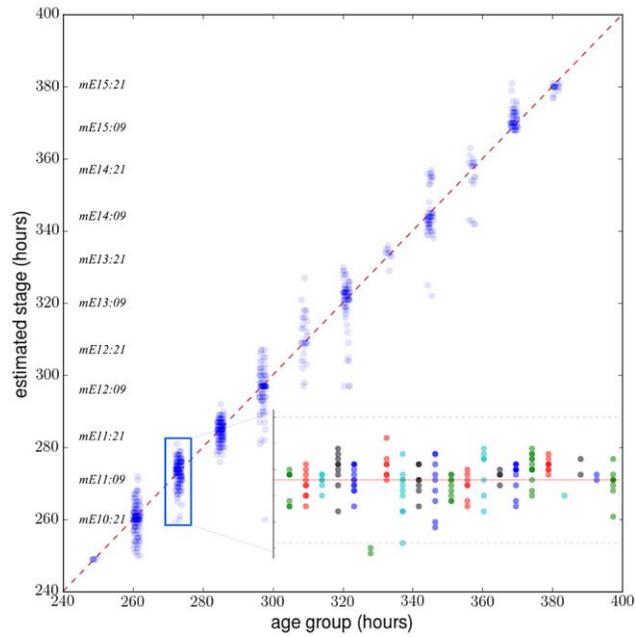
Figure 4: *Distributions of the the measured stages in the dataset.*

The distribution of the total 864 sample of limbs. The y-axis gives the estimated stage while the x-axis corresponds to the 12 harvesting age groups. The inset plot shows the estimated stage of limbs in the age group harvested at 273 h (11 days and 9 hours p.c.) for different litters along the horizontal axis. Each point represents the limb of one individual.
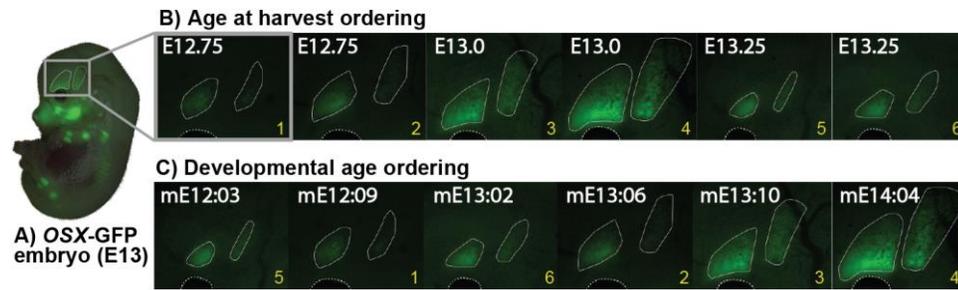
Figure 5: *Comparison of Osx-GFP mice, arranged according to harvesting time and staging time.*

(A) Image of *OSX*-GFP mouse with fluorescent expression of *Osx* in ossification centers (left lateral view).

(B/C) Each image shows the lateral view of an embryo's head with the eye (dashed line) in the bottom left quadrant. Ossification centers of the frontal and parietal bones are outlined to make clear the region of *Osx* expression. This sample includes images from 3 different litters collected at 6 hour intervals (E12.75, E13.0, and E13.25). Yellow numbers (in the bottom right quadrant of each panel) show order of the embryos when their sequence is determined by the harvesting age (ie. top row, B). These numbers are reproduced in the bottom row (C) once the sequence of embryos has been re-ordered based on the morphometric staging system, to highlight how the sequence has changed. The embryos exhibit substantial variation in the rate of development but ordering specimens according to estimated developmental stage (using eMOSS) provides a coherent view of the expected expansion of *Osx* expression over developmental time.
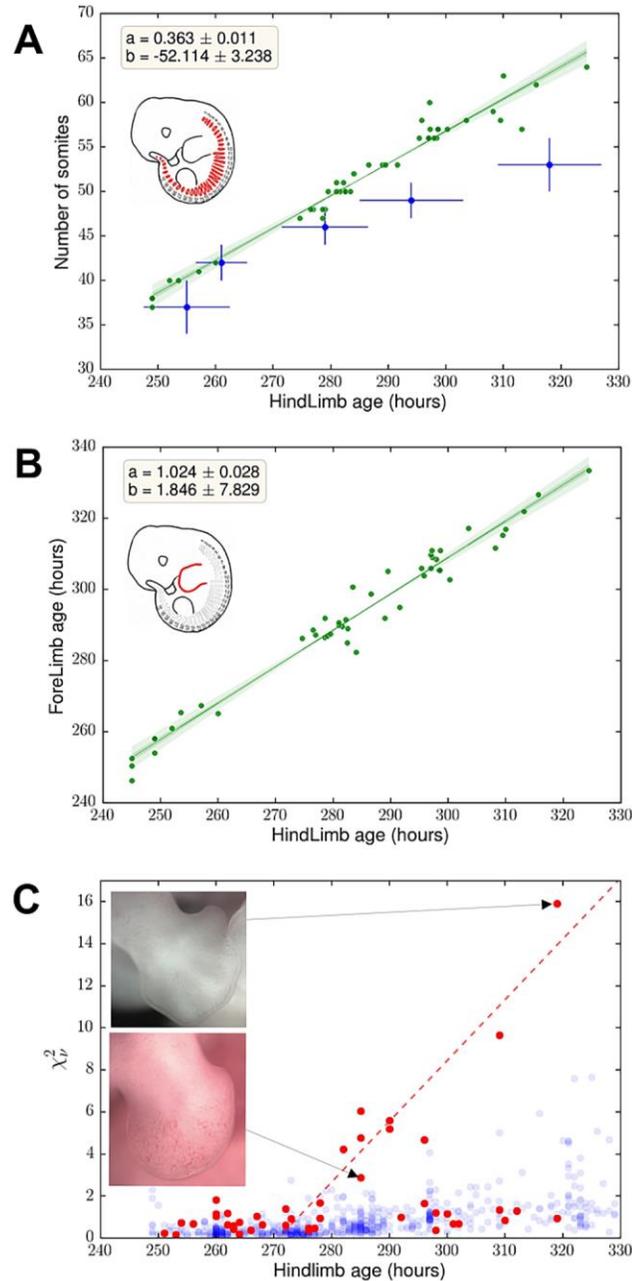
Figure 6: *Correspondence of the staging outcome and other morphological quantities.*

(A) The number of somites of the embryos as a function of the age of the hindlimb. Blue points corresponds to the results in Ref. [12]

(B) Calibration curve between the response of the staging system algorithm on hind and forelimbs.

(C) Distribution of the fit score as a function of age for limbs presenting the Ssq mutation. Points in red color correspond to limbs that carry the mutated gene, superimposed to the wild-type distribution in blue. Red points at low χ_v^2 values correspond to limbs that carry the genetic mutation but have a normal phenotype (as previously reported in [13]). The dashed line corresponds to a linear fit to points above $\chi_v^2 = 2$, where the phenotype is recognizable by visual inspection.