A landmark-free morphometrics pipeline: Application to characterise a Down syndrome model skull phenotype

Nicolas Toussaint[1]*, Yushi Redhead[2,3]*, Wei Liu[4], Elizabeth M. C. Fisher[5], Benedikt Hallgrimsson[4], Victor L.J. Tybulewicz[3,6§], Julia A. Schnabel[1§], Jeremy B.A. Green[2§]

[1] School of Biomedical Engineering and Imaging Sciences, King's College London, United Kingdom
[2] Centre for Craniofacial Biology & Regeneration, King's College London, United Kingdom.
[3] The Francis Crick Institute, London NW1 1AT, United Kingdom.
[4] Department of Cell Biology & Anatomy, University of Calgary, Calgary AB T2N 4N1, Canada
[5] Institute of Neurology, University College London, London, WC1N 3BG, United Kingdom.
[6] Department of Immunology & Inflammation, Imperial College London, London W12 0NN, United Kingdom.

*Equal contribution

[§]Correspondence to:
JAS (email: julia.schnabel@kcl.ac.uk), VLJT (email: Victor.T@crick.ac.uk), or to
JBAG (email: jeremy.green@kcl.ac.uk)

Abstract

Characterising phenotypes often requires quantification of anatomical shapes. Quantitative shape comparison (morphometrics) traditionally uses anatomical landmarks and is therefore limited by the number of landmarks and operator error when landmarks are located manually. Here we create a landmark-free method and use it to characterise the craniofacial skeletal phenotype of the Dp1Tyb mouse model of Down syndrome (DS), comparing it with a landmark-based approach. Both methods identify cranial dysmorphologies in Dp1Tyb mice, especially smaller size and brachycephaly (front-back shortening) homologous to the human phenotype. However, the landmark-free method provides better separation by principal component analysis, identifies differences where landmarks are absent, and better highlights reductions in mid-face structures that are reduced in humans with DS. Moreover, the landmark-free method is less labour-intensive and requires less user training. We conclude that this is a useful method making morphometrics widely-accessible beyond traditional niches in zoology and palaeontology, especially in characterising mutant phenotypes.

Introduction

Morphometrics, the quantitative comparison of biological shapes, is well established in the fields of palaeontology and evolutionary biology[1] and is emerging in genetic studies to quantify and understand morphological phenotypes. Current morphometric methodologies, while powerful, have limitations that reflect a compromise between precision and ease-of-use. Modern morphometrics, otherwise known as geometric morphometrics, uses anatomical landmarks to describe morphological structures and uses computational statistical methods to compare the landmark locations. Landmarks can be histological boundaries (e.g. the intersection between two skull sutures), geometric (e.g. the tip of a bone or the bottom of a sulcus), or relative (e.g. the point on a curved histological boundary furthest away from another landmark). Typically, some tens of landmarks are placed, usually manually, on digital two- or three-dimensional tomographic images (obtained by photography, X-ray or MRI methods) and are then analysed through Euclidean distance matrix analysis (EDMA) or Procrustes Superimposition (PS)[2].

Landmark-based methodologies have limitations. Firstly, they are not easy to apply. There are occasions where an anatomical landmark may be absent from an individual due to natural variation, engineered mutation or pathology. At best this is a nuisance and at worst a landmark of biological relevance is lost, potentially affecting the analysis and interpretation of the data. A second limitation is that smooth surfaces do not have easily defined landmarks. Semi-landmarks, which are points evenly spaced along an outline or curve between two landmarks[3] can reduce this problem to a degree, and a number of computational advances have enabled them to be determined automatically[4,5]. Unfortunately, semi-landmark contours may still leave gaps, potentially missing variation in doming or depression in the triangular surfaces between nearest-neighbour landmark trios[6]. This is a particular problem in soft tissues and embryos, with numerous essentially featureless, smoothly curved surfaces. A third limitation of current landmark-based methods is operator error. Landmark sets for early embryonic mouse head

anatomy are available but were shown to be highly susceptible to both inter- and intra-operator error[7], not just in embryos but also more widely[8]. A study by Shearer et al[9] found that the variance in landmarking Tibetan macaque skulls was equal to or larger than the natural variability found between subjects. In summary, the combined problems of landmark variability, landmark sparseness and operator labour and reliability must be overcome if morphometrics is to be applied widely, especially as a tool to understand the genetics of dysmorphologies.

One of the most common human dysmorphologies is the craniofacial phenotype associated with Down syndrome (DS). Individuals with DS, currently ~1 in 800 births[11], have characteristic features – flattened midface with low nose bridge, front-to-back shortened skull (brachycephaly) and slightly hooded eyelids[10]. Although the craniofacial features affect everyone with DS, this phenotype is not well understood either genetically or developmentally. DS is caused by trisomy of human chromosome 21 (Hsa21) which carries 232 protein-coding genes (Ensembl genome assembly GRCh38)[11,12]. It is thought that the presence of a third copy of one or more of these genes gives rise to the individual defects observed in DS, but the critical dosage-sensitive genes are not known[13–15].

Hsa21 is orthologous to three regions of the mouse genome, located on mouse chromosome 16 (Mmu16), Mmu10 and Mmu17; the largest region on Mmu16 contains 148 protein-coding genes. To model DS, mouse strains have been engineered that carry an extra copy of each of these regions, and between them they recapitulate many aspects of DS [14,16,17]. Morphometrics applied to some of these models, including Ts65Dn [18–20] and Dp(16)1Yey mice [21] showed that trisomy of part or all of the Hsa21-orthologous region of Mmu16 resulted in craniofacial dysmorphology which resembled the DS phenotype. The cranial dysmorphology in Dp1Yey mice was highly consistent (with multiple linear distances between landmarks differing statistically significantly from wild type in all regions measured[21]). Yet the dysmorphic

phenotype was remarkably subtle, with an average landmark-to-landmark distance difference of only 7% between mutant and wild-type (WT) control mice[21].

Landmark-free methods have been developed by the neuroimaging community to quantify the size and shape of the brain precisely because its relatively smooth shape hampers the definition of reliable landmarks[22,23]. However, these methods have not been directly compared to the landmark-based approach. In this paper, we describe a convenient pipeline for landmark-free morphometric analysis based on an approach used for brain imaging[24]. We compare our method to the traditional landmark-based morphometric approach, focusing on the characterisation of the craniofacial phenotype of the Dp1Tyb mouse model of DS which has an additional copy of the entire Hsa21-orthologous region of Mmu16 (ref [14]). We find that the landmark-free analysis gives separation by shape between Dp1Tyb and WT mice that is at least as clear as that achieved by landmark-based analysis, while delivering a number of operational advantages and identifying phenotypic changes not seen in landmark-based analysis. Additionally, we show that the landmark-free method provides a rich output that can be used for higher-resolution statistical and spatial analysis, providing an approach to interpret shape modifications between groups as local tissue effects.

Results

Landmark-based and landmark-free analysis of Dp1Tyb skulls

We used micro-computed tomography (μCT) to acquire images of the skulls of 16-week old WT and Dp1Tyb mice with which to compare the standard landmark-based analysis with our novel landmark-free analysis method. We carried out landmark-based analysis in the conventional way[25], marking the location of 68 landmarks on the cranium and 17 on the mandible (Supplementary Fig. 1). Crania and mandibles were analysed separately since their relative position varies from subject to subject. Landmarks for all crania and mandibles were

aligned using Procrustes Superimposition, and these data were used for further statistical analysis of size and shape.

For the landmark-free approach we developed a pipeline based on previous approaches in morphometrics and neuroimaging (Fig. 1, Table 1, Supplementary Appendix 1). Following thresholding to extract the skull structures from the μCT images, cartilaginous structures were removed (Supplementary Fig. 2) and the images segmented using bone density to separate the mandibles from the crania (step 1). Meshes were generated from the surfaces of the cranium and mandible for all subjects (step 2), aligned using Procrustes Superimposition for scaling (step 3) and these were used for the construction of an atlas (mean shape) for the crania and mandibles of the WT and Dp1Tyb skulls (step 4). Atlas construction was based on the Deformetrica algorithm[24] which, in brief, consists of construction an average mesh of the population, using the mesh to define a flow field that conforms to its shape and quantifies deformations from it to each subject recorded as momentum vectors (momenta - see Methods and Supplementary Appendix 1). The initial output from this atlas consisted of the average mesh for the whole population, a set of control points corresponding to areas with the greatest variability between subjects, and momenta for each control point describing the directional variation of the shape from the average. The average mesh, the control points and the momenta were used for further statistical analysis, with the momenta applied to deform the population average mesh to generate average meshes for each of WT and Dp1Tyb groups preserving one-to-one correspondence of mesh vertices. We performed principal component analyses and used a multiple permutations test on a stratified k fold cross validation classifier to test for significance. To control for overfitting (a risk when the number of measurements substantially exceeds the number of subjects), we compared the PCA difference vector magnitude between the two genotype groups with that of 1000 randomly scrambled groups. We found that the distribution was Normal and that the genotype difference vector was more than 3.5 standard deviations away from the mean vector of the 1000 scrambled groups for both cranium and mandible.

Dp1Tyb mice have smaller crania and mandibles

Centroid size (the mean absolute landmark distance from the landmark-defined centroid) is the most commonly used measure of size in landmark-based geometric morphometrics [26]. We used it here to determine if there were any changes in skull and mandible size between the DS model and controls. Landmark-based comparison of centroid sizes by genotype found that the crania and mandibles of Dp1Tyb mice were both significantly smaller than those of WT mice (Fig. 2A, C), recapitulating the overall reduction in skull size found in humans with DS and as well as in other models of DS[18–21,27]. The landmark-free analysis also showed that Dp1Tyb crania and mandibles were significantly smaller (Fig. 2B, D), with both methods showing a similar ~7% reduction in centroid size of both craniofacial structures in Dp1Tyb mice (Supplementary Table). The landmark-free and landmark-based centroid sizes were different, unsurprisingly given the many extra measurements use in the landmark-free method (~19,000 mesh vertices versus 68 landmarks for cranium and ~16,000 vertices versus 17 landmarks for mandible).

Dp1Tyb mice have altered shapes of crania and mandibles

Both the size difference and gross shape differences were clearly visualised by animated morphing between the mean shapes of WT and Dp1Tyb specimens (generated in the landmark-free pipeline) for the cranium and the mandible (Supplementary Videos 1, 2). The overall decrease in size going from WT to Dp1Tyb crania or mandibles was readily apparent and some shape changes could also be seen, although the latter were more subtle.

To quantify shape differences, these were separated from the size differences by using Procrustes alignment, i.e. scaling the data to equalise centroid sizes. To discover residual shape differences between genotypes, we used principal component analysis (PCA) and applied well established significance tests as described in Methods. Both landmark-based and landmark-free methods showed a statistically significant differences in shape between

Dp1Tyb and WT mice in both crania and mandibles (Fig. 2E-H). Plots of the first two Principal Components identified by the two different methods looked similar, with tighter clustering of specimens for cranium than for mandible.

Dp1Tyb mice recapitulate aspects of human DS craniofacial dysmorphology

To characterise the shape differences in more anatomical detail and to be able to compare the differences in shapes found through the landmark-based method, we used two approaches. First, we simply overlaid the mean landmark configurations from Dp1Tyb and WT crania and mandibles (Fig. 3A-F). Second, we applied an established thin-plate spline shape-interpolation and comparison package (Morpho R - see Methods) to the landmark data to generate displacement heatmaps (Fig. 3G-L). Direct inspection revealed that the morphological differences between Dp1Tyb and WT skulls were broadly distributed and relatively subtle, consistent both with previously reported mouse DS models and with the human phenotype[27–29]. More detailed examination revealed that Dp1Tyb mice have a more domed neurocranium (cyan points at the top-right of Fig. 3A, dark red regions in Fig. 3G, H, J). The cranial doming in combination with the overall smaller size compared to WT mice constitute a net anteroposterior shortening in Dp1Tyb mice, i.e. brachycephaly, a predominant feature of the human DS phenotype.

We also observed an almost unchanged cranial base (Fig. 3I), but some contraction (anterior movement) concentrated around the magnum foramen in the occipital bone of Dp1Tyb crania (points at right of Fig. 3C, blue colour on the right of Fig. 3I and in Fig. 3J). The heatmaps also showed a smaller snout in Dp1Tyb mice as a result of a reduction in size of the nasal bones (Fig. 3G, H). Although not evident in the heatmaps, the Dp1Tyb cranium was wider as can be seen by the displacement of the zygomatic processes' landmarks laterally (Fig. 3C and D). The reduced snout and facial widening in combination with the overall smaller size of the Dp1Tyb crania mimics the "mid-face hypoplasia" of human DS. The Dp1Tyb mandibles had a

small shape change in the alveolar ramus region and the condylar process but these changes were all subtle (less than 5 μm in magnitude) (Fig. 3K, L).

Next we made heatmaps and morphing movies based on the higher-resolution landmark-free method. Displacement maps together with the morphing movies visualised the distances between the two mean meshes. We plotted net displacement rather than movement towards or away from the shape centroid so that only one colour was needed in the maps, using videos to show the direction of differences. The landmark-free analysis showed changes mostly similar to those found using the landmark-based method including the same doming of the neurocranium, slight contraction of the posterior basicranium and shortening of the nasal and maxillary processes (Fig. 4A-D, Supplementary Video 3). It also showed an additional difference: basicranium in Dp1Tyb mice was displaced inferiorly, apparently contributing to the rounding of the cranium.

Differences between Dp1Tyb and WT mandibles were overall much more subtle (Fig. 4E,F, Supplementary Video 4) consisting of a few tens of microns only. Re-scaling the heatmap colour range better visualised the localisation of the differences (Fig. 4G,H) as did generating a video in which the shape change of the mandible was exaggerated by a factor of three (Supplementary Video 5). Thus, we see the expansion buccally of the inferior portion of the ramus, the contraction of the angular process as well as the lingual movement of the molar ridge and widening of the incisor alveolus. The expansion at the inferior portion of the ramus (the portion of the mandible just posterior to the molars) including the anterior-ventral part of the angular process (the most posterior and inferior process of the mandible) was not apparent in the landmark-based method in part because there are no landmarks in this region.

The landmark-free method provides new shape change information: in-plane deformation

Mesh displacement heatmaps primarily represent shape changes that are normal (perpendicular) to the surface of the specimen. However, shape differences arise in development due to different localised growth which is often in the plane, e.g. in planar skull bones (calvaria), and potentially remote from the location of displacement, e.g. proliferation in the growth plates of long bones displacing their tips. Mechanistically it would therefore be useful to capture surface "stretch" as a measure of local growth differences between specimens. This becomes feasible in a landmark-free method where a high density of control points is used to guide an even higher density of mesh vertices. Thus, we calculated and mapped local differences in mesh vertex spacing (Fig. 5A-F, Supplementary Videos 6, 7). The Dp1Tyb skull showed an expansion of the calvaria, most strongly in the lateral-central region of the parietal bones (Fig. 5B, D), which appear to be contributing to the overall doming of the cranial vault observed in both methods. A ~25% contraction (Fig. 5A-C) was also observed in a large region of the snout, not only on the upper surface (Fig. 5A, B) (the nasal bones, maxillary process and the anterior portion of the maxillary zygomatic process) but also on the underside, in the hard palate region (Fig. 5C) (palatine bone and processes). This can be compared to the smaller regions nearer the tip of the snout picked up in the displacement maps (Figs. 3H, 4B) and much better reflects what can be seen in the morphing videos (Supplementary Videos 3 and 6) and is similar to the general midfacial hypoplasia phenotype in humans. There is also contraction in the underside of the auditory bulla (Fig. 5C, arrowheads). In the mandible the main change, a ~10% contraction in the ramus region just anterior of the angular process, corresponds closely to the displacement map, but in addition, the stretch mapping showed a very slight expansion in the diastema (gap between molars and incisor) (Fig. 5E, F).

Discussion

In this paper, we have presented an adaptation and incorporation of the "currents"-based landmark-free shape comparison methodology[24] into a pipeline that can process µCT images into a form (i.e meshes) suitable for this method, and easily compare shapes and provide statistical and other data analyses comparable to more traditional geometric morphometrics. We have used this pipeline to analyse the previously unexamined craniofacial phenotype of Dp1Tyb mice, a relatively new model of DS[14]. We compared the results with a well-established landmark-based method that is commonly used for skeletal shape analysis, particularly in the comparative zoology and paleontology fields[1]. The newer method yielded new biological insights and offers a great deal of promise for use in phenotyping, including for genetic models.

We validated the landmark-free approach through comparison to the classical landmark-based geometric morphometric approach and found similar results; both revealed the novel finding that Dp1Tyb mice present highly consistent size and shape differences when compared to WT mice and these differences parallel those of the craniofacial phenotype observed in human individuals with DS[27–29] and in previously described DS mouse models[18–21]. Both methods also allowed the Dp1Tyb phenotype to be well separated in feature space from WT mice using one or two principal components and both revealed the significantly reduced size of the cranium and mandible of Dp1Tyb mice. We found that the DS model mice show brachycephaly, resolved in the analysis to an overall size reduction plus cranial doming. We also found that the landmark-free approach can provide new dense information about the shape changes in Dp1Tyb mice, allowing us to see changes not visible using the landmark-based approach. Most strikingly, we were able to observe a shape difference in the lower-posterior mandible, where landmarks are absent, and in the snout, cranial base and palate, where landmarks are more abundant but possibly not dense enough to capture the localised in-plane differences. These latter changes in particular indicate homology with the mid-face hypoplasia found in humans with DS.

Methodologically, the landmark-free method has three clear advantages. Firstly, the landmark-free method overcomes the limitations of manual placement of landmarks which is susceptible to inter- and intra-operator error[30]. Secondly, although the landmark-free method requires some manual input in the early stages, particularly in determining image thresholding and in cleaning up imperfect anatomical segmentation, it is substantially less labour-intensive and requires less user training than the landmark-based approach. Thirdly, and perhaps most scientifically significant, the landmark-free method provides much higher resolution and information density than the landmark-based method. In principle, the landmark-free method offers arbitrarily high resolution. In practice we found that decimating the initial mesh from ~2,800,000 and ~200,000 to ~19,000 and ~16,000 vertices and using a kernel size of 2mm to yield 940 and 350 control points for the cranium and mandible respectively [24], captured the interesting anatomical features at high density while avoiding noise, e.g. trivial surface texture differences. Different sizes of specimen will have different optimal spatial parameters. The high density of control points was further refined by having them clustered algorithmically at regions of high variability between samples. This might be contrasted with the inherent bias in landmarking that tends to place shape differences close to landmarks (observable in, for example, Fig. 3A, B). It can also be contrasted with the use of semi-landmarks, which adds landmarks that are evenly distributed across contours or surfaces that are bounded by identifiable landmarks [31,32] but still potentially leaves gaps where landmarks are sparse. A trade-off for all methods that increase the number of discrete observations, however, is the increasing as risk of overfitting as the number of variables (e.g. landmarks or voxels) increases relative to the statistical degrees of freedom, or the “curse of dimensionality” [33]. Controls for overfitting, as we have applied, are therefore, where possible, especially necessary.

Displacement heat maps reveal any kind of displacement, whether in-plane or perpendicular to it. In capturing the shape difference between a subject and control, the landmark-free method retains all the vertices of the mesh. This means that local "stretch" or "compression"

can be mapped as vertex density and are captured in our stretch maps. In effect, each vertex is a landmark in a highly granular landscape. The very short spatial scale of this mapping is likely to be a much better way to capture and localise changes in biologically causal processes, such as cell proliferation or extracellular matrix expansion, compared to net displacement, where the cause could be hundreds of cell diameters away.

The stretch data has both refined the changes found, e.g. in the case of the cranial base showing how the auditory bulla are contracting, and provided new details in the case of the contraction of the inter-sphenoidal synchondrosis and the zygomaticotemporal suture. This type of new information is vital in accurately dissecting the causal developmental changes leading to altered anatomy. We have shown for the first time that not only are the zygomatictemporal sutures displaced laterally (Fig. 3B, C), relative to the rest of the skull, but have a localised contraction suggesting a failure to grow in early development either pre- or postnatally (Fig. 5B, C). It has been suggested that this type of insufficiency at this site could lead to facial dysmorphologies[34,35].

The landmark-free method has previously demonstrated its strength in the analysis of objects with a paucity of landmarks, such as MRI scans of brains where it was first described[24]. We have applied it to a much more landmark-rich type of specimen with two important conclusions. Our first conclusion is that using a landmark-free approach is still advantageous even when traditional landmarking is possible. The second conclusion is that for those working with mouse models in particular, this is a readily useable tool. We believe this pipeline provides a portable tool that can be set up with modest computational expertise that will enable researchers to tackle questions of mutant phenotypes that traditional methods have struggled with, such as in early developmental stages or other biological forms that lack well defined landmarks.

Acknowledgements

We thank Heather Cater and Sara Wells from MRC Harwell for breeding the mice. VLJT and EMCF are supported by the Wellcome Trust (grants 080174, 098327 and 098328) and VLJT by the Francis Crick Institute which receives its core funding from Cancer Research UK (FC001194), the UK Medical Research Council (FC001194), and the Wellcome Trust (FC001194). YR was supported by the Francis Crick Institute and King's College London. JBAG was supported by King's College London. NT and JAS acknowledge support from the Wellcome/EPSRC Centre for Medical Engineering [WT 203148/Z/16/Z] and the Wellcome Trust IEH Award [102431]. BH is supported by a CIHR Foundation grant, NIH R01 R01DE019638 and the Canadian Foundation for Innovation.

Author contributions

NT designed and wrote the code for the analysis. YR and WL carried out experiments. NT, YR and WL analysed the data. YR, NT, VLJT, JAS and JBAG wrote paper. BH, EMCF, VLJT, JBAG and JAS supervised the work.

Competing interests

The authors declare no competing interests.

References

1. Cooke, S. B. & Terhune, C. E. Form, Function, and Geometric Morphometrics. *Anat. Rec.* **298**, 5–28 (2015).
2. Webster, M. & Sheets, H. D. A Practical Introduction to Landmark-based Geometric Morphometrics. *Paleontol. Soc. Pap.* **16**, 163–188 (2010).
3. Bookstein, F. L. Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Med. Image Anal.* **1**, 225–243 (1997).
4. Andresen, P. R. *et al.* Surface-bounded growth modeling applied to human mandibles. *IEEE Trans. Med. Imaging* **19**, 1053–1063 (2000).
5. Frangi, A. F., Rueckert, D., Schnabel, J. A. & Niessen, W. J. Automatic Construction of Biventricular Statistical Shape Models. in *Proceedings of the 2nd international conference on Functional imaging and modeling of the hear* 18–29 (Springer, Berlin, Heidelberg, 2003).
6. Palci, A. & Lee, M. S. Y. Geometric morphometrics, homology and cladistics: review and recommendations. *Cladistics* **35**, 230–242 (2019).
7. Percival, C. J., Green, R., Marcucio, R. & Hallgrímsson, B. Surface landmark quantification of embryonic mouse craniofacial morphogenesis. *BMC Dev. Biol.* **14**, 31 (2014).
8. Von Cramon-Taubadel, N., Frazier, B. C. & Lahr, M. M. The problem of assessing landmark error in geometric morphometrics: Theory, methods, and modifications. *Am. J. Phys. Anthropol.* **134**, 24–35 (2007).
9. Shearer, B. M. *et al.* Evaluating causes of error in landmark-based data collection using scanners. *PLoS One* **12**, e0187452 (2017).
10. Korenberg, J. R. Corrections: Down Syndrome Phenotypes: The Consequences of Chromosomal Imbalance. *Proc. Natl. Acad. Sci.* **91**, 11281a – 11281 (2006).
11. Antonarakis, S. E. Down syndrome and the complexity of genome dosage imbalance. *Nat. Rev. Genet.* **18**, 147–163 (2017).

12. Lejeune, J., Gautier, M. and Turpin, R. Study of somatic chromosomes from 9 mongoloid children. *C. R. Hebd. Seances Acad. Sci.* **248**, 1721–2 (1959).

13. Lana-Elola, E., Watson-Scales, S. D., Fisher, E. M. C. & Tybulewicz, V. L. J. Down syndrome: searching for the genetic culprits. *Dis. Model. Mech.* **4**, 586–95 (2011).

14. Lana-Elola, E. *et al.* Genetic dissection of Down syndrome-associated congenital heart defects using a new mouse mapping panel. *Elife* **5**, 1689–1699 (2016).

15. Watson-Scales, S. *et al.* Analysis of motor dysfunction in Down Syndrome reveals motor neuron degeneration. *PLoS Genet.* **14**, e1007383 (2018).

16. Yu, T. *et al.* A mouse model of Down syndrome trisomic for all human chromosome 21 syntenic regions. *Hum. Mol. Genet.* **19**, 2780–2791 (2010).

17. Herault, Y. *et al.* Rodent models in Down syndrome research: impact and future opportunities. *Dis. Model. Mech.* **10**, 1165–1186 (2017).

18. Richtsmeier, J. T., Zumwalt, A., Carlson, E. J., Epstein, C. J. & Reeves, R. H. Craniofacial phenotypes in segmentally trisomic mouse models for Down syndrome. *Am. J. Med. Genet.* **107**, 317–324 (2002).

19. Hill, C. A., Reeves, R. H. & Richtsmeier, J. T. Effects of aneuploidy on skull growth in a mouse model of Down syndrome. *J. Anat.* **210**, 394–405 (2007).

20. Richtsmeier, J. T., Baxter, L. L. & Reeves, R. H. Parallels of craniofacial maldevelopment in Down syndrome and Ts65Dn mice. *Dev. Dyn.* **217**, 137–145 (2000).

21. Starbuck, J. M., Dutka, T., Ratliff, T. S., Reeves, R. H. & Richtsmeier, J. T. Overlapping trisomies for human chromosome 21 orthologs produce similar effects on skull and brain morphology of Dp(16)1Yey and Ts65Dn mice. *Am. J. Med. Genet. Part A* **164**, 1981–1990 (2014).

22. Bron, E. E. *et al.* Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *Neuroimage* **111**, 562–579 (2015).

23. Routier, A. *et al.* Evaluation of morphometric descriptors of deep brain structures for

the automatic classification of patients with Alzheimer's disease, mild cognitive impairment and elderly controls. *Proc MICCAI Work. Chall. Comput. Diagnosis Dement. Based Struct. MRI Data* 74–81 (2014).

24. Durrleman, S. *et al.* Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *Neuroimage* **101**, 35–49 (2014).
25. Kristensen, E., Parsons, T. E., Hallgrímsson, B. & Boyd, S. K. A novel 3-D image-based morphological method for phenotypic analysis. *IEEE Trans. Biomed. Eng.* **55**, 2826–2831 (2008).
26. Klingenberg, C. P. Size, shape, and form: concepts of allometry in geometric morphometrics. *Dev. Genes Evol.* **226**, 113–137 (2016).
27. Suri, S., Tompson, B. D. & Cornfoot, L. Cranial base, maxillary and mandibular morphology in Down syndrome. *Angle Orthod.* **80**, 861–869 (2010).
28. Fischer-brandies, H., Schmid, R. G. & Fischer-Brandies, E. Craniofacial development in patients with down's syndrome from birth to 14 years of age. *Eur. J. Orthod.* **8**, 35–42 (1986).
29. Fischer-Brandies, H. Cephalometric comparison between children with and without down's syndrome. *Eur. J. Orthod.* **10**, 255–263 (1988).
30. Robinson, C. & Terhune, C. E. Error in geometric morphometric data collection: Combining data from multiple sources. *Am. J. Phys. Anthropol.* **164**, 62–75 (2017).
31. Adams, D. C. & Collyer, M. L. Multivariate Phylogenetic Comparative Methods: Evaluations, Comparisons, and Recommendations. *Syst. Biol.* **67**, 14–31 (2018).
32. Gunz, P., Mitteroecker, P. & Bookstein, F. L. Semilandmarks in Three Dimensions. in *Modern Morphometrics in Physical Anthropology* 73–98 (Kluwer Academic Publishers-Plenum Publishers, 2006).
33. Indyk, P. & Motwani, R. Approximate nearest neighbors. in *Proceedings of the thirtieth annual ACM symposium on Theory of computing - STOC '98* 604–613 (ACM Press, 1998).
34. Jackson, I. T., Bayramiçli, M., Çelebiler, O. & Fendon, T. H. Effect of rigid fixation of

the temporozygomatic suture on craniofacial growth. *Eur. J. Plast. Surg.* **22**, 379–386 (1999).

35. Lu, X. *et al.* Spatial and temporal changes of midface in Apert's syndrome. *J. Plast. Surg. Hand Surg.* **53**, 130–137 (2019).
36. Hallgrímsson, B., Lieberman, D. E., Liu, W., Ford-Hutchinson, A. F. & Jirik, F. R. Epigenetic interactions and the structure of phenotypic variation in the cranium. *Evol. Dev.* **9**, 76–91 (2007).
37. Klingenberg, C. P. MorphoJ: An integrated software package for geometric morphometrics. *Mol. Ecol. Resour.* **11**, 353–357 (2011).
38. Bookstein, F. L. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 567–585 (1989).
39. Baerentzen, J. A. Robust generation of signed distance fields from triangle meshes. in *Fourth International Workshop on Volume Graphics, 2005.* 167–239 (IEEE, 2005).
40. Schlager, S. Morpho and Rvcg -- Shape Analysis in {R}. in *Statistical Shape and Deformation Analysis* (eds. Zheng, G., Li, S. & Szekely, G.) 217–256 (Academic Press, 2017).
41. Mangan, A. P. & Whitaker, R. T. Partitioning 3D surface meshes using watershed segmentation. *IEEE Trans. Vis. Comput. Graph.* **5**, 308–321 (1999).
42. Vollmer, J., Mencl, R. & Muller, H. Improved Laplacian Smoothing of Noisy Surface Meshes. *Comput. Graph. Forum* **18**, 131–138 (1999).
43. Gower, J. C. Generalized procrustes analysis. *Psychometrika* **40**, 33–51 (1975).
44. Fi, M. O. & Garriga, G. C. Permutation Tests for Studying Classifier Performance Markus Ojala. *J. Mach. Learn. Res.* **11**, 1833–1863 (2010).
45. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).
46. Yaniv, Z., Lowekamp, B. C., Johnson, H. J. & Beare, R. SimpleITK Image-Analysis Notebooks: a Collaborative Environment for Education and Reproducible Research. *J. Digit. Imaging* **31**, 290–303 (2018).

47. Schroeder, W., Martin, K. & Lorensen, B. *The Visualization Toolkit: An Object Oriented Approach to 3D Graphics*. **4**, (Kitware, 2006).

48. Nolden, M. *et al.* The Medical Imaging Interaction Toolkit: challenges and advances. *Int. J. Comput. Assist. Radiol. Surg.* **8**, 607–620 (2013).

49. Gori, P. *et al.* A Bayesian framework for joint morphometry of surface and curve meshes in multi-object complexes. *Med. Image Anal.* **35**, 458–474 (2017).

50. Pedregosa, F. Varoquaux, G. Gramfort, A. Michel, V. Thirion, B. Grisel, O. Blondel, M. Prettenhofer, P., Weiss, R. Dubourg, V. Vanderplas, J. Passos, A. & Cournapeau, D. Brucher, M. Perrot, M. Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

Methods

Mice and imaging

C57BL/6J.129P2-Dp(16Lipi-Zbtb21)1TybEmcf (Dp1Tyb) mice[14] were bred at the MRC Harwell Institute. All mice were backcrossed to C57BL/6JNimr for at least 10 generations. All animal work was approved by the Ethical Review panel of the Francis Crick Institute and was carried out under Project Licences granted by the UK Home Office. Heads from 20 mice at 16 weeks of age were used (10 Dp1Tyb and 10 WT, 5 male and 5 female of each genotype). However, 4 subjects were excluded from the analysis due to fractures in either the mandible or skull. Heads were prepared for micro-computed tomography (µCT) by fixation in PFA and then scanned at a 25 µm resolution using a µCT 50 (Scanco).

Landmark-based morphometric analysis

Landmark Acquisition

Three-dimensional locations of 68 anatomical landmarks for the cranium and 17 landmarks for the mandible (Supplementary Fig. 1) as previously defined by Hallgrimsson and colleagues[36] were placed onto 3D reconstructions of µCT images using Microview (Parallax Innovations). Landmarks were placed manually and verified by checking orthogonal planar views of the subject.

Statistical analysis of landmark dataset

Landmarks for all subjects were aligned using Procrustes superimposition, and distances between landmarks for each subject were analysed by Principal Component Analysis (PCA) using MorphoJ[37] to visualise group separation by shape. To quantify significance in shape differences between genotypes we used the Procrustes Distance Multiple Permutations test (1000 iterations) within MorphoJ. Centroid size was calculated as the square root of the sum of the squared distances from each landmark to the centroid i.e. the centre of mass of all

landmarks of a given specimen and was used to compare size differences between WT and Dp1Tyb crania and mandibles. The statistical significance of such size differences was calculated using a 2-tailed unpaired t-test.

Shape changes after Procrustes scaling were visualised as heatmaps. Using the Morpho R package function tps3d (documentation for which can be found at https://www.rdocumentation.org/packages/Morpho/versions/2.6/topics/tps3d), mean landmark sets of each genotype group were used to interpolate an average mesh, using a thin-plate spline method [38], thus providing a mesh for both the mean WT shape and mean Dp1Tyb shape. The Morpho R meshdist function (https://www.rdocumentation.org/packages/Morpho/versions/2.6/topics/meshDist) was then used to create the heatmaps. The function first calculates the distances of the reference meshes vertices (in this case mean WT mesh) to that of the target mesh (mean Dp1Tyb mesh). Then, using a previously proposed algorithm[39], the distances were given a negative value if inside the reference mesh or a positive value if outside. A vector containing blue and red colour values was assigned to the negative and positive values respectively[40].

Landmark-free morphometric analysis

As an alternative to landmark population comparisons, statistical analysis of anatomical shapes can be achieved using so-called atlas-based approaches, which consist of estimating an anatomical model (i.e. template) as the mean of a set of input shapes (rather than point clouds) and quantifying its variation in a test population as deformations. This was previously achieved in a reproducible and robust landmark-free manner by Durrleman et al[24]. This approach bypasses a number of problems associated with mesh point-to-point comparison by representing deformation between shapes as the diffeomorphic transformation of flow fields, i.e. currents over the mesh surface. Currents are parameterised by a set of control points in space and initial velocities, or momenta. By means of a gradient descent optimisation scheme, the method is able to produce a statistical atlas of the population of shapes. An atlas refers to

a mean template shape, a set of final control point positions, and momenta parameterising the displacements between the template to each initial individual shape. In the following sections we describe the different steps to achieve such analysis.

Landmark-free Atlas Construction

Step 1 (Fig. 1, Table 1, Supplementary Appendix 1). Despite the fact that images acquired using µCT show good bone contrast, they often include the presence of artificial objects (noise and debris in the specimen), small holes and cartilage that need be excluded in order to obtain consistently comparable final surface meshes. To extract the surface meshes, a series of image processing steps were applied. After a thresholding operation to extract the skull, "morphological opening" and "closing" were performed on the binary mask to remove internal cartilage structures (Supplementary Fig. 2). Removal of spurious objects was achieved by clustering, categorising all of the connected components in an image by size and retaining only the largest component. Skulls were segmented using bone density to isolate the mandible (whose density is higher than that in the rest of the skull). However, this segmentation of the mandible can happen improperly and may include parts of the temporal bone which must be cleaned and removed manually. Mandible meshes were extracted from the skull binary mask semi-automatically using Watershed segmentation[41].

Step 2. The meshes were produced using marching cubes on the binary images, followed by a surface Laplacian smoothing[42] and finally decimation of a factor of 1/80 to decrease mesh resolution and reduce overall computation time.

Step 3. The atlas construction necessitates the production of aligned meshes from the input µCT images as a pre-processing step. Minimally 3 pairs of landmarks were chosen to align meshes using a Procrustes technique[43] using a rigid body-plus-scaling transformation model (similarity alignment) or a rigid body transformation without scaling (rigid body alignment).

Step 4. As previously described[24], the atlas construction major hyper-parameters consist of (1) the size of the Gaussian kernel used to represent shapes in the varifold of currents, denoted σ_W, and (2) the number of control points, denoted N_{cp}. Control points can be thought of as unbiased landmarks, initially they are spaced on a regular grid evenly but move to areas of greatest variation. Thus as much of the occurring shape change is captured in an unbiased manner. σ_W can be seen as the precision at which the shape deformation is described. N_{cp} denotes the sampling density in space. In our experiments, we fixed σ_W = 2 mm. Using that scale, the derived number of control points Ncp was 350 and 940 for the mandibles and the cranium analyses respectively. All subjects' meshes were used for the construction of the atlas. The optimisation procedure typically converged within 150 iterations in approximately 20h on a 10 core cpu for the mandible. The atlas construction produces the following outputs:

- **Control Points**: For each subject, the N_{cp} final control points locations {c}, that correspond to areas that describe the population's most important variability in shape.
- **Momenta**: For each subject, the N_{cp} final momenta {m} associated to each control point, that describe the shape directional variation from the population average.
- **Template**: Mean shape of the population.

Landmark-free dataset statistical analysis

Step 5 (Fig. 1). The atlas outputs provide a dense amount of information that can be used for various statistical analyses. Centroid size was calculated as the square root of the sum of the squared distances from each mesh node to the centroid of all nodes of a given specimen. Non-linear Kernel PCA with dimension 5 was applied to the set of momenta produced from the atlas of the population, in order to find the principal modes of variation of the entire population. The resulting output provided a way to compare these results with the landmark-based PCA analysis. We projected the subjects onto the feature space for comparison purposes. Such projection provides dense information of shape differences between the two sub-populations. Local magnitude of the momenta interpolated at the template mean mandible (or cranium) mesh point locations allow for additional qualitative interpretation of shape

differences between groups. Stratified k-fold cross-validation analysis was performed on the PCA data to evaluate the statistical power of classification between the two groups. Significance of the classification score was tested using a multiple permutations test at 1000 iterations[44]. To assess overfitting, subjects were randomly partitioned into two groups ("scrambled groups") and PCA analysis performed to generate inter-group vectors. The distribution of the vector magnitudes was tested for Normality using the Shapiro-Wilk test.

Code availability

Python scripts and documentation for the landmark-free morphometric analysis are freely available on GitHub (https://gitlab.com/ntoussaint/landmark-free-morphometry) and their use is also described in Supplementary Appendix 1.

Figure Legends

Figure 1. Stages of processing in the landmark-free approach. Step 1, extraction of region of interest. Initial thresholding of μCT image was used to make a binary mask and regions of the mask were separated by bone density using secondary thresholding, with some manual clean-up based on known anatomy. A region of interest (in this example the mandible in red) was chosen for further analysis. Step 2, mesh generated and decimated by a factor of 0.0125 to reduce data file size. Step 3, meshes of all subjects aligned either using rigid body alignment with no scaling or similarity alignment with scaling. Step 4, atlas construction. Step 5, statistical analysis and visualisation of shape data.

Figure 2. Decreased size and altered shape of Dp1Tyb crania and mandibles. A-D Centroid sizes of WT and Dp1Tyb crania (**A, B**) and mandibles (**C, D**) determined using landmark-based (**A, C**) and landmark-free (**B, D**) methods. Data shown as box and whiskers plots indicating the 25% and 75% centiles (box), range of all data points (whiskers) and the median (black line). Statistical significance was calculated using a two-tailed unpaired t-test; **** $P < 0.0001$. **E-H** Principal component analysis (first two components) of Procrustes aligned shapes of WT and Dp1Tyb crania (**E, F**) and mandibles (**G, H**) determined using landmark-based (**E, G**) and landmark-free (**F, H**) methods. Statistical significance of the differences between the WT and Dp1pTyb groups of samples was calculated using a permutation test, and *P*-values are shown on the plots. Sample size: n = 8 of each genotype.

Figure 3. Visualisation of altered shape of Dp1Tyb crania and mandibles determined using landmark-based analysis. A-F Mean landmark configurations of WT (Orange) and Dp1Tyb (Cyan) crania (**A-D**) and mandibles (**E, F**), showing lateral (**A**), superior (**B**), inferior (**C**) and rear (**D**) views of the cranium and lingual (**E**) and buccal (**F**) views of the mandible. **G-L** Displacement heatmaps after global size differences have been regressed out, produced by

superimposing the mean shapes of the WT and Dp1Tyb crania (**G-J**) and mandibles (**K,L**), showing lateral (**G**), superior (**H**), inferior (**I**) and rear (**J**) views of the cranium and lingual (**K**) and buccal (**L**) views of the mandible. Red and Blue represent the distribution of expansion and contraction respectively in Procrustes (shape) distance. Regions of the Dp1Tyb mesh outside the WT mesh are coloured red, whereas any parts inside the WT mesh are coloured blue, thereby showing displacement relative to WT.

Figure 4. Visualisation of altered shape of Dp1Tyb crania and mandibles determined using landmark-free analysis. Displacement heatmaps after global size differences have been regressed out estimated using output momenta from the current-based atlas construction, showing locations of shape differences between WT and Dp1Tyb crania (**A - D**) and mandibles (**E-H**) , showing lateral (**A**), superior (**B**), inferior (**C**) and rear (**D**) views of the cranium and lingual (**E, G**) and buccal (**F, H**) views of the mandible. **A-F** are on the same colour scale; the colour scale in **G, H** maximises contrast for the mandibles only. Arrows **R** and **A** indicate the ramus and Angular process respectively.

Figure 5. Altered surface stretch in Dp1Tyb crania and mandibles determined using landmark-free analysis. Heatmaps showing surface stretch between WT and crania (**A - D**) and mandibles (**E, F**) , showing lateral (**A**), superior (**B**), inferior (**C**) and rear (**D**) views of the cranium and lingual (**E**) and buccal (**F**) views of the mandible. Stretch changes were estimated using output momenta from the current-based atlas construction. Note that global size changes were regressed out. Arrow **AB** indicates the auditory bulla.

Step	Sub-step	Function	Parameters	Tool	Reference
Step 1. Pre-processing	Binary segmentation	Binary threshold	threshold = 10000	FSL	45
		Closing	kernel size = 5		
		Opening	kernel size = 3		
		Clustering and extraction of largest component			
		Fill holes			
	Object parcelation	Watershed	watershedlevel = 0.22	SimpleITK	46
Step 2. Mesh extraction	Mesh extraction	Marching cubes		VTK	47
		Smoothing	number of iterations = 20 relaxation factor = 0.6		
Step 3. Mesh Alignment	Alignment points	Reference points used to align meshes		MITK	48
	Mesh Alignment	Centroid size		VTK/Numpy	
		vtkProcrustesAlignmentFilter	Rigid body or similarity	VTK	
Step 4. Atlas	Atlas construction	Atlas construction	data-sigma = 10 kernel width (template) = 2.5mm kernel width (subject) = 2mm number of time points = 10 max iterations = 150 number of threads = 19	Deformetrica	49
Step 5. Statistical analysis	Principal component analysis	Kernel PCA	number of components = 5	Scikit-learn	50
		Kernel PCA lambdas (eigen values)			
		Support vector machine			
		Permutation test			
		Cross validation			
	Momenta projection / morph	Shooting	template kernel width = 2.0 deformation kernel width = 1.5	Deformetrica	

Table 1. Summary of steps in the landmark-free analysis pipeline and chosen parameters at each step.

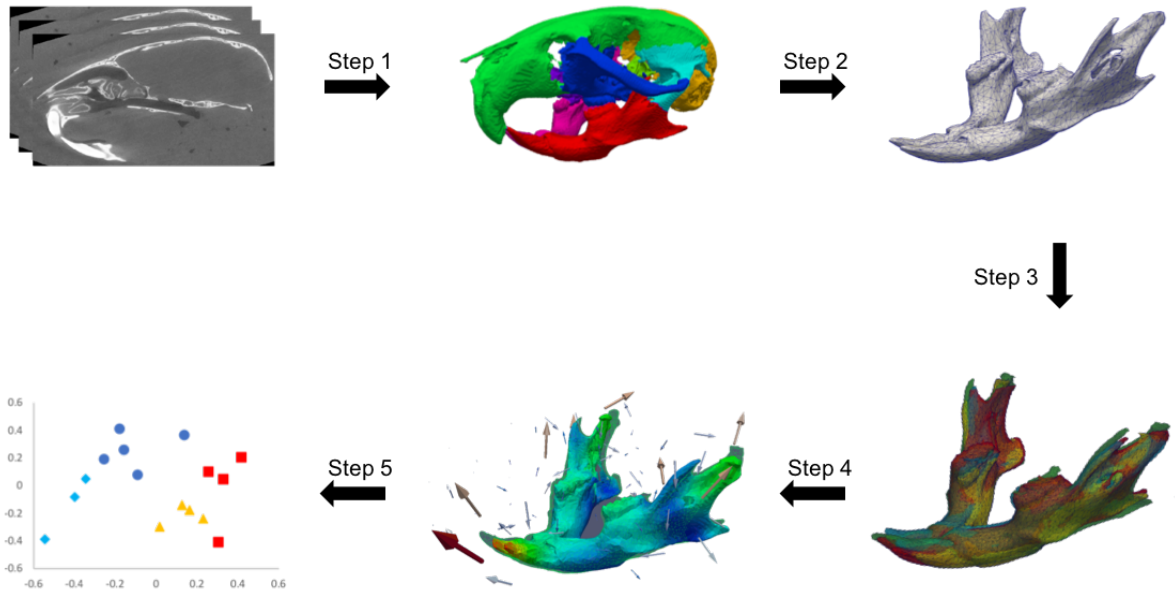

Step 1
Step 2
Step 3
Step 4
Step 5
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.6
-0.4
-0.2
0
0.2
0.4
0.6

Figure 1

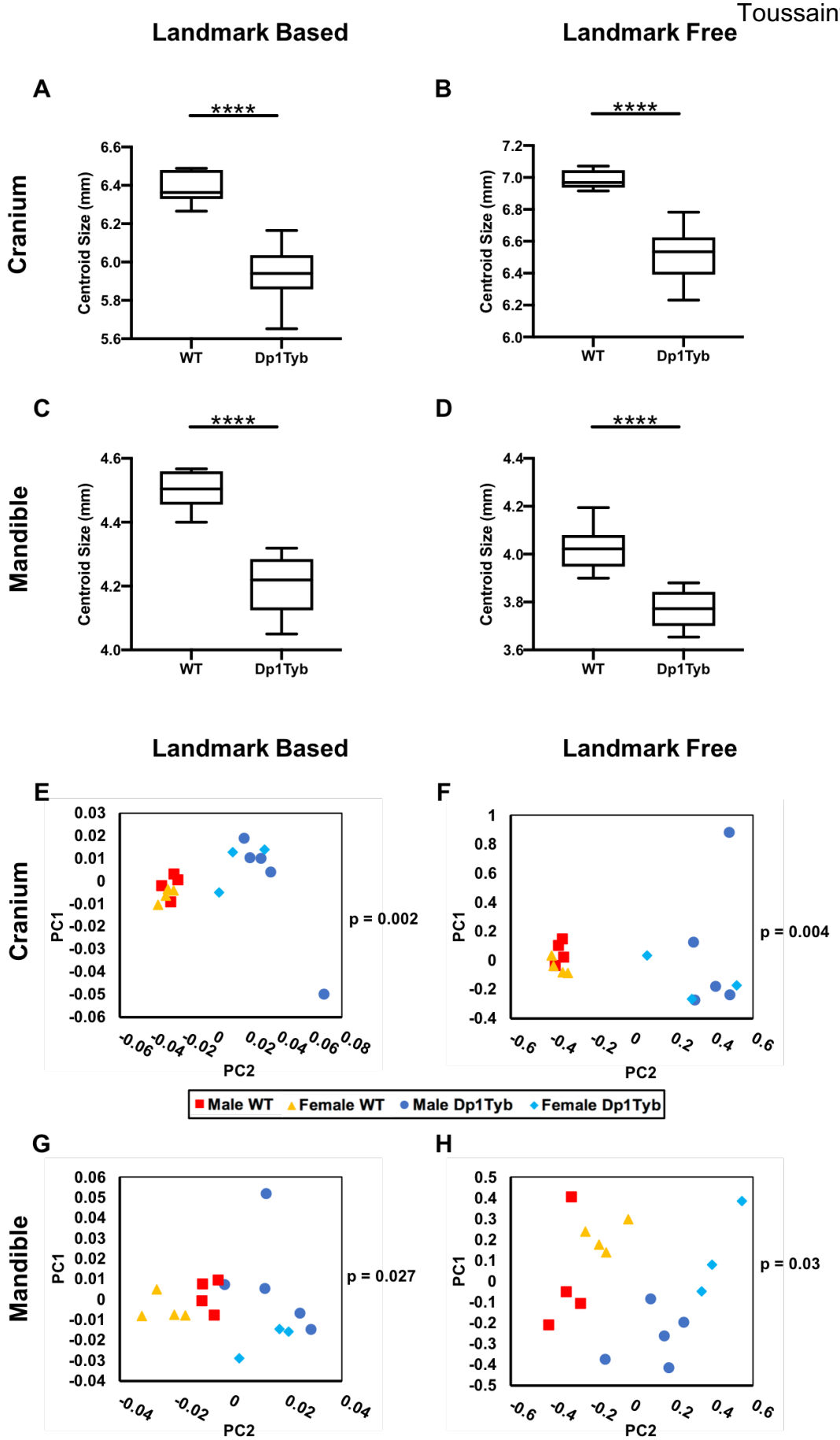

Landmark Based
Landmark Free
A
B
C
D
Cranium
Mandible
Centroid Size (mm)
****
WT
Dp1Tyb
E
F
G
H
PC1
PC2
p = 0.002
p = 0.004
p = 0.027
p = 0.03
Male WT
Female WT
Male Dp1Tyb
Female Dp1Tyb

Figure 2

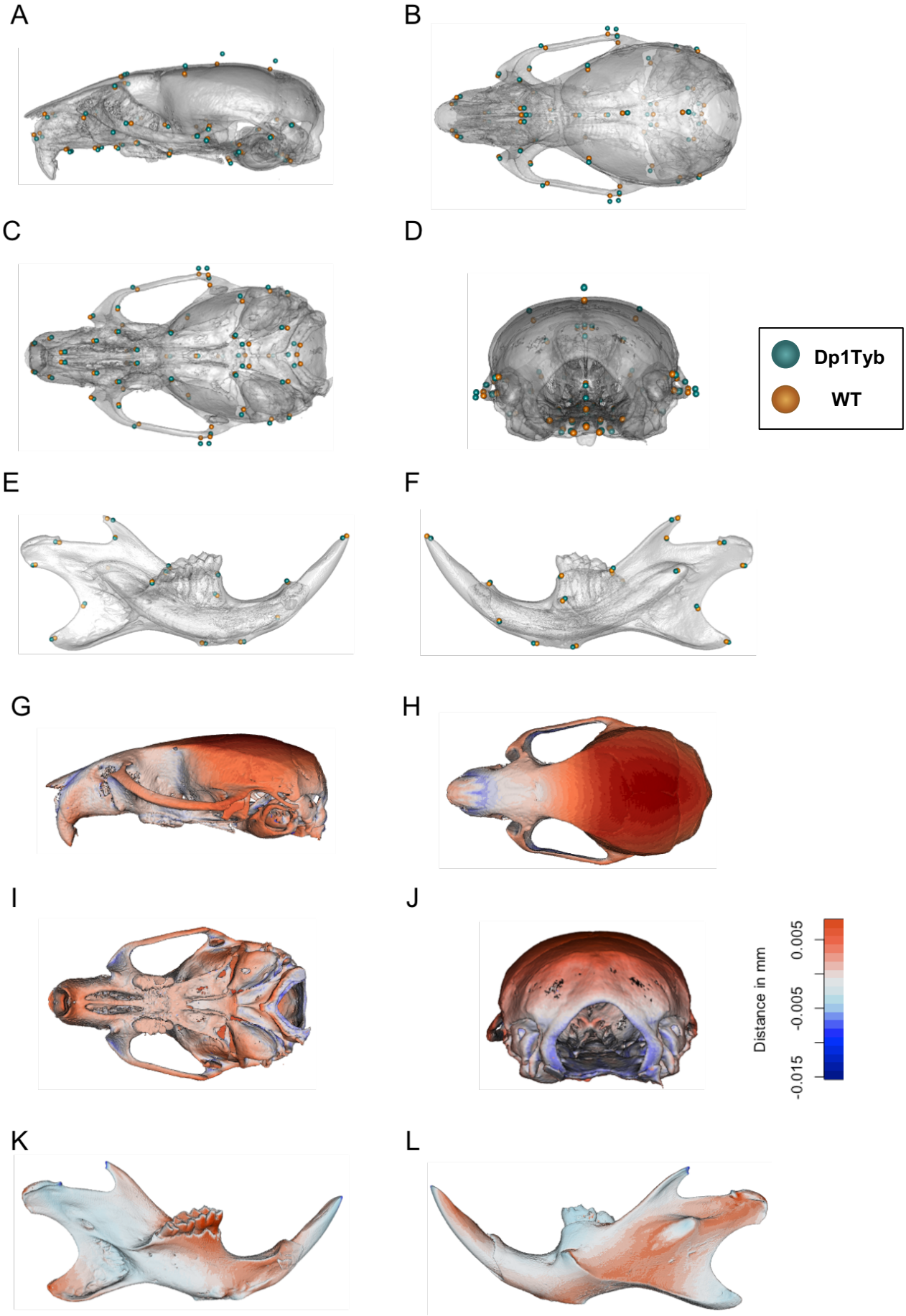


Figure 3

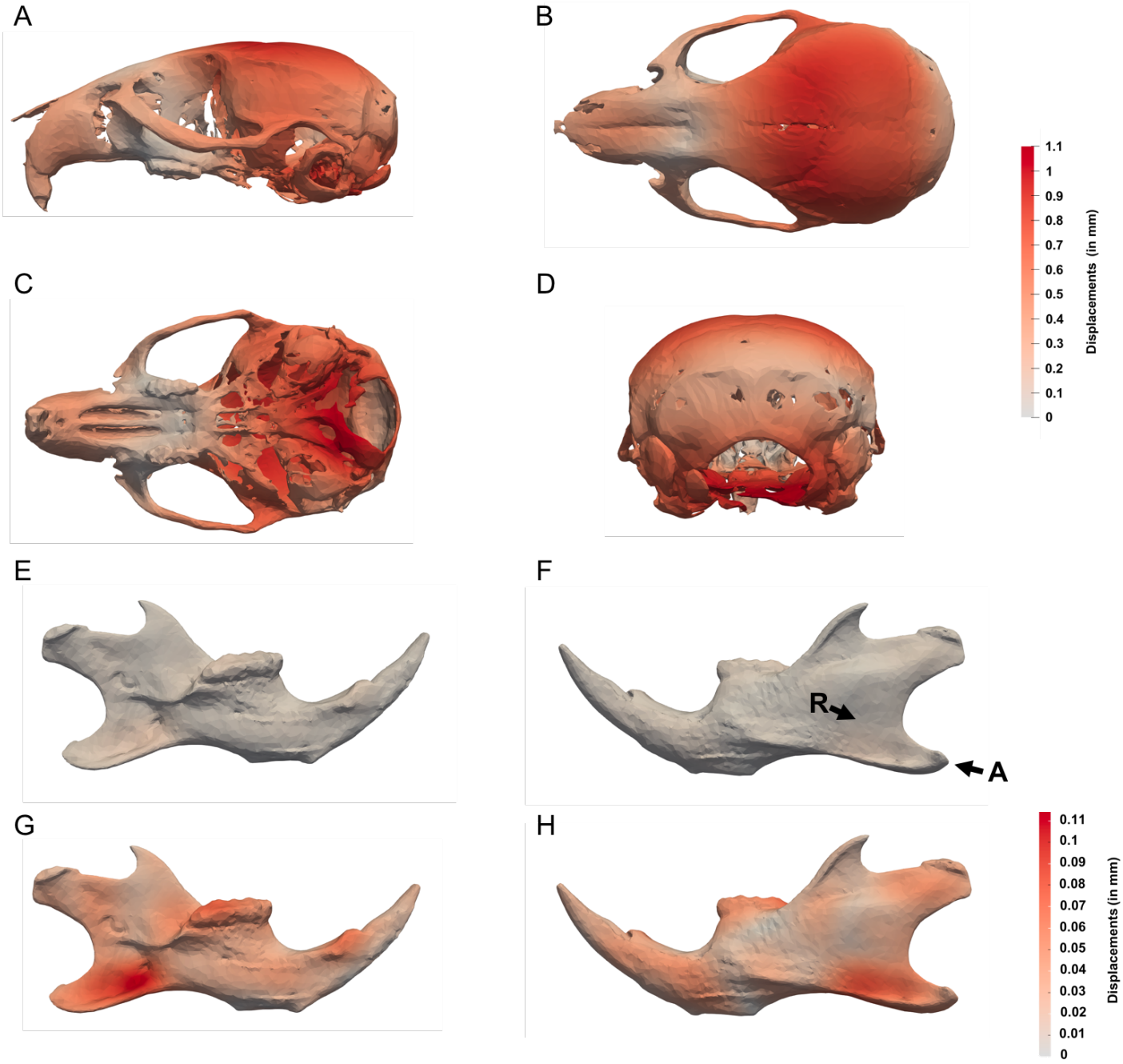

A
B
C
D
E
F
R
A
G
H
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Displacements (in mm)
0.11
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Displacements (in mm)

Figure 4

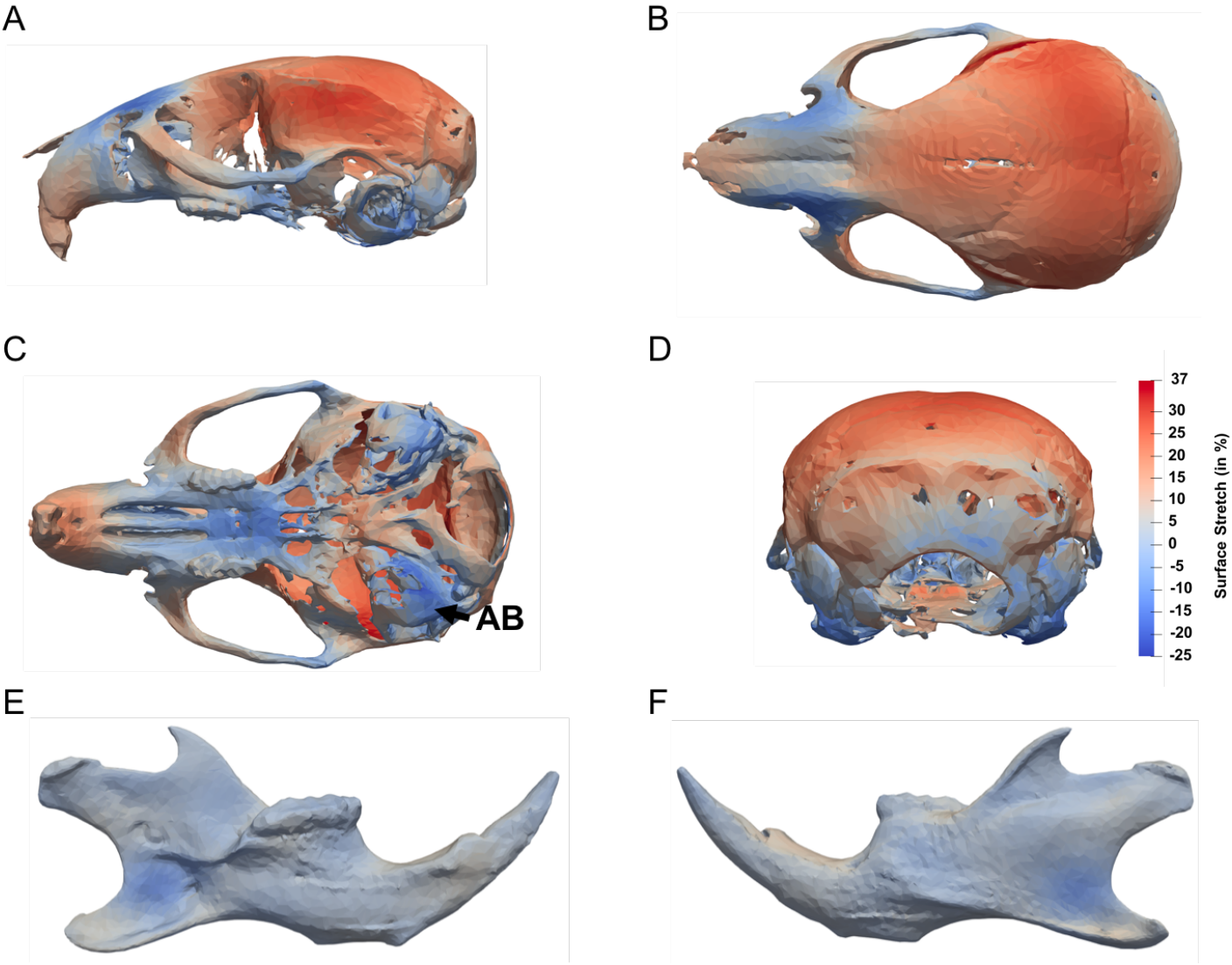


Figure 5.

Supplementary Information

A landmark-free morphometrics pipeline: Application to characterise a Down syndrome model skull phenotype

Nicolas Toussaint[1]*, Yushi Redhead[2,3]*, Wei Liu[4], Elizabeth M. C. Fisher[5], Benedikt Hallgrimsson[4], Victor L.J. Tybulewicz[3,6§], Julia A. Schnabel[1§], Jeremy B.A. Green[2§]

[1] School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
[2] Centre for Craniofacial Biology & Regeneration, King's College London, London, UK
[3] The Francis Crick Institute, London NW1 1AT, UK
[4] Department of Cell Biology & Anatomy, University of Calgary, Calgary AB T2N 4N1, Canada
[5] UCL Institute of Neurology, London, WC1N 3BG, UK
[6] Department of Immunology & Inflammation, Imperial College London, London W12 0NN, UK

*Equal contribution

[§]Correspondence to:
JAS (email: julia.schnabel@kcl.ac.uk), VLJT (email: Victor.T@crick.ac.uk), or to
JBAG (email: jeremy.green@kcl.ac.uk)

Supplementary Videos

Supplementary Video 1.

Cranium morph including size

Cranium morph between the mean WT and Dp1Tyb shapes, where size has not been regressed out. Derived from landmark-free analysis.

Supplementary Video 2.

Mandible morph including size

Mandible morph between the mean WT and Dp1Tyb shapes, where size has not been regressed out. Derived from landmark-free analysis.

Supplementary Video 3.

Cranial morph displacement heatmap

Cranial morph between the mean WT and Dp1Tyb shapes, where size has been regressed out. Red colour change indicates the magnitude of displacement in mm. Derived from landmark-free analysis.

Supplementary Video 4.

Mandible morph displacement heatmap

Mandible morph between the mean WT and Dp1Tyb shapes, where size has been regressed out. Red colour change indicates the magnitude of displacement in mm. Derived from landmark-free analysis.

Supplementary Video 5.

Exaggerated Mandible Morph

Mandible morph between the mean WT and Dp1Tyb shapes, where size has been regressed out and the magnitude of the deformation has been exaggerated by a factor of 3. Derived from landmark-free analysis.

Supplementary Video 6.

Cranial morph stretch heatmap

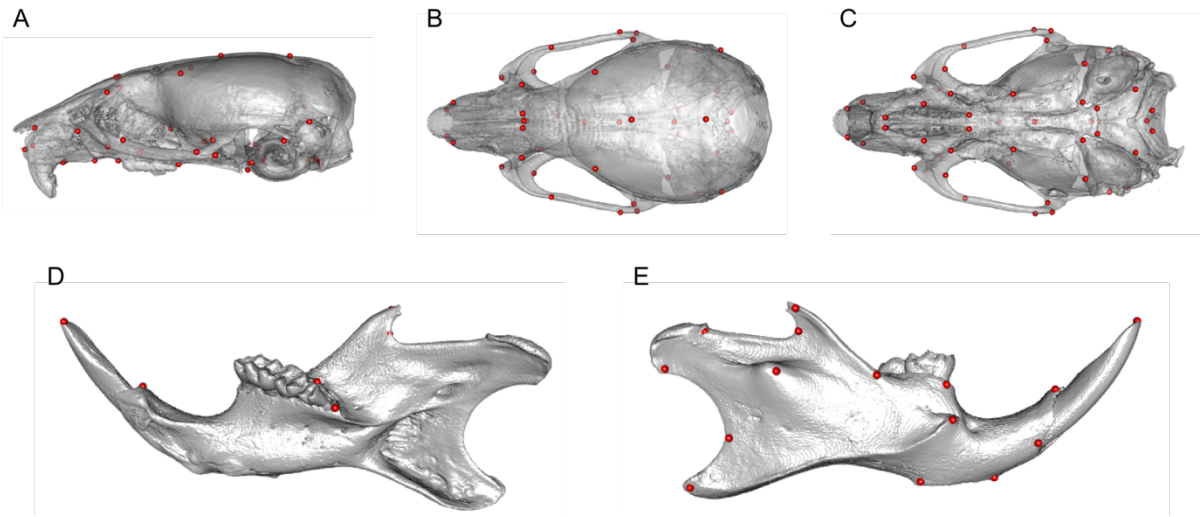
Cranial morph between the mean WT and Dp1Tyb shapes, where size has been regressed out. Red and Blue colour change indicate expansion and contraction respectively as a percentage. Derived from landmark-free analysis.

Supplementary Video 7.

Mandible morph stretch heatmap

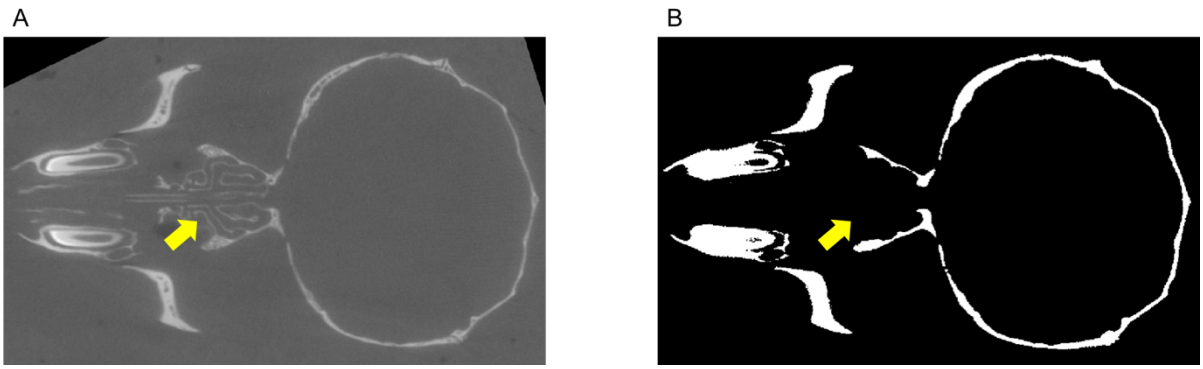
Mandible morph between the mean WT and Dp1Tyb shapes, where size has been regressed out. Red and Blue colour change indicate expansion and contraction respectively as a percentage. Derived from landmark-free analysis.

Supplementary Figures



Supplementary Figure 1.

Location of landmarks used in landmark-based morphometric analysis. A-E Anatomic locations of landmarks shown on a 3D digitized image of a µCT scan as previously defined [1,2]. 68 cranial landmarks are shown on lateral (A), superior (B), and inferior (C) views of the cranium. 17 mandibular landmarks are shown on lingual (D) or buccal (E) views of the mandible.



Supplementary Figure 2.

Planar view of mouse head captured by µCT. A Image pre-processing. **B** Image post-processing. Yellow arrow indicates cartilaginous element (nasal turbinate).

Supplementary Table

		Mean Centroid Size in mm (± SEM)		difference	
		WT	Dp1Tyb	absolute (mm)	percentage
Landmark-based	Cranium	6.387±0.061	5.939±0.061	0.448	7.01%
	Mandible	4.502±0.038	4.204±0.038	0.298	6.62%
Landmark-free	Cranium	6.987±0.063	6.514±0.063	0.473	6.77%
	Mandible	4.025±0.041	3.768±0.041	0.257	6.39%

Supplementary Table.

Comparison of centroid sizes of crania and mandibles from WT and Dp1Tyb mice determined using landmark-based or landmark-free methods.

Supplementary References

1. Hallgrímsson, B., Lieberman, D. E., Liu, W., Ford-Hutchinson, A. F. & Jirik, F. R. Epigenetic interactions and the structure of phenotypic variation in the cranium. *Evol. Dev.* **9**, 76–91 (2007).
2. Hallgrímsson, B., Dorval, C. J., Zelditch, M. L. & German, R. Z. Craniofacial variability and morphological integration in mice susceptible to cleft lip and palate. *J. Anat.* **205**, 501–517 (2004).