

LAMA: Automated image analysis for developmental phenotyping of mouse embryos

Neil R. Horner¹, Shanmugasundaram Venkataraman², Ramón Casero¹, James M. Brown³, Sara Johnson¹, Lydia Teboul¹, Sara Wells¹, Steve Brown¹, Henrik Westerberg¹ & Ann-Marie Mallon^{1*}

¹ Medical Research Council Harwell Institute, Harwell, Oxfordshire OX11 0RD, UK

² MRC Institute of Genetics and Molecular Medicine (IGMM), Edinburgh EH4 2XU

³ School of Computer Science, University of Lincoln, Lincoln LN6 7TS

*corresponding author

Abstract (150)

Advanced 3D imaging modalities such as micro computed tomography (micro-CT), high resolution episcopic microscopy (HREM), and optical projection tomography (OPT) have been readily incorporated into high-throughput phenotyping pipelines, such as the International Mouse Phenotyping Consortium (IMPC). Such modalities generate large volumes of raw data that cannot be immediately harnessed without significant resources of manpower and expertise. Thus, rapid automated analysis and annotation is critical to ensure that 3D imaging data is able to be integrated with other multi-dimensional phenotyping data. To this end, we present an automated computational mouse phenotyping pipeline called LAMA, based on image registration, which requires minimal technical expertise and human input to use. Designed predominantly for developmental biologists, our software performs image pre-processing, registration, statistical and gene function annotation, and segmentation of 3D micro-CT data. We address several limitations of current methods and create an easy to use, fast solution application for mouse embryo phenotyping. We also present a highly granular, novel anatomical E14.5 (14.5 days post coitus) atlas of a population average that integrates with our pipeline to allow a range of dysmorphologies to be automatically annotated as well as results from the validation of the pipeline.

Introduction

A major goal in biomedical research is to assign functional roles to all genes in order to shed light on disease mechanisms, and to identify disease-associated genes and novel drug targets. However, almost two decades since the human and mouse draft genomes were published (Lander et al., 2001; Waterston et al., 2002), the proportion of genes still in the dark genome (minimal gene function or disease association assigned) remains at over 30% (Oprea 2019). One important approach to add functional knowledge to each gene in the genome involves the phenotyping of mutant mice, leveraging the close homology between the mouse and human genomes as well as the relative ease of generating null deletions and the existence of established phenotyping procedures (Hrabě de Angelis 2015; Meehan et al., 2017). The International Mouse Phenotyping Consortium (IMPC) is a high-throughput functional genomics project that is well on its way to generating a genome wide catalogue of gene function by phenotyping all gene knockouts in mice on a uniform genetic background (Brown et al., 2012; Lloyd et al., 2020). To date phenotype data for over 7000 genes are available on the IMPC web portal (mousephenotype.org), data which has already contributed to the identification of many novel candidate disease genes and potential new mouse models of human disease (Cacheiro et al., 2020; Meehan 2017; Bowl et al., 2017; Moore et al., 2018).

Approximately one third of all knockout mouse lines exhibit embryonic or early postnatal, lethality or subviability (Hrabě de Angelis et al., 2015; Dickinson et al., 2016). These genes are of particular importance as they can provide insights into developmental processes and disorders. The IMPC embryo phenotyping pipeline aims to study these genes at key developmental time points, with a major component being the generation and analysis of high resolution, whole embryo, 3D images (Adams et al., 2013).

Embryo anatomy can be manually assigned phenotype annotation by experts with knowledge of mouse anatomy and development (Wilson et al. 2016). However, this is prohibitively time consuming to perform at scale and is subject to user error and bias. Furthermore, manual annotation is susceptible to inter-operator variability, and may miss abnormalities that are not present on a predetermined checklist of structures that reflect annotator expertise. Automated analysis methods that aim to address these issues have been developed and were first applied

to mouse whole embryos, by analysing magnetic resonance images (MRI) from wild type mice at embryonic day 15.5 (E15.5) (Zamyadi et al. 2010). This work was expanded to the analysis of micro-CT images at E15.5 and incorporated the analysis of mutant embryo images (Wong et al., 2014). The latter works by aligning with a non-affine spatial transformation the gross anatomy of mutant embryos that share a genetic alteration and wild type embryos, and then localizing regions of anatomical dysmorphology by modelling the effect of genotype on local volume changes or voxel intensities, in a process called voxel-based morphometry (VBM). The resulting heat maps, representing statistical significance of local voxel deformation, can be overlaid on embryo images to highlight local dysmorphology. Mouse embryo atlases in which visible anatomical structures have been delineated enable the assignment of dysmorphology to anatomical labels as well as the automatic calculation of organ volume by the propagation of atlas labels post-registration back towards the input images. Atlases can be made using a single segmented specimen as a reference, for example those made for the eMouse Atlas Project (EMAP), which created atlases from multiple developmental stages from serial histological sections (Richardson et al., 2014). However the segmentation of organs from micro-CT images and MRI can be more difficult due to increased noise and reduced contrast. A solution to this is to segment a population average consensus image formed from spatially normalising multiple specimens, as shown by (Holmes et al., 1998) in brain MRI, and provides an unbiased registration target from which to propagate anatomical labels (Hammers et al., 2002). Mouse embryo atlases of population average models derived from MRI images (Cleary et al., 2010) and micro-CT (Wong et al., 2012) have been developed for the E15.5 stage where 6 and 48 structures were segmented respectively. These methods have yet to be applied to the analysis of other key developmental stages that are studied within the IMPC, namely E14.5, at which point organogenesis becomes largely complete, and E18.5, the stage immediately prior to birth and which is an important stage for the analysis of perinatal lethality or subviability.

Sample size of both mutant and wild type controls will have a substantial effect on the power of a mouse phenotyping study and its consideration in automated analysis of mouse embryos has not been investigated in-depth. Wong et al., (2012) performed a power analysis using automatically-calculated wild type organ size distributions from four organs (brain, liver, lung and myocardium), and showed that organ volume differences between 9-14% were theoretically detectable using a sample size of eight mutants and controls. Their analysis also showed that increasing sample size may substantially improve sensitivity, but they chose a sample size that was a compromise between sensitivity and computational expense imposed by the analysis pipeline design. Up to

now, however, no empirical assessment of the effect of sample size has been carried out on whole embryo data. Imaging data submitted to the IMPC embryo pipeline often have mutant sample sizes that are significantly lower than the desired eight mutant specimens due to cost and other practical limitations. However, imaging data from wild type control specimens can be much higher as they accumulate over the lifetime of a project. Applying these large numbers of wild type control animals in the automated analysis should increase statistical power, but this has yet to be investigated.

A surprising finding from recent high-throughput studies is that mutant lines generated from isogenic inbred mice frequently exhibit incomplete penetrance and variable expressivity of phenotypes (Wilson et al., 2017; Dickinson et al., 2016). These phenomena potentially complicate efforts to assign phenotype annotations at the line level. The ability to reliably assign phenotype annotations to each animal (specimen-level annotation) would help with cases of penetrance and expressivity variability. In this work we address this question by empirically testing the ability to detect sex differences and phenotypes of knockouts with varying sample sizes.

Another factor to consider is the developmental stage of specimens being tested, as it is known that embryos harvested at E14.5 represent a range of developmental substages due to uncertainty around the exact time of conception. It has been shown that selecting wild type comparison specimens that are stage-matched is vital to prevent misidentification of dysmorphology during manual annotation of embryos (Geyer *et al.*, 2017), but this problem has yet to be studied in the context of automated phenotyping using 3D images.

The issues raised above motivated us to develop a new automated phenotyping pipeline, which we present in this paper. The software pipeline ‘Lightweight Analysis of Morphological Abnormalities’ (LAMA) is freely available software developed for Linux and Windows™ platforms (<https://github.com/mpi2/LAMA>). LAMA is an automated computational phenotyping pipeline based on image registration that performs image pre-processing, registration, statistical annotation and segmentation of 3D image data. LAMA requires minimal technical expertise to install, and can run on a standard desktop PC. The registration strategy that we have employed speeds up the spatial normalization process and enables a greatly increased sample number of baselines to be used as controls, thereby increasing the power to detect abnormalities. We also present results of the validation of the pipeline by testing the ability of LAMA to detect sex-specific differences between sets of E14.5 embryos as well as showing that it can uncover known

anatomical abnormalities in a broadly affected knockout line as well as a knockout line displaying specific, localised dysmorphology. This represents the first time that VBM methods have been applied successfully to whole embryos at this developmental stage. We also show that the developmental substage is an important variable when designing automated phenotyping experiments and we provide a solution to mitigate this confounding effect. We have begun to address the issue of variable expressivity and incomplete penetrance and discuss future areas of research in these domains.

To accompany the pipeline, a novel, highly detailed E14.5 anatomical atlas of a population average is also presented. The population average was computed from micro-CT images of 16 wild type mice at a high resolution of $14\ \mu\text{m}^3$. The atlas consists of the manual segmentation of 184 organs and anatomical structures in the population average representing the most detailed average mouse embryo atlas currently available. The population average and atlas will be a valuable resource for the developmental biology community to aid in both manual and automated analyses. Using them along with LAMA, it is possible for the first time to automatically annotate dysmorphic E14.5 embryo anatomy with ontological EMAPA terms (Hayamizu et al. 2013), which are ready for export to the IMPC for display on the portal, and as we move forward, this novel dataset of ontological phenotype associations will be integrated with other data sets enabling to study the gene function relationships to anatomy and ultimately disease.

Results

Overview of the phenotyping pipeline for E14.5 mouse embryos

LAMA is a voxel-based morphometry approach to automate the detection of anatomical dysmorphology in mouse embryos. It is written in the Python programming language and features spatial normalisation of images using a groupwise registration process to iteratively align micro-CT embryo images into the same coordinate space. Internally, it uses elastix (Klein et al., 2010) for pairwise multi-resolution 3D image registration. Firstly, a population average model embryo must be constructed from wild type derived images (Fig 1A), which serves as an atlas template due to its high contrast and low signal to noise ratio. The population average also acts as a target image for the subsequent spatial normalisation of the phenotyping images (described below) (Fig 1A). The next step involves spatially normalising each wild type and mutant image that will be

used in the downstream statistical analysis by registering it onto the population average (Fig 1B), resulting in homologous anatomical structures occupying identical coordinates and morphology. After spatial normalisation, applying the inverse registration transformation to the population average mask and atlas, aligns them back onto the corresponding specimen image and allows for the calculation of whole embryo volume and each of the atlas labels (Fig 1B). For this paper, we focus on the organ volume analysis as this is linked to the atlas resulting in automated EMAPA-associated phenotype calls that are more robust and interpretable than those at the individual voxel level. However, we also include statistical parametric heat maps from the voxel-level Jacobian determinants (which indicate local volume shrinkage/expansion during spatial normalisation) for illustrative purposes.

Creation of a Novel E14.5 mouse embryo atlas

The E14.5 population average used in this study was created from 16 specimens (8 male and 8 female), with a resulting crown-rump length of 9.18 mm and an isotropic image resolution of 14 μm^3 (Fig. 2A). The visible organs were segmented using a mixture of manual and semi-automatic segmentation (see Methods) producing a segmented E14.5 atlas containing 184 unique labels (Fig. 2B; Table 1). A key requirement for the full utilization of the atlas, is that the segmented labels are assigned appropriate community-accepted ontology terms allowing the automatic integration of the resulting gene-to-phenotype data from LAMA with other data sets. The 184 labels were therefore mapped to gross anatomy terms from the EMAPA developmental ontology (Hyamizu et al., 2013). Through manual inspection of registration results a number of the labels were identified as being too small or thin leading to difficulty in assessing registration accuracy, such as the external carotid arteries (Fig. 2C). With this in mind, we identified such labels in the atlas (see Methods) to exclude them from downstream analyses, resulting in a final set of 103 labels that were used in the current analysis (Table 1). These 103 labels were distributed across the majority of the EMAPA high-level organ system terms (Fig. 2E) and range in size from the largest (forebrain at 35,596,092 μm^3) to smallest (metatarsals 14,420 μm^3) (Fig. 2D) .

Developmental substage

Embryos harvested at E14.5 represent a range of developmental substages (DSS) and have rapidly developing anatomy and so it is crucial for a high throughput data analysis pipeline to

account for this variation in the data. To account for overall size of the embryo, derived organ volumes are normalised by whole embryo volume and the Jacobian determinants are normalised by limiting the analysis of the Jacobian determinants to those generated after the affine registration stage, which accounts for overall embryo size (see Methods) as done previously at E15.5 (Wong et al., 2012). To assess whether this normalization is sufficient to account for differences due to embryo DSS, we made two datasets comprising of wild type specimens where eight of the smallest or largest specimens were labelled as “mutants” (see Methods), thereby simulating the cases of mutant lines containing embryos at early or late E14.5 DSS and we applied the LAMA pipeline statistical analysis to each data set. (Fig 3A). Both tests returned significant Jacobian determinant voxels suggesting that the relabelled wild types had morphological differences that were dependent of the developmental stage of the specimens. We then included a surrogate for DSS, whole embryo volume (WEV), within our linear model analysis, which resulted in the two tests returning no significant voxels for genotype effect, showing that the DSS effect can be controlled for in this manner. To gain a more detailed view on the voxel-level DSS-dependent relative size differences, we fitted Jacobian determinates from 93 wild type specimens to a linear model by WEV only (Fig 3B) which highlighted regions that are proportionally larger at later stages (red) or proportionally smaller at later stages (blue). Similarly, the stage effect on WEV-normalised organ volumes was significant for 78/103 labels (Table 2) including organs that were proportionally larger (n=58) later in development such as thymus and lung lobes (Fig. 3C-D) and those that were proportionally smaller (n=23) including brain ventricles and trigeminal glands (Fig. 3E-F). To summarize, normalizing the Jacobian determinants or organ volumes before statistical analysis is not sufficient to account for DSS and failure to model DSS can lead to false positive results, and one way to account for this is to regress out the DSS effect in the statistical analysis.

Optimal sample size for phenodeviance testing

In order to validate LAMA with a positive control we applied it to wild type embryos where females were relabelled as “mutant”. This gives us a convenient dataset where we expected the only gross differences between the sets to be located at the gonads (see Fig. 4, suppl. 1 for an example of gonad images). A quality control step removed specimens with obvious tissue damage due to sample preparation, extreme imaging artefacts or non-determined sex, resulting in a data set consisting of 89 wild type specimens (49 male, 40 female). Using a linear model (organ volume ~

sex + whole embryo volume), we found that along with the gonad (FDR-corrected q-value = $1.3e^{-25}$) (Fig. 4A), the lens of the eye unexpectedly also had a significant sex effect (FDR-corrected q-value = 0.028) (Fig. 4B; Suppl. Fig 4,2). Therefore, we included the lens along with the gonad as true positives in the following experiment. With this large data set, we were also able to address the effect that sample size has on the sensitivity of phenodeviance detection (in this case the ability to differentiate between male and female gonads and lenses). We ran a series of experiments in which organ volume was first normalised by whole embryo volume and then regressed on whole embryo volume and sex as described previously, in which each experiment had a different number of male and female (relabelled as “mutant”) specimens. Each experiment was repeated 50 times with random specimen selection, and using the permutation approach to correct for multiple testing (see Methods). Due to the large difference between male and female gonad sizes, significant gonad volume differences were identified in almost every replication of each experiment (Fig. 4C), and significant Jacobian determinant voxels were identified within, or close to, the gonad with significant voxels covering a larger area with increasing male sample size (Fig 4E). For the lens of the eye, significant volume differences were detected only with a male sample size of 32 or over, with the maximum male and female sample size (49 male & 8 female) resulting in significant hits in over a half the cases (Fig. 4D). To assess the rate of false positive detection, any significant organ volumes, other than gonad or lens of the eye in the preceding experiments were classed as false positives. The rate of false positives was found to be well controlled with only 1 of 103 organs called as significant in more than 1% of tests (epiglottis in 1.6% of all the replications), and with a mean false positive rate of 0.07% per label. These experiments show that LAMA is able to identify sex-specific differences in wild type embryos and that even with a low mutant sample size, differences in morphology can be detected more reliably as the control sample size is increased.

Automated identification of developmental phenotypes in E14.5 mice embryos

The initial aim of LAMA is to automatically identify dysmorphology from data generated by IMPC and other projects. To demonstrate its effectiveness on IMPC-generated data, we have chosen two exemplar mutant lines that illustrate its use in embryos with multiple dysmorphologies across the body and embryos with very specific, localised abnormalities.

The first example is *Wfdc2*, which encodes a protease inhibitor protein that is expressed in several tissues during mouse development prior to E14.5 (Lizio et al., 2015) including intestines, lungs and pancreas. *WFDC2* plays a role in cancer development (Bingle et al., 2002; Li et al., 2013) and two papers have recently shown *Wfdc2* homozygous mutants display severe pulmonary phenotypes in mice including collapsed lungs at perinatal day 1.5 (P1.5) (Nakajima et al., 2019), alveolar abnormalities, dyspnea and reduced blood oxygen saturation at birth (Zhang et al., 2019), but are otherwise anatomically normal. The IMPC reported a partially penetrant preweaning lethality phenotype for this line with 5.5% of the genotyped pups determined as homozygous for the mutation. Our analysis of four E14.5 *Wfdc2*^{-/-} specimens resulted in significantly smaller organ volumes for the bronchi at the gene-level. (Fig 5A) but no significant specimen-level differences were observed for this gene. The Jacobian determinant analysis identified two significant regions that largely overlap with one of the bronchi after the FDR-corrected p-value threshold was raised to 0.1. This means that for this gene, the whole organ statistics were more sensitive than the Jacobian determinants. In all four mutants, the trachea and bronchus are visibly smaller in diameter (Fig. 5C-D), but otherwise appear normal.

Acan encodes for the protein aggrecan, which is the primary proteoglycan in articular cartilage, is present in the extracellular matrix of long bone epiphyseal growth plates, and is required for normal bone development. *Acan*^{-/-} mice exhibit phenotypes associated mainly with abnormal bone morphology, including long bones, abnormal ribs and vertebrae as well as enlarged liver, and pulmonary hypoplasia (Table 3). Human diseases are associated with *ACAN* mutations including osteochondritis (Stattin et al., 2010) and skeletal dysplasia (Tompson et al., 2009). Along with complete preweaning lethality, the IMPC reports a reduced bone area composition and increased circulating cholesterol levels in adult heterozygous animals. We analysed six E14.5 *Acan*^{-/-} specimens with LAMA, identifying 28 statistically significant organ volume differences. Of these, 13 had a mean volume difference between wild type and mutant of over 15% and are visible by eye (Fig. 6A), and 15 have a mean volume difference of below 15%, which makes visual identification of these differences difficult (Fig 6 suppl. 2). 12 of the significant organs are bones including all those present in the Mouse Genome Informatics (MGI) annotations (Table 3) (see Fig 6C-D for examples of identified dysmorphologies). Of these gene-level annotated organs, the specimen-level analysis assigned annotations to 15 of them (Fig. 6B) showing a considerable overlap between these two analyses. There were three significant organs highlighted by the specimen-level analysis that were not present at the gene-level, representing just four individual calls out of a total of 42 specimen-level calls, including one tail vertebra annotation. The significant

Jacobian determinant voxels indicate a smaller Meckel's cartilage in the mutants, which is not identified in previous literature or by our organ volume difference test, but is visibly different in the mutants (Fig. 6D). We did not find a significantly smaller lung volume difference for any of the lung lobes, which would indicate hypoplasia as previously reported (Houghton et al., 1989), but the inverted lung lobe labels indicate an acceptable registration accuracy (Fig 6 suppl. 1). The position of the lungs within the thoracic cavity however looks altered, possibly due to changes in the thoracic cavity size. We did not identify the remaining previously reported phenotype of tracheal cartilage morphology either.

Discussion

In this work we present a new automated computational phenotyping pipeline for dysmorphology detection in mutant mouse embryos along with a novel, highly-detailed E14.5 anatomical atlas. The LAMA pipeline is easy to install and run, and uses fewer resources than currently available software. We have undertaken validation of the pipeline and provide data on the effect of developmental substage and sample size on phenotyping results, as well as showing that LAMA can uncover previously reported and new phenotypes from E14.5 IMPC knockout mice embryos. This pipeline can accelerate the automatic analysis of 3D embryo data at this developmental stage (and will be adapted for other stages) within large scale projects such as the IMPC and smaller challenge-led projects driving forward the use of disease models in scientific discovery.

We employ an image registration pipeline to spatially normalise wild type and mutants in a similar manner to a previously reported pipeline (Wong et al., 2012; Wong et al., 2014), but with the following two key differences. Firstly, during spatial normalization of the wild type and mutant images, LAMA uses the population average as the fixed image at each stage. This contrasts with Wong et al., (2012) where the wild type and mutant images (8 + 8) are spatially normalised by a combination of initial pairwise registrations in the linear stages and then non-linear registrations. In addition, the fixed image is the average intensity image from the previous stage. Thus, the specimens from each mutant line, and its controls, exist in their own unique coordinate space. By contrast, our approach allows the initial spatial normalization of all the wild type images into the population average space once only (as opposed to registering wild types for each mutant line), and to reuse the resulting data as controls for each mutant line, enabling us to dramatically increase the number of wild type embryos used as baseline comparisons. Secondly, we have shown that even after normalizing volumetric data by whole embryo volume, E14.5 developmental

substage-specific organ volume differences are widespread across the embryo. These stage-dependent growth rate differences can lead to false positive results if the statistical model does include these substage data (Fig 3). Wong et. al., (2012) account for developmental substage variability by normalizing organ volumes and Jacobian determinants to whole embryo volume and report an organ volume standard deviation of 8-13% amongst wild type E15.5 embryos. In the current study, we sought to include as many wild type controls as possible in order to increase the power of our analyses, which increases the range of developmental substages in our control set, and we have adopted a modified linear regression model that controls for this confounding effect. Our choice of whole embryo volume (WEV), as a staging metric was motivated by the ease by which it can be calculated after embryo spatial normalization, and should correspond well to crown-rump length and whole body weight, which have been used previously to stage embryos (Dagg 1963; Peterka et al., 2002). Alternative methods that rely on the appearance of external features of the embryo (Theiler et al., 1989; Boehm et al., 2011; Geyer et al., 2017), may be more accurate, but these are not immediately amenable to automated analysis at E14.5. A promising approach to automated staging involves spatially, and temporally normalizing embryos to a four dimensional population average model embryo, with developmental stage as the fourth dimension (Wong et al., 2015). However the temporal dimension of the reported 4D model, extends only to E14.0 and it is generated from optical projection tomography images, a different modality to that used in this project. During this analysis we identified 81 organs that show a statistically significant staging effect (Table 2) which likely reflects different rates of growth of various organs at different E14.5 developmental substages, and this represents the most detailed embryo-wide data of E14.5 substage-specific organ growth rates that we are currently aware of.

Males and females are sets of animals with convenient and specific gonad differences that enabled us to test the ability of LAMA to identify anatomical differences. We found that we were able to uncover statistically significant organ volume differences in the gonads while keeping false positives low. To assess the performance on mutant data, we tested LAMA on two IMPC-generated knockout lines. The first (*Wfdc2*^{-/-}) was predicted to display specific pulmonary abnormalities and the second (*Acan*^{-/-}), a mutation shown previously to display severely-dysmorphic phenotypes across the whole of the embryo. Our automated analysis of E14.5 *Wfdc2*^{-/-} embryos revealed two significantly different organ volumes: those of the trachea and bronchi, which are novel findings for this gene. These overlap broadly with the locations of the previously reported pulmonary-specific abnormalities in *Wfdc2*^{-/-} mice, including the absence of mature club cells from the bronchi and trachea, postnatally-collapsed lung, reduced lung surfactant levels

(Nakajima et al., 2019), and alveoli abnormalities (Zhang et al. 2020). The novel phenotypes bring forward the time when gross abnormalities due to loss of *Wfdc2* first become visible during embryo development (previously postnatally), and therefore add new temporal information to the role of *Wfdc2* in pulmonary development.

In addition, we have shown that LAMA can recapitulate the majority of the previously reported *Acan*^{-/-} phenotypes, which will greatly speed up the annotation of this severely-affected mutant. There were also four significant organ volume differences for *Acan*^{-/-} that have not previously been reported. It is possible that these are novel phenotypes, but it is difficult to confirm this by looking at the CT images. It is possible that the apparent abnormality is due to proximity to actual severe dysmorphology and during the registration process were warped along with the abnormal organ. For example, two of these organs without previous reports, sympathetic ganglia and the spinal cord marginal layer, are located close to the affected vertebra.

As efforts are underway to reduce the numbers of animals used in scientific experiments we wanted to test whether LAMA could identify dysmorphology with low mutant sample number. In addition, being able to use low sample numbers would let us investigate the effects of incomplete penetrance and variable expressivity as well as providing phenotype data from mutant lines where many specimens do not reach the developmental stage being tested. To begin to answer this, in the sex difference test we show that increasing the control sample size from 8 to 46 greatly increases the power to detect anatomical differences and that by using many controls it is possible to sometimes uncover phenotype information even with a low mutant sample of one. We also show that when applying the specimen-level organ volume analysis to the *Wfdc2*^{-/-} and *Acan*^{-/-} mutant lines, we were able to generate annotations on individual specimens. These specimen level annotations mostly overlapped with those assigned at the gene-level providing support for these phenotypes. These results lead us to recommend the following for our high throughput automated screen of mutant mice: to increase sample size to include as many control specimens as possible, and that specimen-level analysis is feasible at least on a subset of organs, but likely those which are most severely affected.

The choice of registration parameters involves a compromise of balancing good registration accuracy on some organs with misregistration at others. We found that the gonad registration, for example, could be improved by removing much of the registration constraint, but this led to over-warping at the heart. One solution to this could be to use multiple sets of registration parameters,

each optimised to different parts of the atlas. Another approach could involve methods that do not require co-registration of images that could be applied to organs with a high registration error rate. This approach has been previously addressed (Yan et al., 2017; Ashish & Brusniak 2018) but only on a limited number of organs, and these methods are yet to be applied to embryonic mice. LAMA is able to perform statistical analysis on the voxel intensities of the spatially normalised images, but we found that the image data used in this study contained large differences in intensity profiles that were possibly due to the different users and imaging equipment involved in image acquisition over a number of years. For this reason, we have concentrated our current analysis of mutant lines on organ volume differences and Jacobian determinant analysis, which are both more robust to varying intensity profiles. Future work may look towards employing more sophisticated image normalization methods and exploring the analysis of other image features, such as textures, that may be less susceptible to intensity profile differences. Future work will include the adaption of the pipeline to other key developmental stages. E18.5 is a key developmental stage that we are currently working towards as it is important for the analysis of gene mutations that result in perinatal lethality and subviability. Earlier stages may also be amenable such as E12.5, but the earlier stages such as E9.5 may prove difficult for a registration-based approach due to the rapid developmental changes at this time point and the fact that mutations that cause lethality at this stage cause very extreme dysmorphology.

To conclude, we provide a new, easy-to-use pipeline for the automated analysis of E14.5 mouse embryos, along with a highly-detailed atlas that will be useful for both manual and automated studies of mutant mice. We have provided information on the differential growth rate of organs within the E14.5 developmental stage, shed light on the optimal number on the samples to use in mouse embryo phenotyping studies and have shown that LAMA can detect sex-specific differences and abnormal anatomy in E14.5 mutant embryos.

Materials and methods

Mice

All animals were housed and maintained in the Mary Lyon Centre, MRC Harwell Institute under specific pathogen-free (SPF) conditions, in individually ventilated cages adhering to environmental conditions as outlined in the Home Office Code of Practice. The Acan^{tm1b} allele was obtained by cre deletion of C57BL/6N-Acan^{tm1a(EUCOMM)Hmgu/H}

(EM:10224) mice as described in (Biriling et al., 2019). Homozygous mutants are named *Acan*^{-/-} here. The C57BL/6NTac-Wfdc2^{em1(IMPC)H/H} (EM:11407, homozygous mutants named *Wfdc2*^{-/-} here) was obtained by genome editing as described in (Mianné et al., 2017). Lines were maintained by crossing heterozygous animals with inbred C57BL/6N wildtype animals. Mice were euthanised by Home Office Schedule 1 methods.

Micro-CT imaging of whole embryos

14.5 days post coitum (E14.5) female mice were sacrificed by cervical dislocation and the uterine horns removed into ice-cold phosphate buffered saline (PBS). Embryos were extracted and a piece of yolk sac collected for genotype analysis. Embryos were fixed in 4% paraformaldehyde (PFA) at 4°C and left overnight. After fixation, the samples were washed and stored in PBS at 4°C. For staining, samples were rinsed in dH₂O for 10 minutes before being submerged in 50% Lugol's solution and protected from light. Embryos were then left in the contrast agent for 2 days. Following staining, embryos were washed in dH₂O for at least one hour, embedded in 1% agarose (in dH₂O) and left at room temperature for a minimum of two hours.

High resolution micro-CT images (SkyScan 1172, Bruker) of agarose-embedded embryos were acquired at a source voltage of 70 kV, with the current set at maximum (~100 mA). Specimens were imaged, in a standard orientation, at 3 µm with a 0.5 mm aluminum filter. X-ray projections were acquired at 0.25° increments, and reconstructed using the Feldkamp algorithm (Feldkamp et al. 1984) provided by NRecon (Bruker). Ring artefact corrections were applied as necessary. Reconstructions were automatically cropped to remove background and scaled to 14 µm isotropic voxels using the HARP software (Brown et al., 2016).

Registration pipeline implementation

The image registration pipeline was written in the Python programming language (Python 3.6), adapting a modular design that allows for individual components (registration, inversion, statistics etc.) to be run either sequentially or independently using simple TOML configuration files. Individual image registrations are performed using the elastix toolkit (S. Klein *et al.*, 2010; Shamonin *et al.*, 2013). The linear model analysis is implemented in R. All code is available on Github (<https://github.com/mpi2/LAMA/releases/latest>) and is tested to work on Ubuntu 18.04, and Windows. To make the installation of LAMA as easy as possible and to help data

reproducibility, LAMA is available via the PyPi Python package repository (lama-phenotype-detection).

Population average construction

Micro-CT images from 14 specimens of mixed sex were used in the creation of the population average through a multi-stage and resolution registration process. The initial fixed image was chosen at random from the input images and all other images are rigidly-registered towards it. The registered output images were averaged creating a rigidly-aligned blurry average. The outputs from the rigid registration are affinely-registered towards the blurry average to account for differences in overall size. The blurry affine average from this step was then used as the fixed image for deformable registration, which allows non-linear transformations. The deformable B-spline registration process contained five resolutions with decreasing control point spacing and blurring at each resolution, with a final grid spacing of 8 voxels, sequentially aligning smaller anatomical structures (Fig. 1A). The parameter file for our population average can be found here (https://github.com/mpi2/LAMA/configs/301015_pop_avg.toml) (link to population average when available).

Image segmentation/E14.5 Atlas creation

Key anatomical structures within the E14.5 population average were identified manually by referencing the online digitised mouse atlas (Graham et al., 2015), and which itself is based on *The Atlas of Mouse Development* (Kaufman, 1992). Moreover, structures that could be identified were restricted to those that showed good contrast and resolution within the population average. ITKSnap was used (Yushkevich et al., 2006; www.itksnap.org) to create the segmentations, and dependent on size and complexity of an anatomical structure, were produced using semi-automated and manual methods and combined into a single label file until 184 structures were segmented at E14.5. These structures were then merged with some previous segmentations of brain structures derived from an E15.5 atlas (Wong et al., 2012), to give a total of 184 anatomical components. Small, spindly labels in the atlas were flagged by taking the mean of a distance transform for each label, and flagging labels with a value < 1.5. The atlas and associated metadata file are available (link when available).

Generation of data for phenotype detection

Baseline and mutant specimens are registered onto the previously created population average. The outputs of this registration include the non-affine spatial transformations, co-registered images and the Jacobian determinants $\det(J_F)$, a scalar field which describes the local volume change in each voxel. Statistical analysis of these outputs (described below) produce statistical parametric heat maps that can be overlaid onto the population average applying the computed spatial transformations, or superimposed onto the input images applying the inverse of the spatial transformations. The statistical parametric heat maps can be viewed with the Volume Phenotype Viewer (VPV) described previously (Brown *et al.*, 2016; <https://github.com/mpi2/vpv>)

A typical project involves generating baseline data from wild type animals, before running all the mutant lines. For each line, the mutant and all the baseline data at each voxel, within a mask region, are fitted to a linear model with genotype and staging metric (whole embryo volume) as fixed effects, returning a t-statistic and p-value for the genotype effect.

Statistical analysis

Multiple linear regression analysis was conducted in R (R Core Team, 2018) using the `lm` function from the MASS package (Venables & Ripley, 2002). Benjamin-Hochburg FDR (BH-FDR) correction was done using the `padjust` R package (Benjamin & Hochburg 1995).

Voxel-level data

Jacobian determinants, generated at each voxel within the population average mask, provide information about how the registration has behaved at discrete locations in an image. The scalar value of the Jacobian determinant at a given location is the factor by which that region has expanded or shrunk in volume during registration. This approach, known as tensor-based *morphometry*, can be used to reveal biologically significant localised shape or size changes within a population (Ashburner & Friston, 2000). As well as the Jacobian determinants, the raw registered intensity images are analysed. To account for image intensity differences between the input images, LAMA includes an optional image intensity normalization step. Images are linearly normalised to the mean intensity of voxels across all the inputs or alternatively a region of interest (ROI) specified in the target coordinate space. To account for small registration inaccuracies, a Gaussian blur of full-width-half-maximum (FWHM) 100 μm is applied to voxel-level data. Each voxel is fitted to a linear model (*R* notation: $\text{voxel} \sim \text{genotype} + \text{whole embryo volume}$). We use

whole embryo volume as a proxy for developmental stage, and the addition of it as a fixed effect models out changes that are due to developmental stage only. To account for multiple testing, the resulting p-value maps are corrected using the Benjamini Hochberg method. The final parametric heat maps are made by thresholding the t-static volume at $q > 0.05$. This is output as a volume that can be overlaid onto the target or registered image in VPV.

Whole organ volume analysis

The atlas is inverted back onto the rigidly-aligned volumes, after registration, which effectively segments these volumes. These segmentations are used to calculate organ volumes for each specimen as well as providing a visual atlas for identifying structures in the original inputs and determining registration accuracy. The organ volume data is much less complex than the voxel-based data (the number of organs vs number of voxels) allowing us to employ a more robust method for multiple testing correction described previously by Hrabě de Angelis et al., (2015). To summarize, organ-specific null distributions are generated by sampling synthetic mutants from the baseline data in such a way as to match the distribution of mutant specimens per line. Synthetic mutants are fitted to a linear model as described previously (*organ volume ~ genotype + whole embryo volume*) and the genotype effect p-values are stored. Alternative distributions are made by collecting genotype effect p-values from testing the real mutants of each line. To obtain a data set-wide p-value threshold per organ, p-values for the organ are ranked and a descending p-value search is conducted starting at $p = 0.05$ until a value is found where the proportion of alternative p-values under the threshold divided by the proportion of null p-values is < 0.05 . Mutant p-values below this threshold are assigned as significant, which sets the organ-specific FDR to 5%.

Developmental Substage analysis

To simulate the effect of analysing mutant lines containing specimens at early or late E14.5 developmental substage, two datasets were created from 99 wild type specimens by relabeling subsets based on the calculated whole embryo volume (WEV) of each specimen. In the first set, eight specimens with the lowest WEV were relabelled as mutant. In the second set, eight specimens with the highest WEV were relabelled as mutant. To test whether a substage effect could be detected, the organ volume or Jacobian determinants data points from each set were then analysed by using the linear regression models (*organ volume ~ genotype*) or (*Jacobian determinant ~ genotype*). To test whether including WEV as a fixed effect could remove the

substage confound, the organ volume or Jacobian determinants from each set were then analysed by using the linear regression models (*organ volume* ~ *genotype* + *WEV*) or (Jacobian determinant ~ *genotype* + *WEV*). The resulting organ volume, or Jacobian determinant, genotype effect p-values were corrected for multiple testing using the Benjamini Hochberg method.

To assess the contribution of developmental substage on local relative volume differences across the embryo, Jacobian determinants or organ volumes from 99 wild type specimens were analysed by using the linear regression model (*organ volume* ~ *whole embryo volume*). The resulting organ volume, or Jacobian determinant, genotype effect p-values were corrected for multiple testing using the Benjamini Hochberg method.

Detection of sex-specific differences

The spatially normalised data from 93 wild type specimens were split by sex. For each experiment, a different number of males and females were chosen at random without replacement and the females were relabelled as mutants. These data were then analysed by using the linear model (*organ volume* ~ *sex* + *whole embryo volume*). For organ volume analysis, organ-specific p-value thresholds were generated using the permutation based method as described above. Only experiments where combinations of male and female sample size allowed at least 500 unique null permutations were included. Each experiment was repeated 50 times to gain a robust measure of the effect of sample number. For each set of experiments, the organs that were reported as having a significant organ volume difference were recorded.

For the Jacobian determinant analysis q-values were generated from the resulting p-values using the Benjamini Hochberg method and threshold to $q < 0.05$. False positive rates for organs other than gonad or lens of the eye was calculated by dividing the total number of times a given organ passed the organ-specific significance threshold by the total number of tests performed.

Variable penetrance and low *N*

To identify potentially variable expressivity or incomplete penetrance of organ volume phenotypes, specimens were treated as for the line-level analysis above. With the number of mutants reduced to one the FDR threshold was increased to 20%. The voxel-level data is similarly processed as the line-level voxel data and the voxels are thresholded to an FDR of 5%.

Optimization and quality control

At each stage of the registration process, the image similarity metric output by elastix is plotted against iteration number for each spatial resolution allowing the user to select an optimal number of iterations per stage.

Image registration can sometimes fail to produce acceptable results, for example when the moving image is over fitted to the fixed image, producing unrealistic warping of images. To check for issues such as these, after each registration stage an additional HTML report is generated which contains mid-sagittal slices for each registered image for rapid quality control. Another issue that can be encountered using bspline-based image registration is folding of the deformation field which prevents topology preservation, reducing the accuracy of the propagation of labels from the atlas to the specimen images. This problem can be detected by the presence of negative Jacobian determinants. In the case of negative Jacobian determinants, instead of outputting log-transformed Jacobian determinants, only the negative Jacobian determinant regions are displayed, allowing the user to quickly identify problematic regions within the registered images. The registration stage of the analysis is the most time-consuming part of the pipeline, and so it is a requirement to be able to optimize registration parameters, especially within a high throughput context.

Comparing phenotypes to known phenotypes

Tables of known phenotypes were generated by querying MGI phenotype pages for the gene of interest (Acan: www.informatics.jax.org/marker/phenotypes/MGI:99602, Wfdc2: www.informatics.jax.org/marker/phenotypes/MGI:1914951). Only phenotypes generated from a homozygous null strains were kept to aid in the comparison. Duplicate and redundant phenotypes (e.g. *abnormal bone structure* if more specific bone phenotypes were present). Also removed were phenotypes that might not translate to a gross anatomical dysmorphism that could be potentially detected by LAMA (e.g. deafness).

Figures

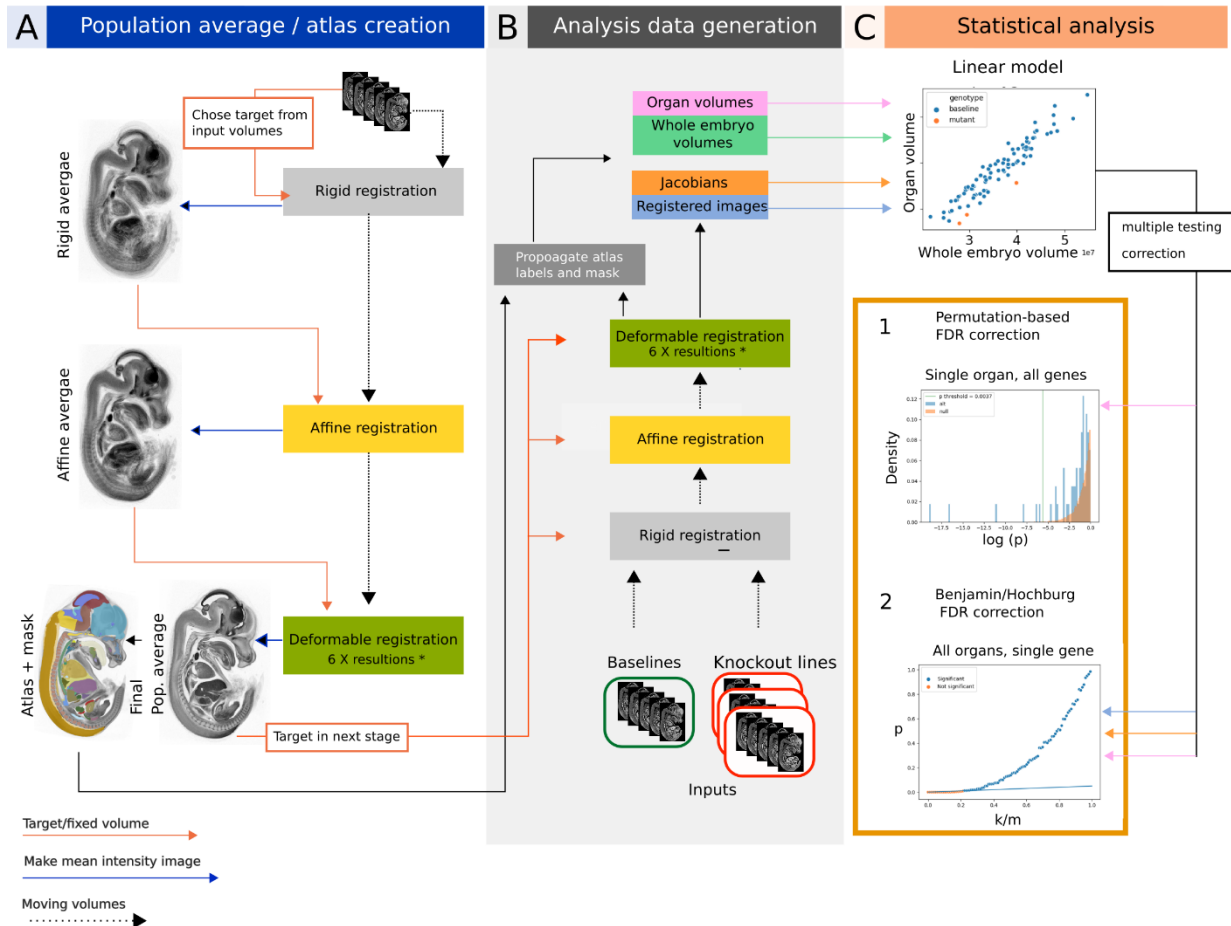


Figure 1. LAMA pipeline workflow. (A) Population average construction. An initial target from the inputs is used to rigidly align all inputs creating a rigid population average. This is repeated with affine registration and deformable registration at each stage using the population average as the target from the subsequent stage. (B) Generation of data for phenotype detection. The same process as in A, but each stage uses the population average form A as the target. Jacobian determinant volumes, registered images and organ volumes are statistically analysed. (C) Organ volumes (shown here) or voxel values are fitted to a linear model by genotype and whole embryo volume. The resulting genotype effect p-values are corrected for multiple testing either by permutation-based FDR correction (option 1, organ volumes only) the orange histogram is the null distribution from permuting the wild type organ volumes and the blue histogram is the alternative distribution derived from testing the organ volume from all mutant lines tested. The vertical green line indicates the calculated p-value threshold for this organ, with values lower than this annotated as significant. Option 2 is to apply Benjamini Hochberg FDR correction for voxel-level or organ volume data. k/m = p-value rank divided by number of values. The straight blue line indicates the threshold under which values are annotated as significant.

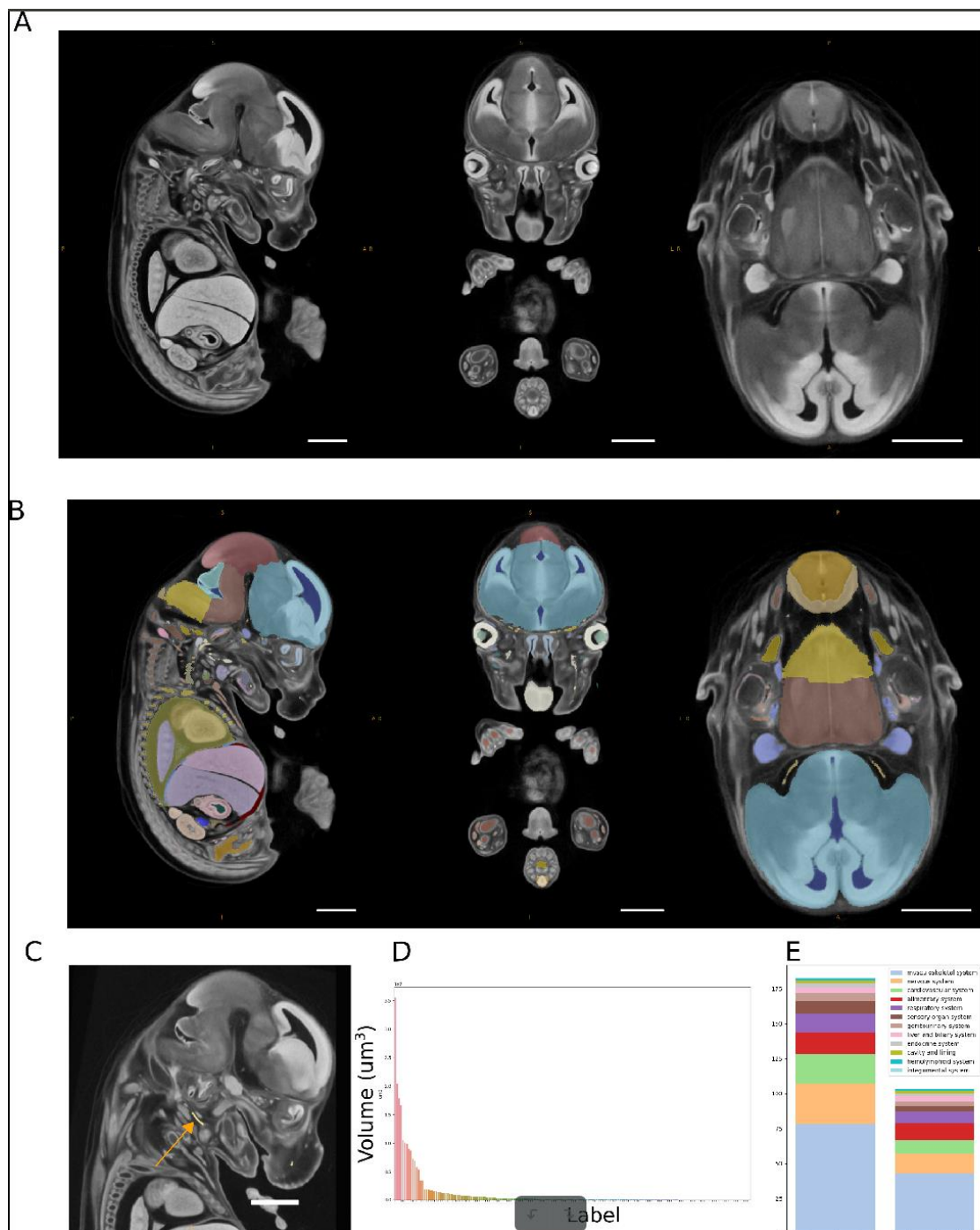


Figure 2. E14.5 atlas creation. (A) Population average created from 16 wild type mixed sex E14.5 C57BL/6 mice. (B) E14.5 atlas consisting of 184 individual structures overlaid on the population average. (C) An example of a small size-excluded label, external carotid artery (arrow). Scale bar = 1mm. (D) Plot showing the organ volumes for each of the 184 E14.5 atlas labels. (E) Counts for each high level organ system applied to labels in the E14.5 atlas.

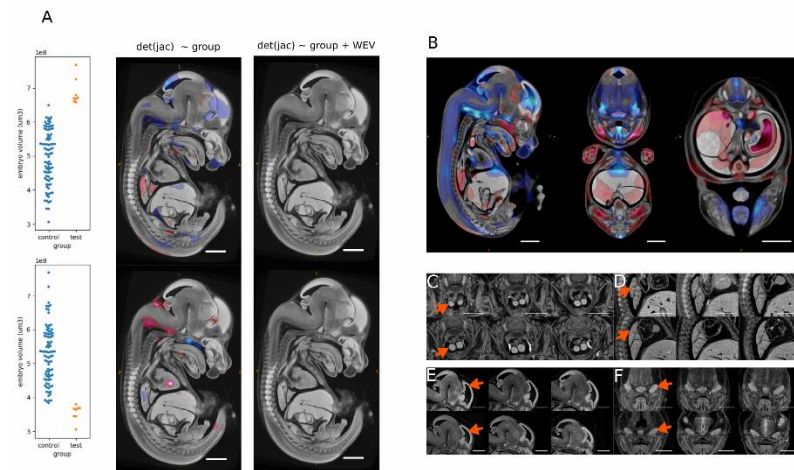


Figure 3. Effect of E14.5 developmental substage on local volume changes detected by LAMA. (A) Simulation of the analysis of mutant lines with large (top panel) or small (bottom panel) wild types relabelled as mutants. Genotype effect t-statistics where q (FDR-corrected p-values) < 0.05 were overlaid on the E14.5 population average. Left panel shows results from the model: *Jacobian determinant* \sim *genotype* with red voxels highlighting regions that are significantly larger in the test group and blue voxels highlighting regions that are significantly smaller in the test group. Right panel shows results from model: *Jacobian determinant* \sim *genotype* + *WEV*, where no voxels passed the $q < 0.05$ threshold. (B) Result from linear model: *Jacobian determinant* \sim *WEV*, indicating regions where the normalised organ volume is directly (red) or inversely proportional (blue) to WEV. (C-F) Illustrative examples of differences in organ size relative to embryo volume. Each panel shows three representative wild types from the smallest set (top) and largest set (bottom) of specimens. Images are affinely registered towards the population average to account for overall embryo volume. Arrows indicate relevant anatomy. (C-D) Larger relative thymus size and lung volume (D), (E-F) Smaller relative lateral ventricles and trigeminal gland volume. Scale bar = 1mm.

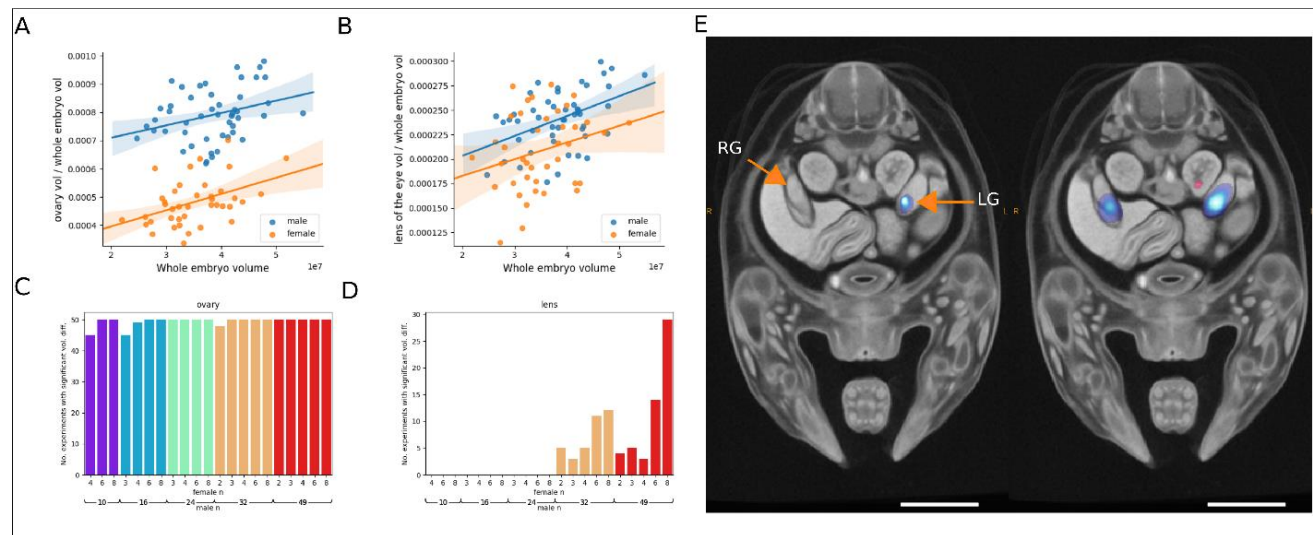


Figure 4. Identification of sex differences in wild type E14.5 mice. (A) A series of statistical tests were carried out with various combinations of male and female wild type sample size (with females relabelled as 'mutant'). Each test was repeated 50 times with randomly selected specimens. The values reported are the number of times the gonad volume was reported significantly different ($p < \text{organ } p \text{ threshold}$) in each set of tests. (B) Example results of Jacobian determinant analysis using 8 females and 8 males (left) and 8 females and 46 males (right). Scale bar = 1mm. RG: right gonad, LG = left gonad.

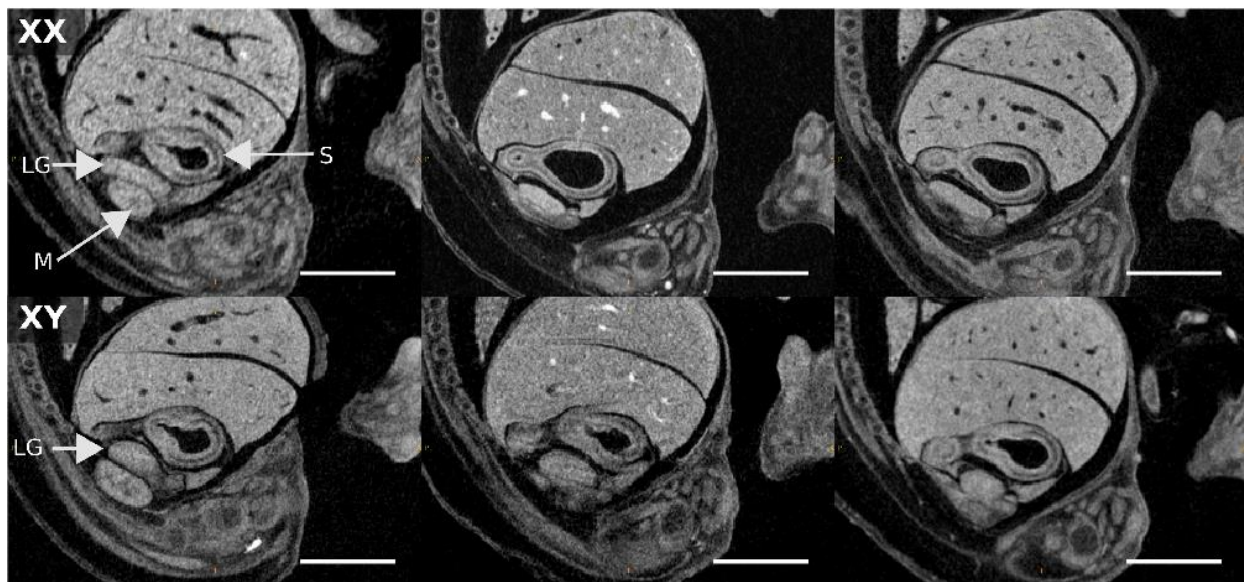


Figure 4, supplementary 1. Wild type embryos gonad examples.

Sagittal sections of rigidly-aligned wild type embryo images showing the left gonad (LG). The stomach (S) and metanephros (M) are also indicated. Females (top) and males (bottom). Scale bar = 1mm.

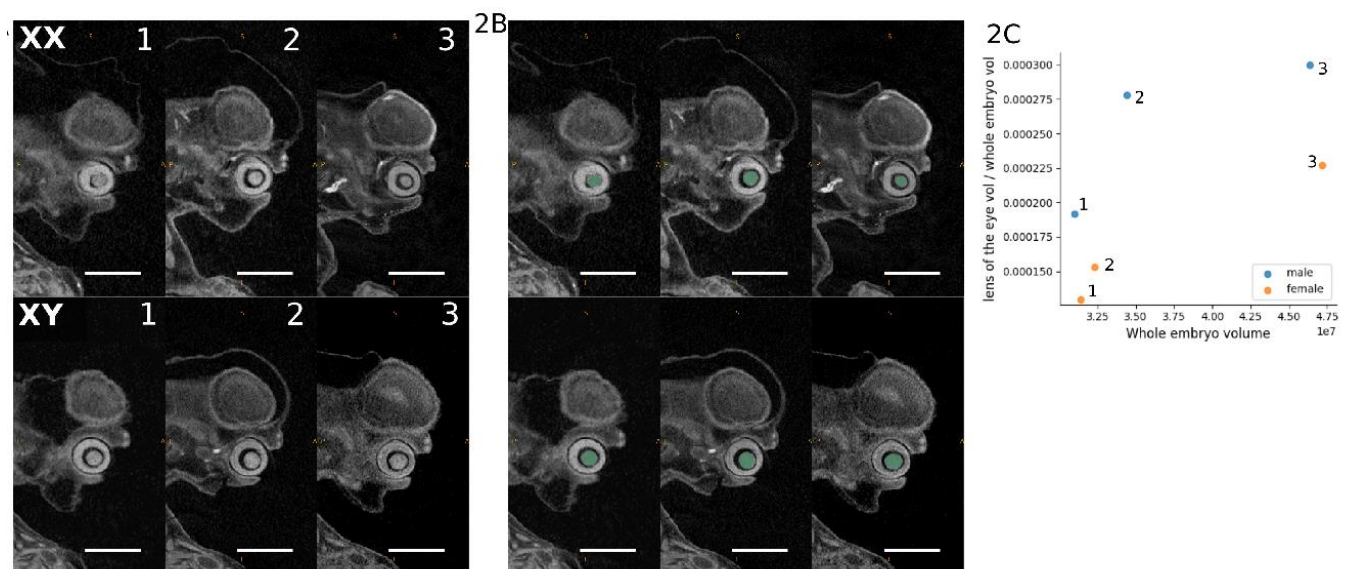


Figure 4, supplementary 2. Differences in wild type lens volume

Representative sagittal sections of affinely-registered (normalised for overall embryo size) wild type specimens illustrating difference in lens sizes between females (top rows) and males (bottom rows). (A) Unlabeled affinely-registered images (B) Affinely-registered images overlaid with inverted lens label (green). (C) Plot of whole-embryo normalised organ volume against whole embryo volume.

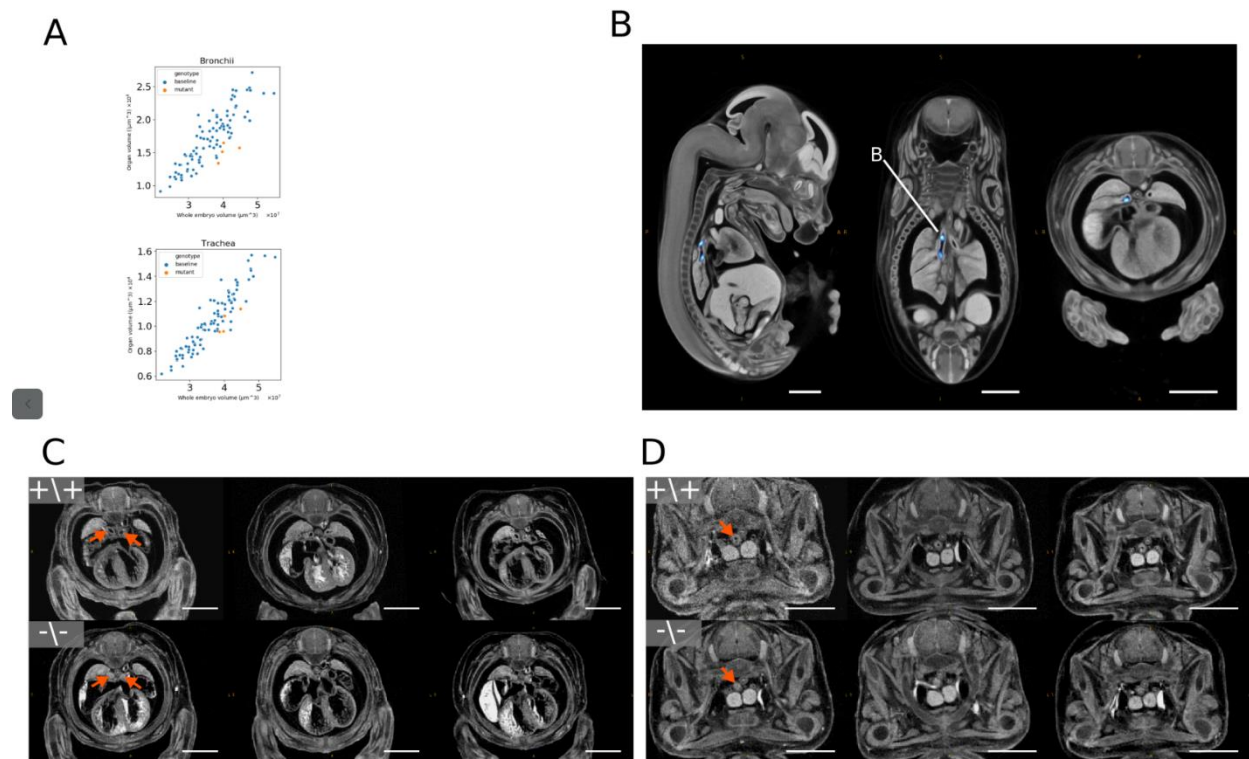


Figure 5. Automated identification of pulmonary system phenotypes in a *Wfdc2* knockout mouse line.

(A) Organ volume plots from statistically significant organs. (B) Jacobian determinant analysis t-statistics (FDR corrected to $q < 0.1$) overlaid on the E14.5 population average. Blue regions indicate smaller bronchi in the mutants. (C-D) Axial slices of rigidly-aligned *Wfdc2* mutants (bottom) and whole embryo volume-matched wild type rigidly-aligned specimens (top). Arrows indicate the location of affected organs (C: bronchi, D: trachea).

Scale bar = 1mm.

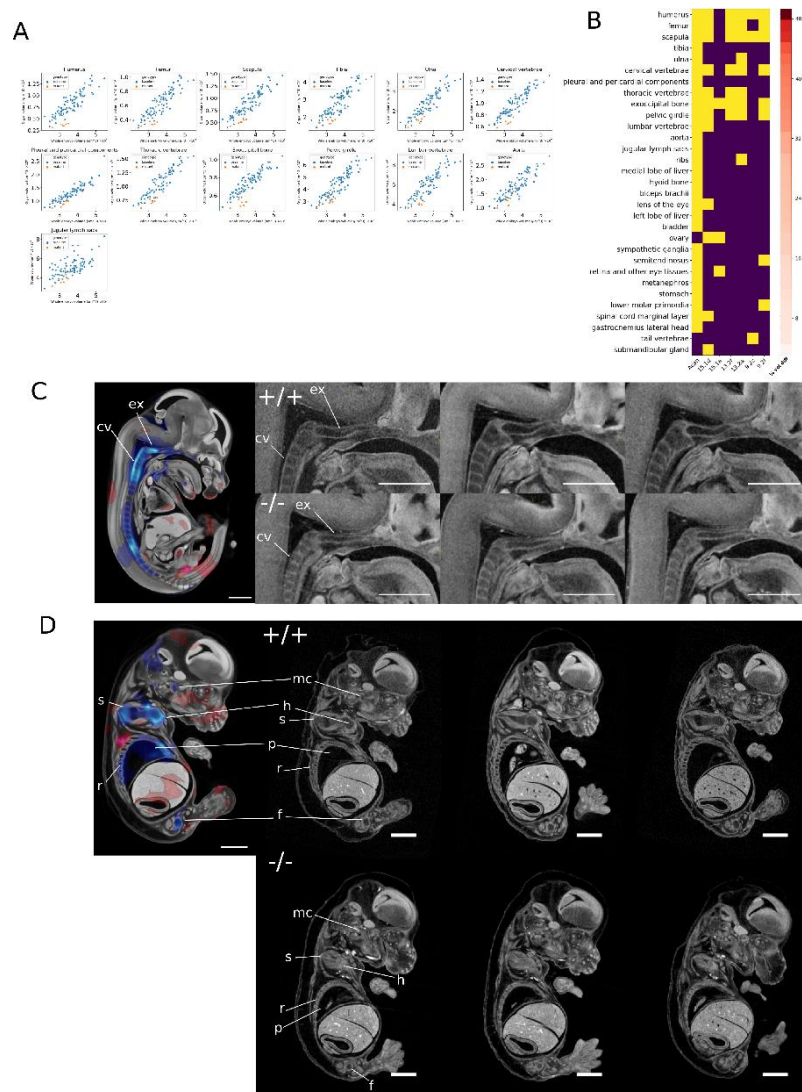


Figure 6. Analysis of *Acan*^{-/-} mutants by LAMA.

(A) Scatter plots showing organ volume in voxels vs whole embryo volume in voxels for organs with a significant genotype effect. Only organs with a mean whole embryo-normalised volume difference, between controls and mutants, of > 15% shown (B) Comparison of line-level (*Acan*) and individual specimen-level results. Yellow cells indicate statistically significant organ volume differences for the genotype effect. The column '% vol diff' indicates the mean organ volume difference (whole embryo normalised) between wild type and *Acan*^{-/-} mutants (C-D) Illustrative 2D sagittal slices highlighting identified dysmorphology. Left panels: Jacobian determinant t-statistics (FDR corrected to $q < 0.05$) overlaid onto the E14.5 population average. Right panels: rigidly-aligned *Acan*^{-/-} specimens (bottom), and stage-matched wild type rigidly aligned specimens (top)

Scale bar = 1mm

Legend: mc: Meckels's cartilage, ex: exoccipital bone, cv: cervical vertebrae, h: humerus, s: scapula, p: pleural and pericardial components, r: ribs, f: femur

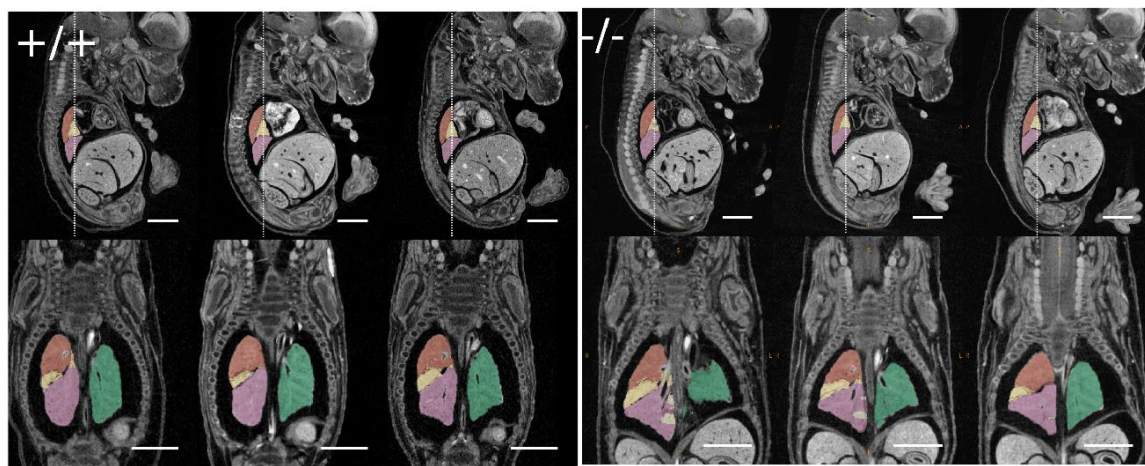


Fig 6. Supplementary 1. Comparison of wild type and *Acan*^{-/-} mutant lungs

Rigidly-aligned specimens overlaid with individual lung lobe labels inverted from the atlas showing mutants (left) and stage-matched wild types (right). Top row are sagittal sections, bottom row are coronal sections. The dotted line in the sagittal sections corresponds to the section chosen for the coronal view.

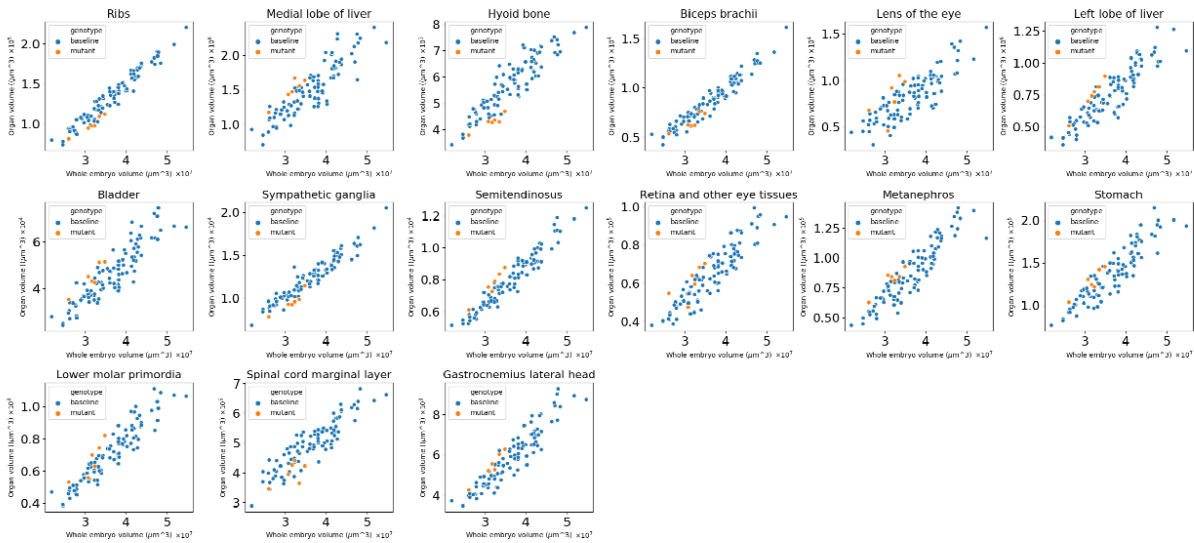


Fig 6. Supplementary 2. Significant organ volume differences in *Acan*^{-/-} null mutants.

Scatter plots showing organ volume in voxels vs whole embryo volume in voxels for organ with a significant genotype effect. Organs highlighted here have a mean whole embryo-normalised difference of less than 15% between wild types and mutants. Blue markers are wild type controls, orange markers are mutants.

label	label name	EMAPA Anatomical structure	EMAPA organ system	Used in current analysis
3	forebrain	16895	nervous system	TRUE
21	cerebellar primordium	17787	nervous system	TRUE
23	midbrain	10974	nervous system	TRUE
24	pons	17963	nervous system	TRUE
25	medulla oblongata	17950	nervous system	TRUE
26	retina and other eye tissue	17160	sensory organ system	TRUE
27	lens of the eye	17838	sensory organ system	TRUE
35	submandibular gland	18812	alimentary system	TRUE
36	Meckel's cartilage	17635	alimentary system	TRUE
37	mandible	18290	musculoskeletal system	TRUE
38	lower molar primordia	17621	alimentary system	TRUE
40	stomach	17621	alimentary system	TRUE
41	stomach lumen	18395	alimentary system	TRUE
42	foregut	18548	alimentary system	TRUE
43	mid hindgut	No term	alimentary system	TRUE
44	rectum	17890	alimentary system	TRUE
45	pancreas	17903	alimentary system	TRUE
47	medial pterygoid muscle	25141	musculoskeletal system	TRUE
48	masseter muscle	25135	musculoskeletal system	TRUE
49	tongue	17885	alimentary system	TRUE
51	vomeronasal organ	17612	sensory organ system	TRUE
52	medial lobe of liver	18318	liver and biliary system	TRUE
53	left lobe of liver	18307	liver and biliary system	TRUE
54	right lobe of liver	18311	liver and biliary system	TRUE
55	caudate lobe of liver	18311	liver and biliary system	TRUE
56	bladder	18321	genitourinary system	TRUE
58	metanephros	17373	genitourinary system	TRUE
61	left lung	17653	respiratory system	TRUE
62	right lung cranial lobe	17993	respiratory system	TRUE
63	right lung middle lobe	17997	respiratory system	TRUE
64	right lung caudal lobe	17980	respiratory system	TRUE
65	right lung accessory lobe	17981	respiratory system	TRUE
66	bronchus	32089	respiratory system	TRUE
67	trachea	18053	respiratory system	TRUE
70	nasal cavity epithelium	17605	sensory organ system	TRUE
71	future anterior pituitary	30996	endocrine system	TRUE
72	brain ventricles	32474	nervous system	TRUE
73	trigeminal V1 ganglia	18797	nervous system	TRUE
77	superior ganglion of gloss	17154	nervous system	TRUE
78	inferior ganglion of gloss	17153	nervous system	TRUE
79	superior cervical sympathetic	18441	nervous system	TRUE
80	sympathetic ganglia	17157	nervous system	TRUE
83	spinal cord mantle layer	17961	nervous system	TRUE
84	spinal cord marginal layer	17963	nervous system	TRUE
93	heart	10105	cardiovascular system	TRUE
96	aorta	18061	cardiovascular system	TRUE
97	ascending aorta	17014	cardiovascular system	TRUE
98	descending aorta	18060	cardiovascular system	TRUE
99	aortic arch	17611	cardiovascular system	TRUE
110	ductus arteriosus	17858	cardiovascular system	TRUE
111	ductus venosus	17343	cardiovascular system	TRUE
113	jugular lymph sacs	18249	cardiovascular system	TRUE
114	inferior vena cava	18416	cardiovascular system	TRUE
115	superior vena cava	18417	cardiovascular system	TRUE
118	adrenals	18420	endocrine system	TRUE
119	spleen primordium	18335	alimentary system	TRUE
120	thymus primordium	17923	hemolymphoid system	TRUE
122	metacarpals	30136	musculoskeletal system	TRUE
124	radius	19303	musculoskeletal system	TRUE
125	ulna	19104	musculoskeletal system	TRUE
126	humerus	19105	musculoskeletal system	TRUE
127	triceps brachii	19111	musculoskeletal system	TRUE
128	biceps brachii	19108	musculoskeletal system	TRUE
130	flexor carpi ulnaris	30138	musculoskeletal system	TRUE
142	pronator teres	No term	musculoskeletal system	TRUE
144	teres major	19110	musculoskeletal system	TRUE
145	latissimus dorsi	19178	musculoskeletal system	TRUE
146	hindlimb phalanges	30243	musculoskeletal system	TRUE
147	metatarsals	30139	musculoskeletal system	TRUE
148	tarsals	25072	musculoskeletal system	TRUE
149	tibia	19341	musculoskeletal system	TRUE
150	tibia	30179	musculoskeletal system	TRUE
151	femur	19343	musculoskeletal system	TRUE
152	ilacrus	19104	musculoskeletal system	TRUE
153	quadratus femoris	30241	musculoskeletal system	TRUE
155	poos major	18169	musculoskeletal system	TRUE
161	semimembranosus	30242	musculoskeletal system	TRUE
162	semitendinosus	30782	musculoskeletal system	TRUE
164	tibialis anterior	30660	musculoskeletal system	TRUE
170	vastus intermedius	30243	musculoskeletal system	TRUE
172	vastus lateralis	30244	musculoskeletal system	TRUE
173	gastrocnemius lateral head	30249	musculoskeletal system	TRUE
174	gastrocnemius medial head	30250	musculoskeletal system	TRUE
175	flexor digitorum longus	30248	musculoskeletal system	TRUE
176	popliteus	30240	musculoskeletal system	TRUE
180	pectoralis	No term	musculoskeletal system	TRUE
183	cervical vertebrae	17674	musculoskeletal system	TRUE
184	thoracic vertebrae	18011	musculoskeletal system	TRUE
185	lumbar vertebrae	18007	musculoskeletal system	TRUE
186	sacral vertebrae	18009	musculoskeletal system	TRUE
187	tail vertebrae	18074	musculoskeletal system	TRUE
188	fibula	18010	musculoskeletal system	TRUE
189	exoccipital bone	18708	musculoskeletal system	TRUE
191	hyoid bone	18050	musculoskeletal system	TRUE
195	epiglottis	18289	alimentary system	TRUE
196	clavicle	18721	musculoskeletal system	TRUE
197	scapula	18722	musculoskeletal system	TRUE
198	pelvic girdle	18028	musculoskeletal system	TRUE
199	diaphragm	17703	respiratory system	TRUE
200	cervical dig	10060	nervous system	TRUE
205	ovary	17962	genitourinary system	TRUE
206	peritoneal cavity	18138	cavity and lining	TRUE
207	pleural and pericardial cavities	No term	cavity and lining	TRUE
31	semicircular canal	32832	sensory organ system	FALSE
32	endolymphatic ducts	18980	sensory organ system	FALSE
33	sublingual duct	18809	alimentary system	FALSE
34	submandibular duct	35833	alimentary system	FALSE
39	lower incisor primordia	17918	alimentary system	FALSE
40	thyroid	17008	alimentary system	FALSE
56	genioglossus muscle	18276	musculoskeletal system	FALSE
57	bladder lumen	18321	genitourinary system	FALSE
59	ureteric trunk	28494	genitourinary system	FALSE
60	nephric ducts	18577	genitourinary system	FALSE
68	optic II nerve	17575	sensory organ system	FALSE
69	eye muscles	35335	sensory organ system	FALSE
72	infundibulum of pituitary	35998	endocrine system	FALSE
74	facial VII ganglia	17969	nervous system	FALSE
76	vestibulocochlear VIII ganglia	17571	nervous system	FALSE
81	oculomotor III nerve	17574	nervous system	FALSE
82	trigeminal V1 nerves	17798	nervous system	FALSE
83	trigeminal V2 nerves	17799	nervous system	FALSE
84	trigeminal V3 nerves	17800	nervous system	FALSE
85	vestibulocochlear VIII nerve	17801	nervous system	FALSE
86	anterior left vagal X	25300	nervous system	FALSE
87	posterior right vagal X	17275	nervous system	FALSE
88	spinal accessory XI nerve	17267	nervous system	FALSE
89	hypoglossal XII nerves	17269	nervous system	FALSE
90	spinal nerves	18989	nervous system	FALSE
91	left recurrent laryngeal nerve	25300	nervous system	FALSE
92	right recurrent laryngeal nerve	17275	nervous system	FALSE
100	anterior cerebral arteries	17854	cardiovascular system	FALSE
101	middle cerebral arteries	18239	cardiovascular system	FALSE
102	posterior cerebral arteries	17861	cardiovascular system	FALSE
103	common carotid arteries	17855	cardiovascular system	FALSE
104	external carotid arteries	18511	cardiovascular system	FALSE
105	internal carotid arteries	20328	cardiovascular system	FALSE
106	communicating arteries	18598	cardiovascular system	FALSE
107	pulmonary arteries	17008	respiratory system	FALSE
108	umbilical arteries	18331	cardiovascular system	FALSE
109	vertebrobasilar artery	38022	cardiovascular system	FALSE
112	jugular veins	18638	cardiovascular system	FALSE
116	pulmonary veins	18645	respiratory system	FALSE
117	subclavian veins	17621	cardiovascular system	FALSE
121	forelimb phalanges	30590	musculoskeletal system	FALSE
123	carpals	25098	musculoskeletal system	FALSE
129	flexor carpi radialis	30137	musculoskeletal system	FALSE
131	flexor digitorum superficialis	No term	musculoskeletal system	FALSE
132	flexor digitorum profundus	No term	musculoskeletal system	FALSE
133	flexor digitorum profundus	30200	musculoskeletal system	FALSE
134	flexor digitorum profundus	30201	musculoskeletal system	FALSE
135	flexor digitorum superficialis	No term	musculoskeletal system	FALSE
136	extensor carpi ulnaris	30130	musculoskeletal system	FALSE
137	extensor digitorum lateralis	30132	musculoskeletal system	FALSE
138	extensor digitorum communis	30131	musculoskeletal system	FALSE
139	extensor pollicis longus	No term	musculoskeletal system	FALSE
140	extensor carpi radialis longus	No term	musculoskeletal system	FALSE
141	pronator quadratus	30206	musculoskeletal system	FALSE
143	palmaris longus	30205	musculoskeletal system	FALSE
154	biceps femoris	35170	musculoskeletal system	FALSE
156	poos minor	18170	musculoskeletal system	FALSE
157	ligamentum patellae	18569	musculoskeletal system	FALSE
158	peroneus longus	30233	musculoskeletal system	FALSE
159	peroneus digiti quarti	30200	musculoskeletal system	FALSE
160	caudofemoralis	30226	musculoskeletal system	FALSE
163	gluteus maximus	18327	musculoskeletal system	FALSE
165	tibialis posterior	30205	musculoskeletal system	FALSE
166	gracilis anticus	30232	musculoskeletal system	FALSE
167	gracilis posticus	30233	musculoskeletal system	FALSE
168	pectineus	30238	musculoskeletal system	FALSE
169	extensor digitorum longus	35330	musculoskeletal system	FALSE
171	vastus medialis	30243	musculoskeletal system	FALSE
177	plantaris	30234	musculoskeletal system	FALSE
178	soleus	35786	musculoskeletal system	FALSE
179	tensor aponeurosis plantae	30205	musculoskeletal system	FALSE
181	trapezius	18183	musculoskeletal system	FALSE
182	serratus anterior	35769	musculoskeletal system	FALSE
189	basisphenoid bone	18706	musculoskeletal system	FALSE
192	thyroid cartilage	18698	respiratory system	FALSE
193	cricoid cartilage	18696	respiratory system	FALSE
194	arytenoid cartilage	18695	respiratory system	FALSE
204	tail dig	18372	nervous system	FALSE

Table 1. E14.5 Atlas label information

label: label number in atlas images, label_name: the descriptive name, EMAPA Anatomical structure: associated EMAPA anatomical term, EMAPA organ system: the top level EMAP organ system terms, Used in current analysis: If FALSE not used in the current study.

label	label_name	p	q	t	label	label_name	p	q	t
128	biceps_brachii	1.3E-14	1.3E-12	9.18988	71	future_anterior_pituitary	0.00054	0.00102	-3.5862
35	submandibular_gland	9.9E-14	5.0E-12	8.76173	147	metatarsals	0.00056	0.00102	-3.5791
172	vastus_lateralis	1.8E-13	5.0E-12	8.63642	26	retina_and_other_eye_tissues	0.00063	0.00115	3.53917
174	gastrocnemius_medial_head	2.0E-13	5.0E-12	8.62129	164	tibialis_anterior	0.00068	0.0012	3.51948
37	mandible	6.2E-13	1.3E-11	8.37966	45	pancreas	8.40E-04	0.00147	3.45413
73	brain_ventricles	9.0E-13	1.6E-11	-8.3023	78	inferior_ganglion_of_glossopharyngeal_IX_nerves	9.70E-04	0.00167	-3.4099
62	right_lung_cranial_lobe	2.4E-12	3.5E-11	8.10042	40	stomach	1.00E-03	0.00167	3.401
61	left_lung	1.0E-11	1.3E-10	7.79203	198	pelvic_girdle	1.01E-03	0.00167	3.39898
127	triceps_brachii	2.7E-11	3.1E-10	7.58555	146	hindlimb_phalanges	1.07E-03	0.00174	-3.3812
120	thymus_primordium	4.1E-11	4.3E-10	7.49823	162	semitendinosus	1.34E-03	0.00216	3.30884
64	right_lung_caudal_lobe	7.0E-11	6.5E-10	7.3866	95	heart	4.46E-03	0.00707	-2.9163
75	trigeminal_V_ganglia	1.7E-10	1.5E-09	-7.1971	23	midbrain	7.42E-03	0.01148	-2.7389
145	latissimus_dorsi	2.3E-10	1.9E-09	7.12757	99	aortic_arch	7.47E-03	0.01148	-2.7365
161	semimembranosus	4.0E-10	2.9E-09	7.01389	36	Meckel_s_cartilage	1.02E-02	0.01544	2.6238
155	psaos_major	4.8E-10	3.3E-09	6.97466	97	ascending_aorta	1.11E-02	0.01649	-2.5943
130	flexor_carpi_ulnaris	5.3E-10	3.4E-09	6.95087	44	rectum	1.16E-02	0.01709	2.57575
118	adrenals	8.1E-10	4.9E-09	6.85987	207	pleural_and_pericardial_components	1.20E-02	0.01741	2.56361
188	ribs	1.3E-09	7.7E-09	6.75128	191	hyoid_bone	1.38E-02	0.01976	-2.5108
65	right_lung_accessory_lobe	1.6E-09	8.5E-09	6.71782	98	descending_aorta	1.59E-02	0.02242	-2.4574
58	metanephros	1.8E-09	9.0E-09	6.69273	66	bronchii	1.67E-02	0.02325	2.43816
41	stomach_lumen	2.1E-09	1.0E-08	6.65099	67	trachea	1.98E-02	0.02725	2.37115
144	teres_major	4.0E-09	1.9E-08	6.51352	110	ductus_arteriosus	2.82E-02	0.0382	-2.2303
153	quadratus_femoris	4.1E-09	1.9E-08	6.50373	195	epiglottis	3.48E-02	0.04656	-2.1427
24	pons	2.6E-08	1.1E-07	-6.0971	122	metacarpals	3.66E-02	0.0483	-2.1217
25	medulla_oblongata	5.2E-08	2.1E-07	-5.9399	200	cervical_drg	4.05E-02	0.05285	-2.0779
77	superior_ganglion_of_glossopharyngeal_IX_nerves	6.5E-08	2.6E-07	-5.8887	186	sacral_vertebrae	5.05E-02	0.06505	1.98174
48	masseter_muscle	1.2E-07	4.7E-07	5.73888	185	lumbar_vertebrae	5.73E-02	0.07283	1.92562
21	cerebellar_primordium	1.3E-07	4.9E-07	-5.7233	47	medial_pterygoid_muscle	7.13E-02	0.0895	1.8252
94	spinal_cord_marginal_layer	1.7E-07	6.1E-07	-5.6631	199	diaphragm	9.20E-02	0.11413	1.70306
53	left_lobe_of_liver	1.8E-07	6.3E-07	5.64832	51	vomeronasal_organ	1.06E-01	0.12944	1.63473
125	ulna	1.9E-07	6.4E-07	5.63678	70	nasal_cavity_epithelium	1.54E-01	0.18673	1.43718
152	iliacus	2.2E-07	7.0E-07	5.60969	96	aorta	1.66E-01	0.19923	1.39521
93	spinal_cord_mantle_layer	3.3E-07	1.0E-06	-5.5133	3	forebrain	1.81E-01	0.21417	-1.3483
63	right_lung_middle_lobe	4.6E-07	1.4E-06	5.43097	206	peritoneal_cavity	2.09E-01	0.24512	1.2641
113	jugular_lymph_sacs	7.8E-07	2.2E-06	-5.3078	79	superior_cervical_sympathetic_ganglia	2.42E-01	0.28046	-1.1768
126	humerus	7.8E-07	2.2E-06	5.30579	184	thoracic_vertebrae	2.47E-01	0.28249	1.16556
190	exoccipital_bone	8.1E-07	2.2E-06	-5.2992	175	flexor_digitorum_longus	2.81E-01	0.31819	-1.0843
38	lower_molar_primordia	8.2E-07	2.2E-06	5.2952	176	popliteus	2.98E-01	0.33394	1.04613
149	fibula	1.2E-06	3.1E-06	5.2085	49	tongue	3.70E-01	0.4093	-0.9018
54	right_lobe_of_liver	1.3E-06	3.3E-06	5.1882	148	tarsals	4.88E-01	0.53457	-0.6965
151	femur	1.8E-06	4.5E-06	5.10603	80	sympathetic_ganglia	5.42E-01	0.58789	0.61175
180	pectoralis	2.5E-06	6.1E-06	5.02603	183	cervical_vertebrae	5.92E-01	0.63318	0.53846
173	gastrocnemius_lateral_head	5.7E-06	1.4E-05	4.81899	42	foregut	5.96E-01	0.63318	-0.5316
197	scapula	8.7E-06	2.0E-05	4.71349	115	superior_vena_cava	6.83E-01	0.71815	0.40929
150	tibia	1.2E-05	2.7E-05	4.63301	187	tail_vertebrae	6.96E-01	0.72425	-0.3918
124	radius	1.4E-05	3.1E-05	4.59813	111	ductus_venosus	0.81439	0.83811	-0.2354
55	caudate_lobe_of_liver	1.8E-05	4.0E-05	4.52535	114	inferior_vena_cava	0.82183	0.83811	-0.2258
170	vastus_intermedius	2.3E-05	5.0E-05	4.46207	196	clavicle	0.93811	0.94731	-0.0779
27	lens_of_the_eye	3.2E-05	6.8E-05	4.37406	119	spleen_primordium	0.94772	0.94772	0.06575
56	bladder	6.8E-05	0.00014	4.17571					
43	mid_hindgut	8.4E-05	0.00017	4.1183					
205	ovary	0.00011	0.00022	4.04187					

Table 2. Wild type staging effect linear model outputs.

Results from fitting calculated organ volumes to linear model : organ volume ~ whole embryo volume (WEV). p: WEV effect p-value.
q: Benjamini-Hochburg FDR-corrected p-values, t: WEV effect t-statistics.

Allelic Composition	Genetic Background	Annotated Term	Identified with LAMA organ volume analysis	Identified with LAMA Jacobian analysis	Phenotype Summary Category	MGI publication reference
Acan-cmd-Bco/Acan-cmd-Bco	BAU/Gallic-Acan-cmd-Bco	defect palate	No atlas label	Yes	craniofacial	158755
Acan-cmd-Bco/Acan-cmd-Bco	BAU/Gallic-Acan-cmd-Bco	short snout	No atlas label	Yes	craniofacial	158755
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	short basicranium	Smaller organ volume (exoccipital)	Yes	craniofacial	130795
Acan-cmd-Bco/Acan-cmd-Bco	BAU/Gallic-Acan-cmd-Bco	distended abdomen	No atlas label	Yes	growth/size/body region	158755
Acan-cmd-Bco/Acan-cmd-Bco	BAU/Gallic-Acan-cmd-Bco	protruding tongue	No	No	growth/size/body region	158755
Acan-cmd-Bco/Acan-cmd-Bco	BAU/Gallic-Acan-cmd-Bco	small thoracic cavity	Yes (pleural & pericardial cavities)	Yes	growth/size/body region	165316
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	flattened snout	Not in atlas	Yes	growth/size/body region	15952
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal limb morphology	Yes	Yes	limbs/digits/tail	130795
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	absent caudal vertebrae	Yes (specimen level)	Yes	limbs/digits/tail	130795
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	brachydactyly	No (digits not analyzed)	Yes	limbs/digits/tail	130795
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	enlarged liver	Yes	Some significant voxels in forelimb phalanges	liver/biliary system	15952
Acan-cmd-Bco/Acan-cmd-Bco	BAU/Gallic-Acan-cmd-Bco	pulmonary hypoplasia	No	No	respiratory system	165316
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal lung morphology	No	Yes	respiratory system	123353
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal tracheal cartilage morphology	No	Yes	respiratory system	15952
Acan-cmd-Bco/Acan-cmd-Bco	C57BL/6N-Acan-cmd-Bco	abnormal bone structure	Yes	Yes	skeleton	1211773
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal cartilage development	Yes (abnormal bones)	Yes	skeleton	130795
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal craniofacial bone morphology	No	Yes	skeleton	130795
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal rib morphology	Yes	Yes	skeleton	130795
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal trabecular bone morphology	Yes	Yes	skeleton	15952
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	abnormal vertebral morphology	Yes	Yes	skeleton	15952
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	decreased length of long bones	Yes	Yes	skeleton	15952
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	short femur	Yes	Yes	skeleton	15952
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	short humerus	Yes	Yes	skeleton	15952
Acan-cmd-Bco/Acan-cmd-Bco	STOCK T hloww tpr3-4tf	short vertebral column	Yes	Yes	skeleton	15952

Table 3. Phenotypes recorded for Acan^{-/-} mutant mice in the MGI database. Green cells indicates concordance with recorded phenoty

Funding

The National Institutes of Health (grant number 1 U54 HG006370-01). Lydia Teboul was supported by a Medical Research Council Strategic Award.

Acknowledgements

We would like to thank James Cleak and and Zsombor Szoke-Kovacs for generating image data and helpful discussions. Also thanks to the husbandry team from the Mary Lyon Centre for generation and maintenance of all the mouse specimens used in the study.

References

1. Adams D, Baldock R, Bhattacharya S, et al. Bloomsbury report on mouse embryo phenotyping: recommendations from the IMPC workshop on embryonic lethal screening. *Dis Model Mech*. 2013;6(3):571-579. doi:10.1242/dmm.011833
2. Ashish N, Brusniak M-Y. Automated Mouse Organ Segmentation: A Deep Learning Based Solution. April 2018. <http://arxiv.org/abs/1804.09205>. Accessed January 29, 2020.
3. Benjamini Y, Hochberg Y. Benjamini-1995.pdf. *J R Stat Soc B*. 1995;57(1):289-300. doi:10.2307/2346101

4. Bingle L, Singleton V, Bingle CD. The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. *Oncogene*. 2002;21(17):2768-2773. doi:10.1038/sj.onc.1205363
5. Birling M-C, Yoshiki A, Adams DJ, et al. A resource of targeted mutant mouse lines for 5,061 genes. *bioRxiv*. 2019. doi:10.1101/844092
6. Boehm B, Rautschka M, Quintana L, Raspopovic J, Jan Ž, Sharpe J. A landmark-free morphometric staging system for the mouse limb bud. *Development*. 2011;138(6):1227-1234. doi:10.1242/dev.057547
7. Bowl MR, Simon MM, Ingham NJ, et al. A large scale hearing loss screen reveals an extensive unexplored genetic landscape for auditory dysfunction. *Nat Commun*. 2017;8(1):886. doi:10.1038/s41467-017-00595-4
8. Brown JM, Horner NR, Lawson TN, et al. A bioimage informatics platform for high-throughput embryo phenotyping. *Brief Bioinform*. 2018;19(1):41-51. doi:10.1093/bib/bbw101
9. Brown SDM, Moore MW. The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. *Mamm Genome*. 2012;23(9-10):632-640. doi:10.1007/s00335-012-9427-x
10. Cacheiro P, Muñoz-Fuentes V, Murray SA, et al. Human and mouse essentiality screens as a resource for disease gene discovery. *Nat Commun*. 2020;11(1):655. doi:10.1038/s41467-020-14284-2
11. Cleary JO, Modat M, Norris FC, et al. Magnetic resonance virtual histology for embryos: 3D atlases for automated high-throughput phenotyping. *Neuroimage*. 2011;54(2):769-778. doi:10.1016/j.NEUROIMAGE.2010.07.039
12. Dagg CP. The interaction of environmental stimuli and inherited susceptibility to congenital deformity. *Integr Comp Biol*. 1963;3(2):223-233. doi:10.1093/icb/3.2.223
13. Dickinson ME, Flenniken AM, Ji X, et al. High-throughput discovery of novel developmental phenotypes. *Nature*. 2016;537(7621):508-514. doi:10.1038/nature19356
14. Dorr AE, Lerch JP, Spring S, Kabani N, Henkelman RM. High resolution three-dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice. *Neuroimage*. 2008;42(1):60-69. doi:10.1016/j.neuroimage.2008.03.037
15. Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *J Opt Soc Am A*. 1984;1(6):612. doi:10.1364/josaa.1.000612
16. Hayamizu TF, Wicks MN, Davidson DR, Burger A, Ringwald M, Baldock RA. EMAP/EMAPA ontology of mouse developmental anatomy: 2013 update. *J Biomed Semantics*. 2013;4(1):15. doi:10.1186/2041-1480-4-15
17. Holmes CJ, Hoge R, Collins L, Woods R, Toga AW, Evans AC. Enhancement of MR images using registration for signal averaging. *J Comput Assist Tomogr*. 1998;22(2):324-333. doi:10.1097/00004728-199803000-00032

18. Houghton MJ, Carey JC, Seegmiller RE. Pulmonary Hypoplasia in Mice Homozygous for the Cartilage Matrix Deficiency (Cmd) Gene: A Model for a Human Congenital Disorder. *Pediatr Pathol*. 1989;9(5):501-512. doi:10.3109/15513818909026909
19. Hrabě de Angelis M, Nicholson G, Selloum M, et al. Analysis of mammalian gene function through broad-based phenotypic screens across a consortium of mouse clinics. *Nat Genet*. 2015;47(9):969-978. doi:10.1038/ng.3360
20. Kaufman MH. *The Atlas of Mouse Development*. Academic Press; 1992.
21. Kovačević N, Henderson JT, Chan E, et al. A three-dimensional MRI atlas of the mouse brain with estimates of the average and variability. *Cereb Cortex*. 2005;15(5):639-645. doi:10.1093/cercor/bhh165
22. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062
23. Li J, Chen H, Mariani A, et al. HE4 (WFDC2) Promotes Tumor Growth in Endometrial Cancer Cell Lines. *Int J Mol Sci*. 2013;14(3):6026-6043. doi:10.3390/ijms14036026
24. Lizio M, Harshbarger J, Shimoji H, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol*. 2015;16(1):22. doi:10.1186/s13059-014-0560-6
25. Lloyd KCK, Adams DJ, Baynam G, et al. The Deep Genome Project. *Genome Biol*. 2020;21(1):18. doi:10.1186/s13059-020-1931-9
26. Meehan TF, Conte N, West DB, et al. Disease model discovery from 3,328 gene knockouts by the International Mouse Phenotyping Consortium. *Nat Genet*. 2017;49(8):1231-1238. doi:10.1038/ng.3901
27. Mianné J, Codner GF, Caulder A, et al. Analysing the outcome of CRISPR-aided genome editing in embryos: Screening, genotyping and quality control. *Methods*. 2017;121-122:68-76. doi:10.1016/j.ymeth.2017.03.016
28. Moore BA, Leonard BC, Sebbag L, et al. Identification of genes required for eye development by high-throughput screening of mouse knockouts. *Commun Biol*. 2018;1(1). doi:10.1038/s42003-018-0226-0
29. Nakajima K, Ono M, Radović U, et al. Lack of whey acidic protein (WAP) four-disulfide core domain protease inhibitor 2 (WFDC2) causes neonatal death from respiratory failure in mice. *Dis Model Mech*. 2019;12(11):dmm040139. doi:10.1242/dmm.040139
30. Oprea TI. Exploring the dark genome: implications for precision medicine. *Mamm Genome*. 2019;30(7-8):192-200. doi:10.1007/s00335-019-09809-0
31. Peterka M, Lesot H, Peterková R. Body Weight in Mouse Embryos Specifies Staging of Tooth Development. *Connect Tissue Res*. 2002;43(2-3):186-190. doi:10.1080/03008200290000673
32. R Core Team. R: A Language and Environment for Statistical Computing. 2018. <https://www.r-project.org/>.
33. Richardson L, Venkataraman S, Stevenson P, et al. EMAGE mouse embryo spatial gene expression database: 2014 update. *Nucleic Acids Res*. 2014;42(D1). doi:10.1093/nar/gkt1155

34. Stattin E-L, Wiklund F, Lindblom K, et al. A Missense Mutation in the Aggrecan C-type Lectin Domain Disrupts Extracellular Matrix Interactions and Causes Dominant Familial Osteochondritis Dissecans. *Am J Hum Genet.* 2010;86(2):126-137. doi:10.1016/j.ajhg.2009.12.018
35. Theiler K. Procedure. In: *The House Mouse*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1989:1-2. doi:10.1007/978-3-642-88418-4_3
36. Thompson SW, Merriman B, Funari VA, et al. A recessive skeletal dysplasia, SEMD aggrecan type, results from a missense mutation affecting the C-type lectin domain of aggrecan. *Am J Hum Genet.* 2009;84(1):72-79. doi:10.1016/j.ajhg.2008.12.001
37. Venables WN, Ripley BD. *Modern Applied Statistics with S*. Fourth. New York: Springer; 2002. <http://www.stats.ox.ac.uk/pub/MASS4>.
38. Waterston RH, Lindblad-Toh K, Birney E, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature.* 2002;420(6915):520-562. doi:10.1038/nature01262
39. Wilson R, Geyer SH, Reissig L, et al. Highly variable penetrance of abnormal phenotypes in embryonic lethal knockout mice. *Wellcome Open Res.* 2017;1:1. doi:10.12688/wellcomeopenres.9899.2
40. Wilson R, McGuire C, Mohun T, et al. Deciphering the mechanisms of developmental disorders: Phenotype analysis of embryos from mutant mouse lines. *Nucleic Acids Res.* 2016;44(D1):D855-D861. doi:10.1093/nar/gkv1138
41. Wong MD, Dorr AE, Walls JR, Lerch JP, Henkelman RM. A novel 3D mouse embryo atlas based on micro-CT. *Development.* 2012;139(17):3248-3256. doi:10.1242/dev.082016
42. Wong MD, Maezawa Y, Lerch JP, Henkelman RM. Automated pipeline for anatomical phenotyping of mouse embryos using micro-CT. *Development.* 2014;141(12):2533-2541. doi:10.1242/dev.107722
43. Wong MD, Van Eede MC, Spring S, et al. 4d atlas of the mouse embryo for precise morphological staging. *Dev.* 2015;142(20):3583-3591. doi:10.1242/dev.125872
44. Yan D, Zhang Z, Luo Q, Yang X. A novel mouse segmentation method based on dynamic contrast enhanced micro-ct images. *PLoS One.* 2017;12(1). doi:10.1371/journal.pone.0169424
45. Zamyadi M, Baghdadi L, Lerch JP, et al. Mouse embryonic phenotyping by morphometric analysis of MR images. *Physiol Genomics.* 2010;42 A(2):89-95. doi:10.1152/physiolgenomics.00091.2010
46. Zhang T, Long H, Li J, Chen Z, Wang F, Jiang SW. WFDC2 gene deletion in mouse led to severe dyspnea and type-I alveolar cell apoptosis. *Biochem Biophys Res Commun.* 2020. doi:10.1016/j.bbrc.2019.11.011