Edinburgh Research Explorer

A landmark-free morphometric staging system for the mouse limb bud

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher final version (usually the publisher pdf)

A landmark-free morphometric staging system for the mouse limb bud

Bernd Boehm[1,2], Michael Rautschka[1,2], Laura Quintana[1], Jelena Raspopovic[1], Žiga Jan[1] and James Sharpe[1,3,*]

SUMMARY

We have created a 2D morphometric analysis of the developing mouse hindlimb bud. This analysis has provided two useful resources for the study of limb development. First, a temporally accurate numerical description of shape changes during normal mouse limb development. Second, a web-based morphometric staging system, which has the advantage of being easy to use, and with a reproducibility of about ±2 hours. It allows users to upload a dorsal-view photo of a limb bud, draw a spline curve and thereby stage the bud within a couple of minutes. We describe how the system is constructed, its robustness to user variation and illustrate one application: the accurate tracking of spatiotemporal dynamics of gene expression patterns.

INTRODUCTION

Determining the developmental stage of an embryo is essential to developmental biology. Although the age of an embryo (in days or hours) is sometimes used for a rough estimate of developmental stage, genuine staging systems are preferable because different embryos (even genetically identical ones) may develop at different speeds – two littermates, harvested at the same time will often display different stages and can therefore not be directly compared in terms of developmental processes. A wide variety of staging systems have therefore been elaborated that cover the most common model organisms (*C. elegans*, *Drosophila*, Zebrafish, *Xenopus*, Chick, Mouse) (Hall, 2008; Hamburger and Hamilton, 1992; Campos-Ortega and Hertenstein, 1993; Nieuwkoop and Faber, 1956; Theiler, 1972), or even individual organs such as the developing eye or respiratory system (O'Rahilly, 1983; O'Rahilly and Boyden, 1973). At the present time, developmental stages of mammalian embryos are either defined by gestational time, by a suitable staging system (such as the Theiler stages) or by somite counting where possible.

Staging systems are generally based on the observation of qualitative features. For example, one of the features that distinguishes mouse Theiler Stage 19 (TS19) from the previous stage is that the lens vesicle has become completely closed and separated from the ectoderm (Theiler, 1972). Another feature is that the forelimb becomes visibly composed of two parts (i.e. the handplate is visible), whereas in the hindlimb this does not occur until TS20. These two examples illustrate some of the advantages of a classical staging system. The feature is a binary description and therefore relatively simple to classify – the lens vesicle is either closed or not, the handplate is either visible or not. Another advantage is that it does not rely on any special technology – one can simply examine the embryo with a dissecting microscope.

However, classical staging systems also suffer a number of disadvantages. The first is subjectivity – two different researchers will not always agree on whether the handplate is yet visible. Another is the possible lack of developmental synchrony – if limb bud development is not strictly synchronous with the embryo as a whole (or even between left and right limb buds) then a staging system based on features of the whole embryo will have a limited temporal resolution for staging the limb buds. The degree of natural variation in limb bud development is highlighted in Fig. 1 where simple measurements were carried out on a carefully selected dataset of mice of equal gastrulational ages (E10.75-E12.75) (Fig. 1B,C). Even buds which fall into the same Theiler Stage show a surprising degree of variability (Fig. 1D). A third major problem is temporal resolution – some developmental events in mammals occur over a period of hours, while the Theiler Stages generally last for about half a day. Finer staging is often achieved through the use of somite counting (Michos et al., 2004); however, this is not possible for later stages of development (as the earliest formed somites have already differentiated by this point), and it also does not overcome the potential problem of developmental asynchrony.

A fourth problem of these staging systems is a relatively new issue, just emerging as we enter an era of developmental systems biology. Data-driven computer simulations need to track temporal dynamics in explicit units of time (i.e. hours or minutes), see for example Boehm et al. (Boehm et al., 2010). An ideal resource therefore, for any given model species, would be an explicit continuous description of the developmental state of the embryo over time (e.g. its changing anatomy or morphology). The fact that embryos can development at slightly different speeds means that this standard developmental trajectory will have to represent an average individual.

In 1989, a limb-specific morphology-based staging system was defined (Wanek et al., 1989) covering the period from E10 until after birth. However, this framework has a rather low temporal resolution of 12 hours and it has not proven simple enough to be used routinely. We therefore wished to develop a new system that would be both more accurate and objective (by employing morphometric techniques) but also simpler to use. As different mouse strains can vary in size and in speed of developmental, we chose a cross between two different strains

[1]EMBL-CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), UPF, Calle Dr. Aiguader 88, Barcelona 08003, Spain. [2]MRC Human Genetics Unit, Edinburgh EH4 2XU, UK. [3]ICREA Professor, Centre for Genomic Regulation (CRG), UPF, Barcelona 08003, Spain.

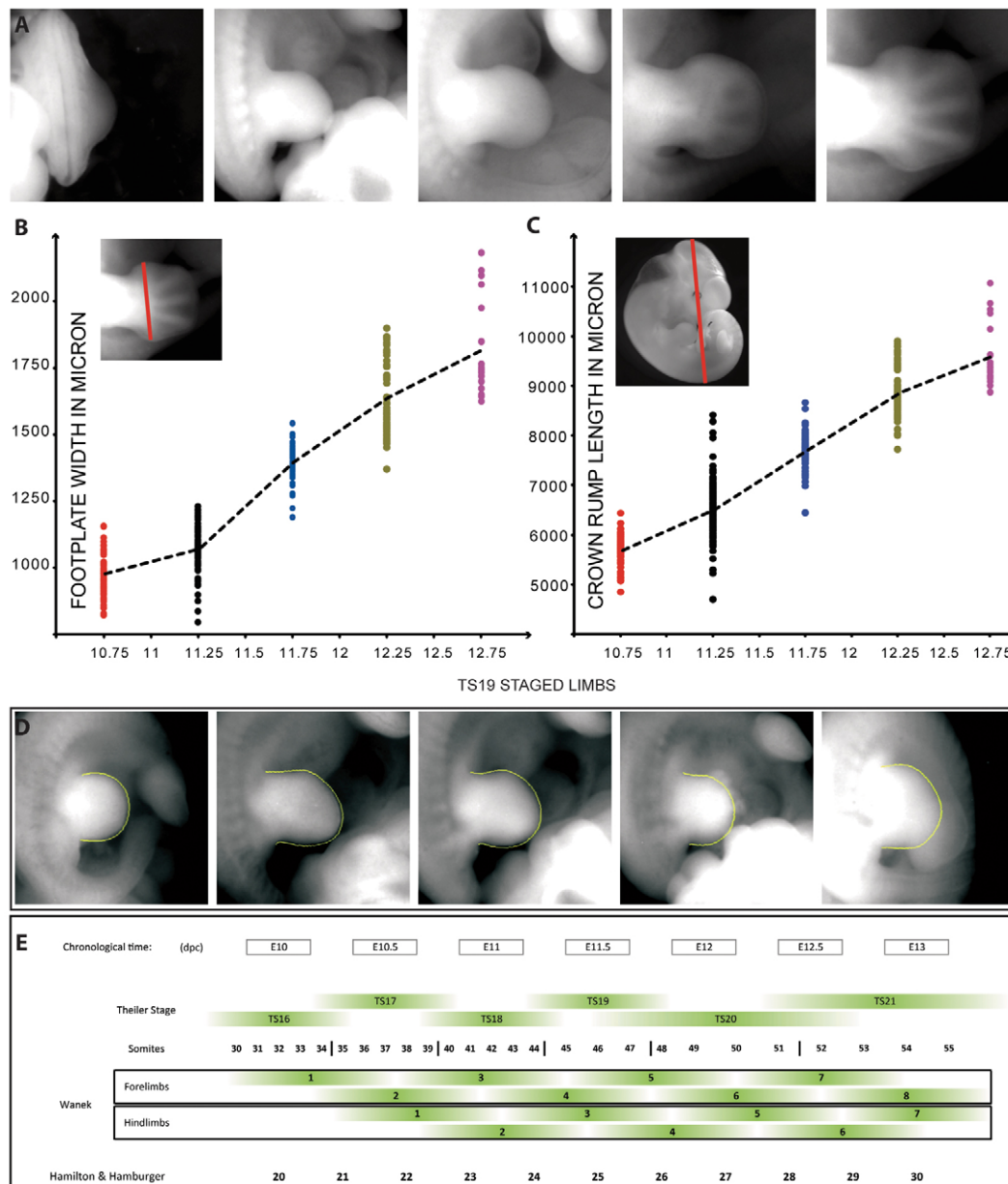*Author for correspondence (james.sharpe@crg.es)

Fig. 1. Current staging approaches for limb development. (**A**) A selection of right hindlimb buds photographed from the dorsal side to illustrate the sequence of normal limb development. (**B,C**) The high degree of size variability in embryo size is shown by linear measurements of footplate width (B) and crown-rump length (C) for 663 embryos harvested at six fixed timepoints after mating. Each embryo is represented as one datapoint and the huge degree of overlap between adjacent 12-hour timepoints is evident. (**D**) Five limb buds from embryos all staged as TS19 illustrate the low temporal precision of the Theiler staging system (Theiler, 1972). (**E**) A comparison of relevant staging systems highlights their strengths and weaknesses. Both Theiler stages and somite stages vary in length during development (becoming slower at later timepoints). The Wanek system (Wanek, 1989) is limb specific and therefore can stage fore- and hindlimbs independently, but has low temporal resolution. A rough alignment with the chick Hamilton and Hamburger system (Hamburger and Hamilton, 1992) is also displayed for comparison.

to maximize the applicability of the results across general laboratory strains (CBA×C57Bl6). To obtain a suitable degree of morphological variation within the dataset, we chose the F2 generation of this cross [as was chosen for the Edinburgh Mouse Atlas Project (EMAP) (Baldock et al., 1992)], as this produces a controlled source of genetic variation between the two original strains (Kaufmann, 1998). We focused on the hindlimb between stages E10.5-E12.5, as this covers the period of greatest morphology diversity.

MATERIALS AND METHODS
Calculating curvature
Cubic splines were implemented as described previously (Press, 1992). Each spline was sampled at equidistant points (20 μm apart) and the set of intervening line segments were used to calculate local curvature (α) using the following equation (Eqn 1):

$$\alpha = \arccos\left(\frac{\vec{a} \cdot \vec{b}}{|\vec{a}| \cdot |\vec{b}|}\right),$$ (1)

where a and b are sequential segment vectors. The full length of each spline was therefore represented as a sequential array of curvature values. This method is sensitive to the location of start and endpoints, and to whether the outline is traced in a clockwise or counter-clockwise direction (anterior to posterior or vice versa). As this function is affected by inaccuracies in the digitization of the outline, we applied additional smoothing by a moving average algorithm to reduce noise (Rohlf, 1990).

Shape comparison

The curvature arrays for two shapes were compared by iterating through every possible overlapping alignment of the two arrays, and then summing the differences of the aligned values as follows (Eqn 2):

$$\sum_{i=0}^{n} \left| ca_i - cb_{i+j} \right|, \qquad (2)$$

where ca and cb are the curvature arrays for shape a and b, n is the size of the larger array, i is the index for comparing curvature values, and j is the offset representing the shift of array b relative to array a (j ranges from $-n$ to n). With this offset, we compensate for the problem that no unique start or endpoint of the shape can be expected and still find homologous semi-landmarks. This calculation results in an age bias, such that well-matching older limb buds always have a larger difference value than well-matching younger limb buds, because their shapes are more complex. We therefore normalized the difference value by dividing by the total curvature gradients for both shapes. The local gradient of curvature along the spline is a successful choice of normalization parameter, because it increases with the complexity of the shape (e.g. the digits) while a simple summation of total curvature does not. For the size-independent version, the same shape-comparison process described above is performed multiple times. The query graph is first scaled to the same size as the reference graph. It is then re-scaled to multiple new sizes from 75% of original length to 125%, in steps of 1%.

Multiple alignment

The alignment of multiple shapes is an extension of the shape comparison described above. Within a nested loop, a graph (a) is compared to each of the others within the group (b). This results in a series of shift values, each of which would create the best alignment between a and b. These shift values are averaged and applied to a, before moving to the next in the list. Iterations of this spring-analogy algorithm are repeated until equilibrium is reached.

Staging new limb buds

New limbs are staged by comparing them with every graph in the standard trajectory. The only difference from a normal shape comparison (above) is that the graphs in the standard trajectories are smoother than single limbs because they were generated as the average of 20-40 real limb shapes. Consequently, new limb graphs have to be smoothed before being compared with the standard trajectory. We systematically explored a wide range of smoothing values and determined that an averaging window of 19 was optimal when applied to the curvature array. The averaging window was applied as follows (Eqn 3):

$$c_i = \frac{\sum_{j=i-s}^{j=s} c_j}{2s+1} , \qquad (3)$$

where c_i is the curvature at the i-th position of the spline, and s is the half-width of the averaging window ($s=9$).

RESULTS
A landmark-free method for limb bud shape comparison

Numerical methods for comparing two shapes (geometric morphometrics) typically depend on information about which points in each shape correspond to each other – known as 'landmarks' (Zelditch, 2004). However, for some biological structures, identifying suitable landmarks can be challenging, and in the case of the mouse limb bud this task appears almost

impossible. Externally the bud displays a smooth rounded shape, containing almost no singular points, and internally no easily identified landmarks can be seen, so our goal was to develop an approach that would not require the user to choose landmarks.

We wished our morphometric staging system to be as easy to use as possible, and therefore sought an approach which could extract the necessary measurements from a standard 2D photo of the limb bud. Although clear landmarks are lacking, it is nevertheless evident that the general profile of the limb bud (as seen from the dorsal side) progresses through a series of continuous yet stereotypical shape changes during development (Fig. 1A). The outline of the limb bud can be described by a single line that curves smoothly in 2D (within the plane of the photo). A convenient mathematical tool to capture such an arbitrary smooth curve is a spline, and we explored the use of a cubic spline for this purpose (Press, 1992). The splines for three limb buds of different ages are shown as yellow lines in Fig. 2A,E,I.

In addition to its ease of use, a key advantage of the spline is that it generates a smooth 2D shape that can be reduced to a 1D dataset, making numerical comparison easier. Alignment (or 'superimposition') is a generic requirement for shape comparison to remove the influence of location and orientation, but comparing two landmark-free shapes in 2D involves large degrees of freedom (including both rotation and translations) and is therefore a complex task that often has no unique solution (Zelditch, 2004). By contrast, comparing 1D datasets is much easier (as for example seen in DNA sequence alignment algorithms). If the absolute image scale is known, a linear shift of one set relative to another is sufficient to explore all possible alignments. In this case, we chose to use the local curvature along the spline as a unique and non-redundant quantification of limb-bud shape. In other words, a curvature graph can be generated for each spline, and the graph itself contains sufficient information to recreate the original shape unambiguously (i.e. a 1-to-1 mapping). The curvature graphs for the splines shown in Fig. 2A,E,I are displayed in Fig. 2B,F,J, and illustrate the main features of the graphs: concavities such as the wrist regions or the interdigital regions are recorded as negative curvatures (to the left of the vertical graph origin), while convexities such as the digits themselves are positive. Curvatures were calculated from a set of equidistant (20 μm) sampling positions along each spline (see Materials and methods). This method was adapted from the classical $\Phi^*(t)$ shape function by Zahn and Roskies (Zahn and Roskies, 1972) and Bookstein's tangent-angle function (Bookstein, 1978).

It is then relatively easy to derive a numerical comparison for the 1D curvature graphs. The absolute scaling of graphs (in μm) is known by recording the pixel dimension of the image when limb bud photos are taken. We can then compare two graphs by translating one relative to the other (in effect 'sliding' one graph along the other). For each step of this sliding procedure, we obtain a difference value (as the sum of curvature differences every 20 μm along the graph) and the minimum difference reflects how different the two limb bud shapes are. We have also developed a size-independent method, in which the graphs are fitted to each other by a combination of 1D translation (as above) and also an optimal scaling. This version is described below.

Calculating an average shape for standard timepoints

Owing to the variability of embryo development (discussed in the Introduction) the first step towards a continuous description of development is to determine the average limb shapes at a discrete
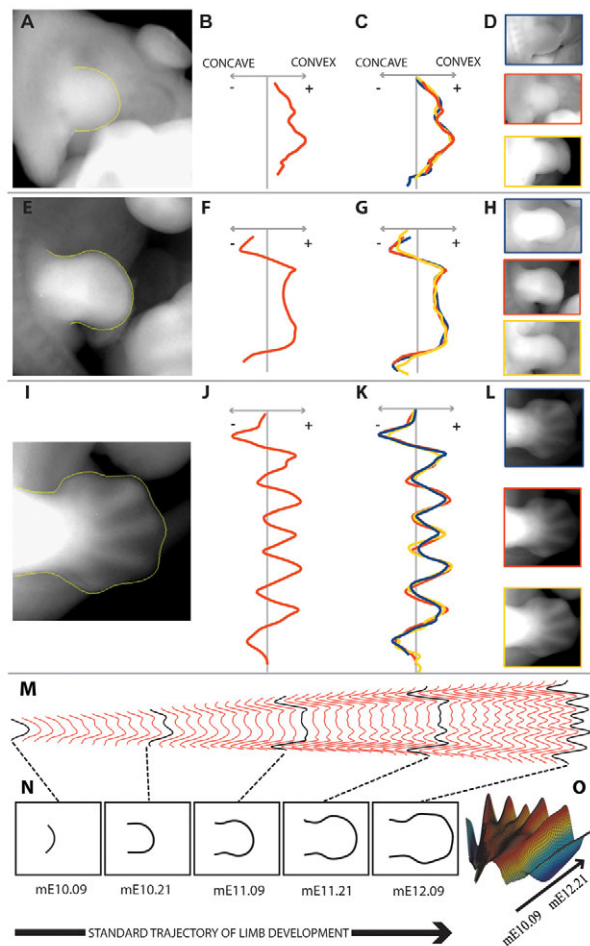
Fig. 2. Quantifying limb bud shape over time. (**A**) The outline of a young limb was determined using a spline curve (yellow line). (**B**) The curvature graph derived from this spline. (**C,D**) Multiple shapes of a similar age can be compared to each other by aligning their curvature graphs (C), extracted from their photos (D). (**E-L**) The same set of splines and alignments is shown for two older limb buds. Wrist regions can always be seen as negative curvatures (concavities) to the left of the vertical origin. Digits produce positive curvatures (convexities) to the right of the vertical origin. (**M**) The full interpolated series of curvature graphs is shown, with the primary timepoints in black. (**N**) Any of the curvature graphs can be converted back into a 2D limb bud shape, creating the full trajectory of normal developmental shapes. (**O**) The standard trajectory can be visualized in 3D view, the convex regions such as digits (red) and concave regions (blue) highlight the dramatic shape changes over 36 hours of development.

number of equally spaced timepoints. Unlike amphibians or birds, where the initiation of development can be controlled relatively accurately (by external fertilization or by shifting the temperature), the internal fertilization of mouse oocytes means that the precise moment of conception cannot be known. Overnight matings only limit the possible time of fertilization to an approximate 8-hour window, and the resulting variability therefore requires a large sample size to enable an average to be calculated. We therefore chose to collect many embryos for just four primary timepoints that were 12 hours apart, rather than collecting fewer embryos for more timepoints, in which the error for each timepoint would be greater. Based on the variability analysis (described below), 140 hindlimb

buds were chosen as a suitable sample size for each timepoint. The four timepoints were: (1) 9pm on the 10th day after mating; (2) 9am on the 11th day after mating; (3) 9pm on the 11th day after mating; and 9am on the 12th day after mating.

Although this spanned our main period of interest, we also collected a lower number of embryos at two extra time-points (12 hours before and 12 hours after the range) so that limb buds younger or older than the primary range would not be artificially 'clamped' to the minimum or maximum stage.

For each timepoint the collection of 140 limb bud shapes displays two sources of variability: stage dependent and stage independent. In principle, averaging shapes that only vary in a stage-independent manner (i.e. all those limb buds that are genuinely the same stage) would indeed produce our desired result: a representative average shape for that stage. However, as our collection includes a distribution of genuinely different stages, and the shape changes across this distribution are fairly arbitrary, there is no guarantee that averaging shapes of different stages will produce a representative shape for the average stage. As an example, averaging limb shapes from a collection that includes TS19, TS20 and TS21 will not necessarily recreate the average shape for the central stage TS20 (see Fig. S1 in the supplementary material). For this reason, we sought a method to exclude those outliers that represented older or younger limb buds while trying to retain the genuine stage-independent variation. We developed a successive approximation approach, in which an average shape is calculated from the full set of 140 limb buds (through a multiple alignment algorithm – see Materials and methods) and then the 20% of shapes most divergent from this are discarded before repeating the calculation. To determine how many times to repeat this procedure, we monitored the rate of change of the average shape over the iterations. As expected, over the first few iterations (seven) the rate of change of the average shape decreased and stabilized, suggesting that stage-dependant variations were being successfully filtered-out, and we were left with 20-40 shapes whose variation represents the true average stage of the group. Beyond iteration seven, the rate of shape change increased again, indicating that the process had gone too far, and that we were now simply reducing the sample size of the useful stage-independent variation, thereby shifting the calculated average shape away from the genuine stable average. We therefore chose the shape calculated at iteration seven as the representative shape for this timepoint.

Interpolating to a full developmental trajectory

Once an average curvature graph had been calculated for the four primary timepoints, we required a method to smoothly interpolate these graphs into a more finely sampled sequence, i.e. to fill in the 12 hour intervals. A necessary task therefore was to align curvature graphs for stages that are 12 hours apart from each other. A major challenge here is that in addition to having different shapes (e.g. the older graphs contain digit convexities, while the younger ones do not), the graphs are also of significantly different sizes (due to limb bud growth). We therefore devised a method to find the best fit that explores not only the best alignment of graphs (as before) but also the optimal scaling of each younger primary graph to fit the next older one. Once the optimal 'stretched' alignment had been found, equivalent positions along the splines were determined by mapping the set of curvature points from the older graph directly onto the scaled younger one (see Fig. S1 in the supplementary material). This was carried out for all six timepoints, effectively generating a series of equivalent curvature-based points across the

whole data set. For each point, we now had a curvature value for six timepoints during development (12 hours apart). We therefore chose to use a cubic spline to interpolate these six curvature values across the equivalent landmarks for all the 1-hour timepoints (55 interpolations in total). We now had a total of 61 curvature graphs representing the gradually changing limb bud shapes over time; however, as a last step before using this dataset we had to rescale the graphs back to their correct size. The scaling factors for the six primary graphs was determined by the stretching method mentioned above, and so we again used a cubic spline method to interpolate these scaling factors across the full dataset.

The resulting set of curvature graphs is shown in Fig. 2M and represents the first continuous numerical description of limb bud shape change – the standard developmental trajectory. This resource will be especially useful for computer modelling (e.g. Marcon et al., 2011), in particular because it is an accurate mapping of shape change onto real time (measured in hours). This will allow significant improvement over previous computer simulations in which it has not been possible to determine accurate rates of shape change.

We also developed a naming convention closely related to the traditional concept of embryonic days. Traditionally, conception is considered to occur at midnight and 12 noon the following day is labelled as E0.5. On the 10th day, a 9pm harvest would therefore lie halfway between E10.75 and E11.0. As we now have shapes for every 1-hour timepoint, we choose to call this timepoint mE11:21. The 'm' denotes that the system is morphometric and the colon indicates that the second number (21) is the hour of the day (9pm) rather than the fraction of the day implied by the more traditional nomenclature.

The staging system

We next wished to convert the developmental trajectory into a practical staging system easily accessible to the community. We therefore designed software which allows users to access the staging system using any standard web-browser connected to the internet. The software comprises three main parts (see Fig. S2 in the supplementary material). (1) A graphical user interface (GUI) written as a Java applet (which runs on the client-side, in the web browser). This allows users to upload their photos and draw the splines (Fig. 3A). Once a spline is drawn, the data is sent to the second part of the software (2), which is a server-side PHP program. This program calculates the stage of each new limb bud by performing the shape comparison function described above, to compare the new limb to all 61 standard limbs in the full trajectory (see Materials and methods). The best fitting curvature graph indicates the most probable stage of the new limb bud. Last, (3) we also developed a MySQL database to store all the photos, spline data and metadata (e.g. username, whether the limb bud is a left or right bud, etc.). Each user subsequently has access to all their photos, splines and predicted stages. This online resource is publicly available at http://limbstaging.crg.es.

An important test of the system is to explore how robust or sensitive it is to user variation. We use this criteria as the most practical definition of how accurate the system can be. Our first test was to determine how crucial a flat orientation of the limb bud is when photographed. We used a 3D computer model of a limb bud obtained by OPT (Sharpe et al., 2002) so that its orientation could be accurately rotated in 5° steps. We performed this test rotating about two axes: the anteroposterior axis and the proximodistal axis. For both cases, an image was rendered at a nominal flat angle, and then further images were generated for clockwise and anticlockwise rotations (Fig. 3B). This demonstrated that a rotation of up to ±15° generated a variation of just ±1 hour in the predicted age of the limb bud. Misalignments of 15° are relatively easy to spot by eye while positioning the bud for photo capture, and so the system is robust enough to cope with typical residual variation in limb bud orientation. Another important test is how much the results vary between different users of the system. Sixteen limb buds were chosen spanning the full range of developmental stages. Five users were asked to draw splines for all the limb buds, and the consistency in their results can be seen in Fig. 3C. The average standard deviation in the predicted stage showed that the reproducibility of the system is about ±2 hours, although it may be possible to improve this in the future.

Another practical issue is how limb bud samples should be prepared for photographing. The standard trajectory was created from freshly dissected limb buds, but it is useful to assess what effect typical treatments such as fixation or dehydration could have on the predicted stage. Sixteen limb buds around mE11.09 were staged after each step of a typical preparation process: (1) freshly dissected, (2) fixing in 4% paraformaldehyde, (iii) dehydrating in 100% methanol (e.g. for sample storage) and then (iv) rehydrating back into PBS. Fixation alone had little impact on the predicted stage (a 1-hour increase for some of the specimens), suggesting that limb buds could be routinely fixed before being staged. However, dehydration had a consistent and significant impact on the results – on average, limb buds were staged almost 7 hours younger than freshly dissected. Rehydration almost restored the results back to their original estimates (an average of ~1.5 hours younger than the freshly dissected values).

We investigated the cause for the inconsistent staging of dehydrated limb buds and found that it was largely due to the smaller size of the buds, rather than an altered shape. We therefore sought to create an alternative version of the staging system, in which the comparison considers only the shapes of the limbs such that size has no impact on the results. This was accomplished by modifying the algorithm so that it can re-scale the lengths of the curvature graphs (as well as translating them relative to each other) to find the optimal fit. Testing this on the same set of dehydrated limbs shows a dramatic improvement (Fig. 3D) – limb buds maintain their predicted stage across the various histological treatments. It is likely that by explicitly ignoring size information the new staging results display greater variability than before. We therefore tested this prediction by applying the new algorithm to the same test results as before (the five users on 16 limb buds), and we thus confirmed that indeed the standard deviation was now increased to ±3 hours (see Fig. S3 in the supplementary material).

Once the staging system was functional, we decided to explore the possible presence of any left-right asymmetries in the speed of development (assuming that left and right limb buds go through the same shape changes, but not necessarily with perfect synchronization), and in particular we asked two questions: (1) is there a detectable left-right bias in developmental stage during this part of limb bud development and (2) are there any detectable changes in left-right synchronicity over time? For each embryo, we staged the left and right hindlimbs independently, calculated the asymmetry as $Stage_{LEFT} - Stage_{RIGHT}$, and the stage for the embryo as a whole as the average of the two buds. The results show an average asymmetry of nearly 2 hours, and an upper s.d. of nearly 4 hours, which suggests that staging individual limb buds is preferable to extrapolating the stage form the body (e.g. somite counting). However, no significant bias towards the left or right
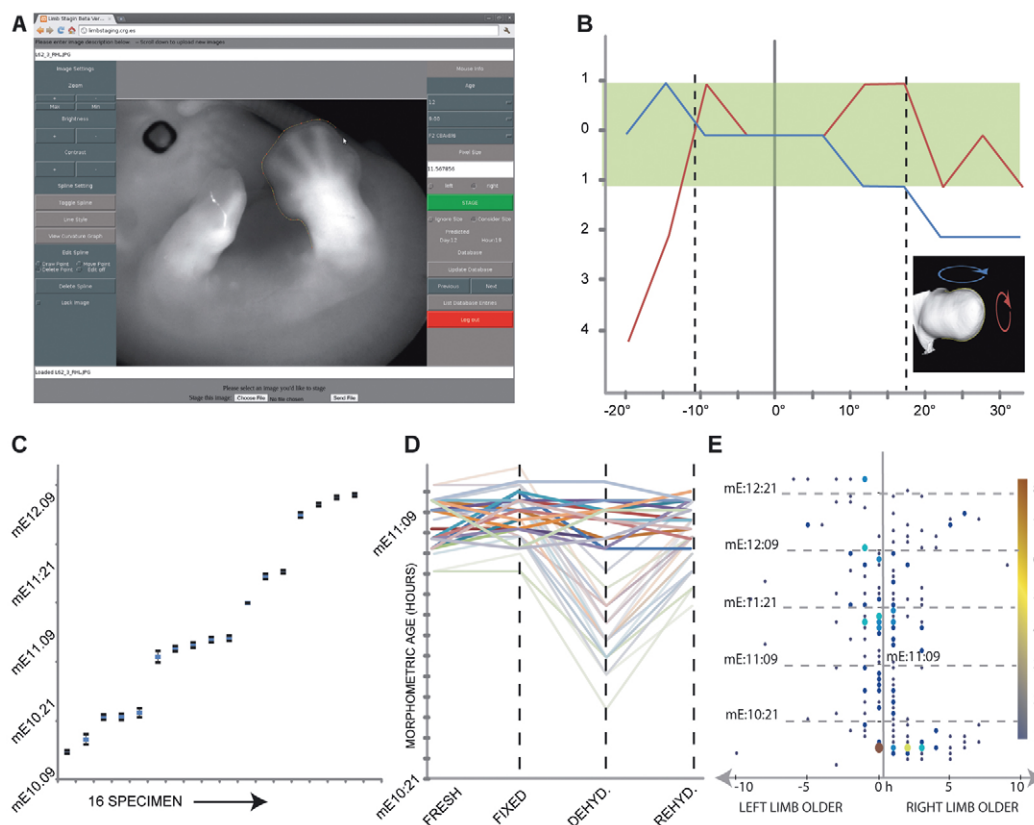
Fig. 3. The morphometric staging system. (**A**) The live screenshot show the GUI a user operates in order to stage a limb bud (publicly available at http://limbstaging.crg.es). The applet offers control over brightness, contrast and magnification, and the image can freely be moved in the window. The spline is drawn in the 'drawing mode' by clicking on the outline of the limb bud. A control point is added and the control points are connected by a spline interpolation. Additional data about the limb (such as strain, gastrulational age, its position and the pixel size of the microscope image) can be added and are stored in a database. The green 'stage' button opens a new window/tab with the staging result. (**B**) Graph illustrating the robustness of the system to limb orientation during photography. Within a range of 30°, the result varies by just ±1 hour. (**C**) Graph illustrating the robustness to user-dependent variations from five novice users. y-axis labels are 12 hours apart, and the average s.d. across the 12 limb buds was ~45 minutes. (**D**) The influence of tissue processing on the staging results. Sixteen specimens of ~mE11:09 were dissected, fixed, dehydrated and rehydrated. For each step, a photo was taken and staged. Results are shown for two versions of the staging algorithm. In pale colours are the size-dependant results that show a strong dip in predicted stage upon dehydration, whereas in bolder colours are the results of the size-independent algorithm. (**E**) The training dataset was analysed for left-right asymmetry. Size and colour show the number of limb pairs staged the same. There is an average asymmetry of nearly 2 hours, but no strong bias to one side.

sides was observed, and no strong trend of changing asymmetry over time (Fig. 3E), although this remains an interesting issue for future studies.

Spatiotemporal dynamics of gene expression

One valuable application for an accurate organ-specific staging system is to help resolve the spatiotemporal dynamics of gene expression patterns. Some genes display patterns that are spatially complex and change rapidly over time. A staging system that is accurate to only ±6 hours will group together stages that display detectable differences in the pattern, making it hard to resolve the true evolution of gene expression. One of the most dynamically changing gene in the developing limb bud is Sox9, which progressively defines the emerging skeletal pattern over time. We thus performed whole-mount in situ hybridization for this gene on a collection of C57Bl6 embryos. Fig. 4A highlights the value of the staging system by showing the left and right limb buds from embryos of a single litter. Despite all coming from the same litter, a wide spread of stages is evident. The morphometric staging system is able to reveal that the embryos vary in age by up to 24

hours (from mE11:21 at the top of the figure to mE12:21 at the bottom) and confirms that the progression of the Sox9 pattern matches the estimated stages closely. The value of the system for comparing multiple genes assayed in different specimens, is highlighted by the example in which the left and right limb buds of one specimen are 4 hours different in stage (4th row in Fig. 4A) and the development of their Sox9 patterns also reflects this asynchrony (cleaner separation of digit expression in the older left bud and weaker zeugopod expression).

We next used Sox9 expression patterns to ask another question. Will different strains of mouse show the same correlation between morphological development and gene expression patterns? This would be a useful way to address the more general practical issue of whether our staging system can be used for many different mouse strains. By performing whole-mount in situ hybridization for Sox9 on two further mouse strains (52 OF1s and 48 CD1s) we were able to build up a direct comparison of the patterns at three different stages (roughly 6 hours apart). Although the embryos and adults of the three strains are known to vary in average size (and for example CD1s and OF1s possess a 19-day gestation, whereas
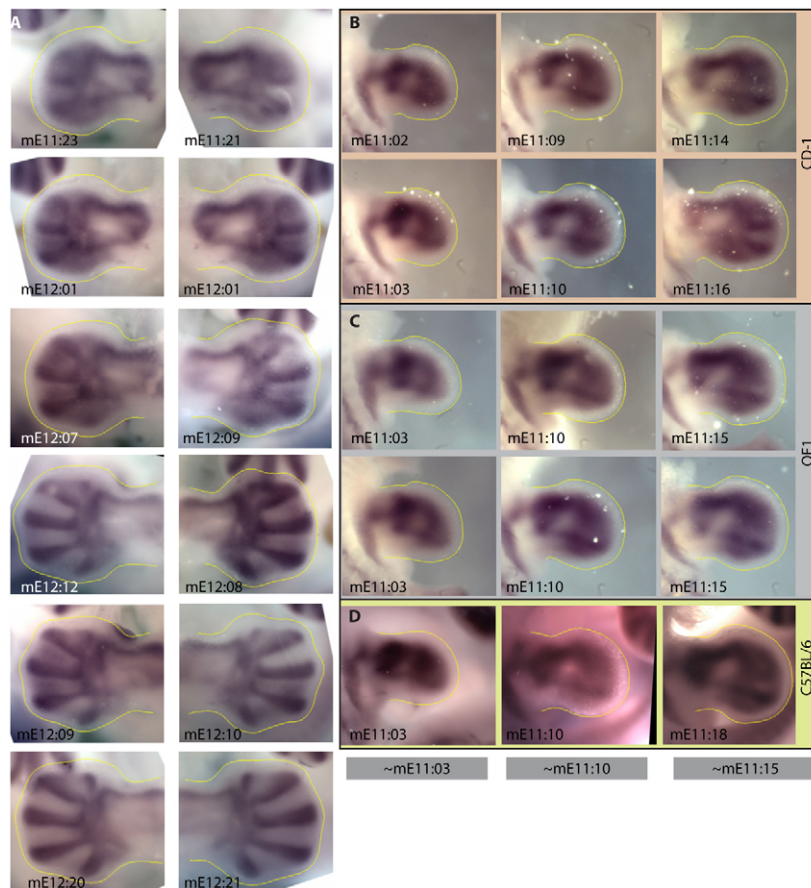
Fig. 4. The dynamics of Sox9 expression. (**A**) A series of six left-right pairs of Sox9 patterns all taken from a single litter. The youngest limb buds are roughly 24 hours younger than the oldest buds. By ordering the limb buds based on their morphometric stage, we see the progression of a subtle aspect of the Sox9 pattern. Strikingly, in the fourth pair of buds, the left bud is staged 4 hours older than the right (mE12:12 versus mE12:08) and the Sox9 pattern is slightly more developed (the expression domain connecting digits 1 and 2 is weaker, which is more similar to the older patterns of the litter). (**B-D**) A comparison of Sox9 progression between three different strains of mouse: (B) CD1, (C) OF1 and (D) C57Bl6. Although individual variations exist in the subtle details of the pattern, an agreement between pattern and morphological stage is seen across the three timepoints that were roughly 6 hours apart. Two limb buds are shown for the OF1s and CD1s of each stage to highlight the slight individual-to-individual variation. Further details of the full Sox9 timecourse are provided in Fig. S4 in the supplementary material.

Bl6 have a 21-day gestation), nevertheless our staging system found a close correspondence across the three strains for the relationship between morphometric stage and the Sox9 pattern (Fig. 4B). This appears to support our original choice to build the standard trajectory from the F2 generation of a cross between C57Bl6 and CBA, rather than from a single inbred strain.

DISCUSSION

A major advantage of our morphometric system is that it achieves alignment and comparison of shapes that possess no obvious landmarks, and without forcing the user to pick pseudo-landmarks. This also makes photographing the samples easier, as the limb buds do not have to be oriented at any special angle. The system may therefore be suitable to be applied to other landmark-free organs in the future. It should be noted that, for convenience, we have given each stage a name based on the average age from our training data set. However, our system is nevertheless a genuine staging system – not an 'ageing system'. As an example, not all embryos staged as mE11:09 will reach this point at exactly 11 days and 9 hours after conception. Nevertheless, we consider this nomenclature sensible as it does provide a rough indication of the embryo age. It is also important to note that although the systems provides an estimate to the nearest hour, the reproducibility is approximately ±2 hours for the size-dependant method and ±3 hours for the size-independent method. It is hoped that it may be possible to improve on this for future versions of the system.

The general principle applied here is not restricted to the mouse or even to limb buds. A similar staging system could be developed for chick limb development or the morphogenesis of any embryonic structure that does not possess obvious landmarks.

Additionally, a useful application of the limb-specific system could be the staging of manipulated embryos (mutants, or other experiments) in which Theiler Staging or somite counting is no longer possible owing to the resulting phenotypic changes across other parts of the embryo (e.g. if the mutant alters somitogenesis). In this respect the limb buds may be particularly practical, as they can be dissected off the main embryo and stage-analysed separately.

Finally, in addition to a practical and objective staging system, our analysis has created the first accurate trajectory of average shape changes during development, which is mapped to the progression of time in explicit units (hours). This will be an invaluable resource for computer models of limb development, as seen by Marcon et al. (Marcon et al., 2011), who were able to build a method for clonal analysis on the basis of this morphometric trajectory.

Competing interests statement
The authors declare no competing financial interests.

Supplementary material
Supplementary material for this article is available at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.057547/-/DC1

DEVELOPMENT

References

Baldock, R., Bard, J., Kaufman, M. and Davidson, D. (1992). A real mouse for your computer. *BioEssays* **14**, 501.

Boehm, B., Westerberg, H., Lesnicar-Pucko, G., Raja, S., Rautschka, M., Cotterell, J., Swoger, J. and Sharpe, J. (2010). The role of spatially controlled cell proliferation in limb bud morphogenesis. *PloS Biol.* **8**, e1000420.

Bookstein, F. (1978). The measurement of biological shape and shape change. New York: Springer-Verlag New York.

Campos-Ortega, J. A. and Hartenstein, V. (1993). *Atlas of* Drosophila *Development*. Plainview, NY: Cold Spring Harbor Laboratory Press.

Hall, D. H. and Altun, Z. F. (2008). C. elegans *Atlas*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Hamburger, V. and Hamilton, H. L. (1992). A series of normal stages in the development of the chick embryo. 1951. *Dev. Dyn.* **195**, 231-272.

Kaufman, M. H. (1992). *The Atlas of Mouse Development.* London, UK: Academic Press.

Marcon, L., Arque's, C. G., Torres, M. S. and Sharpe, J. (2011). A computational clonal analysis of the developing mouse limb bud. *PloS Comp. Biol.* **7**, e1001071.

Michos, O., Panman, L., Vintersten, K., Beier, K., Zeller, R. and Zuniga, A. (2004). Gremlin-mediated BMP antagonism induces the epithelial-mesenchymal feedback signaling controlling metanephric kidney and limb organogenesis. *Development*. **131**, 3401-3410.

O'Rahilly, R. (1983). The timing and sequence of events in the development of the human eye and ear during the embryonic period proper. *Anat. Embryol.* **168**, 87-99.

O'Rahilly, R. and Boyden, E. (1973). The timing and sequence of events in the development of the human respiratory system during the embryonic period proper. *Anat. Embryol.* **141**, 237-250.

Nieuwkoop, P. D. and Faber, J. (1956). *Normal Table of* Xenopus laevis. Amsterdam: North-Holland.

Press, W. H., Teukolsky, S. A. Vetterling, W. T. and Flannery, B. P. (1992). Numerical recipes. In *C: the art of scientific computing*, 2nd edn, p. 113. Cambridge, UK: Cambridge University Press.

Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sorensen, J., Baldock, R. and Davidson, D. (2002). Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science* **296**, 541.

Theiler, K. (1972). *The House Mouse: Development and Normal Stages from Fertilization to 4 Weeks of Age*. Berlin, Germany: Springer-Verlag.

Wanek, N., Muneoka, K., Holler-Dinsmore, G., Burton, R. and Bryant, S. V. (1989). A staging system for mouse limb development. *J. Exp. Zool.* **249**, 41-49.

Zahn, C. and Roskies, R. (1972). Fourier descriptors for plane closed curves. *Computers IEEE Trans.* **100**, 269-281.

Zelditch, M. (2004). *Geometric Morphometrics for Biologists: A Primer*. London, UK: Academic Press.