Published in final edited form as:
Alcohol. 2013 August ; 47(5): 405–415. doi:10.1016/j.alcohol.2013.04.005.

Effect of Prenatal Alcohol Exposure on Bony Craniofacial Development: A Mouse MicroCT Study

Li Shen, PhD[1,2], **Huisi Ai, MS**[1], **Yun Liang, PhD**[1], **Xiaowei Ren, MS**[1,3], **Charles Bruce Anthony, PhD**[4], **Charles R. Goodlett, PhD**[2,5], **Richard Ward, PhD**[6], and **Feng C. Zhou, PhD**[2,4]

Feng C. Zhou: imce100@iupui.edu

[1]Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN 46202, USA

[2]Stark Neurosciences Research Institute, Indiana University School of Medicine, Indianapolis, IN 46202, USA

[3]Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN 46202, USA

[4]Department of Anatomy and Cell Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

[5]Department of Psychology, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA

[6]Department of Anthropology, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA

Abstract

Craniofacial bone dysmorphology is an important but under-explored potential diagnostic feature of fetal alcohol spectrum disorders. This study used longitudinal MicroCT 3D imaging to examine the effect of prenatal alcohol exposure on craniofacial bone growth in a mouse model. C57BL/6J dams were divided into 3 groups: alcohol 4.2% v/v in PMI® liquid diet (ALC), 2 weeks prior to and during pregnancy from embryonic (E) days 7-E16; pair-fed controls (PF), isocalorically matched to the ALC group; chow controls (CHOW), given *ad libitum* chow and water. The MicroCT scans were performed on pups on postnatal days 7 (P7) and P21. The volumes of the neurocranium (volume encased by the frontal, parietal, and occipital bones) and the viscerocranium (volume encased by the mandible and nasal bone), along with total skull bone volume, head size, and head circumference were evaluated using general linear models and discriminant analyses. The pups in the alcohol-treated group, when compared to the chow-fed controls (ALC vs. CHOW) and the isocaloric-fed controls (ALC vs. PF), showed differences in head size and circumference at P7 and P21, the total skull volume and parietal bone volume at P7, and volume of all the tested bones except nasal at P21. There was a growth trend of ALC < CHOW and ALC < PF. While covarying for gender and head size or circumference, the treatment affected the total skull and mandible at P7 (ALC > CHOW), and the total skull, parietal bone, and

Please address correspondence to: Li Shen, PhD, Center for Neuroimaging, Department of Radiology and Imaging Sciences, IU School of Medicine, 355 W 16th Street, Suite 4100, Indianapolis, IN 46202, USA, Phone: (317) 963-7504, Fax: (317) 963-7547, shenli@iu.edu.

occipital bone at P21 (ALC < CHOW, ALC < PF). While covarying for the P7 measures, the treatment affected only the 3 neurocranial bones at P21 (ALC < CHOW, ALC < PF). Discriminant analysis sensitively selected between ALC and CHOW (AUC = 0.967), between ALC and PF (AUC = 0.995), and between PF and CHOW (AUC = 0.805). These results supported our hypothesis that craniofacial bones might be a reliable and sensitive indicator for the diagnosis of prenatal alcohol exposure. Significantly, we found that the neurocranium (upper skull) was more sensitive to alcohol than the viscerocranium (face).

Keywords

prenatal alcohol; craniofacial bone; facial dysmorphology; diagnosis

Introduction

Heavy maternal alcohol abuse during pregnancy can result in fetal alcohol syndrome (FAS), in which the affected children have facial dysmorphology, neurodevelopmental deficits, and growth retardation. The key facial features used for clinical diagnosis of FAS include short palpebral fissures, a thin upper vermillion, and a smooth philtrum (Astley and Clarren, 2001). The prevalence of FAS has been estimated to be between 0.5-2 per 1000 live births (May and Gossage, 2001; Sampson et al., 1997), and more recently as high as 2 to 7 per 1000 children in US school-age populations (May et al., 2009). However, neurodevelopmental deficits resulting from prenatal alcohol exposure may affect as many as 10 times as many children but this cannot be accurately determined because many of these children do not express the facial dysmorphology which is necessary for diagnosis of FAS (Hoyme et al., 2005; May et al., 2009; Sampson et al., 1997). The term fetal alcohol spectrum disorders (FASD) is now used to include the non-dysmorphic alcohol-affected children together with children with full or partial FAS. Because the non-dysmorphic children are the majority of FASD children and they cannot be identified on the basis of the traditional facial dysmorphology, the early identification of these FASD children poses a significant challenge to the field.

One approach to the early identification of FASD is to validate reliable measures of prenatal alcohol effects on growth and development that may be used as indicators of developmental damage. Although clinical assessment of facial dysmorphology relies primarily on the configuration of soft tissue, quantitative analysis of the craniofacial bones underlying this soft tissue may provide a reliable and previously under-utilized diagnostic tool to estimate the quantity, frequency, pattern, and duration of prenatal alcohol exposure. Craniofacial bones are derived from either neural crest cells during neurulation (Huang and Saint-Jeannet, 2004) or from splanchnic and lateral plate mesoderm (Moore et al., 2011), and are highly sensitive to environmental insult leading to cytotoxicity, retarded migration, and apoptosis (Chen and Sulik, 1996; Chen et al., 2011; Debelak and Smith, 2000; Hassler and Moran, 1986; Rovasio and Battiato, 1995, 2002; Sulik et al., 1988).

The current study asked if craniofacial bone growth would be affected by prenatal alcohol exposure, and whether different craniofacial bones showed different degrees of dysmorphology over the course of development. Experimental animal models can display the effects of the relationships between developmental timing, dosage of alcohol exposure, and the variable phenotypes of FASD (Goodlett et al., 2005). The craniofacial dysmorphology induced by prenatal alcohol exposure has been modeled in the mouse both for facial and skeletal development and has been found to closely parallel the dysmorphology observed in humans (Hernandez-Guerrero et al., 1998; Robin and Zackai, 1994; Sulik, 1984; Sulik and Johnston, 1983). It has also been demonstrated that head

circumference is significantly reduced in most children with FAS (Rosett, 1980), and this is now a standard proxy measure for reduced brain growth used for diagnosis of FAS (Hoyme et al., 2005). Analysis of the structure of the mouse skull could lead to improvements in identifying children affected by FASD, via detection of dysmorphology in the bony structures of the head and face.

Micro-computed tomography (MicroCT) is used in this study because bone tissue produces CT images that are distinct from CT images produced by all other tissues. The high resolution capability of the MicroCT system is thus suited to quantify skeletal morphometry. MicroCT images mineralized animal tissues (Neues and Epple, 2008) and can also provide quantitative 3D images of soft tissues using a small number of contrast stains (Parsons et al., 2008). It has been used to study craniofacial dysmorphology (Kaminen-Ahola et al., 2010; Metscher, 2009) and embryonic morphology (Schmidt et al., 2010) in mouse models.

As an initial step toward a more complete understanding of prenatal alcohol effects on craniofacial bone development, this study examined the entire skull comprising the bones of the neurocranium (frontal, parietal and occipital), and 2 of the 3 viscerocranial bones of the face (the nasal bone and the mandible). (The maxilla could not be included due to the difficulty of defining its boundaries with MicroCT). We hypothesized that the development of the craniofacial skeleton would be adversely affected by prenatal alcohol exposure.

Materials and Methods

Animals and Treatments

C57BL/6J (Jackson Laboratory) mice were used. All animals were handled in accordance with Indiana University Animal Care and Use Guidelines (IACUC). Adequate measures were taken to minimize pain or discomfort. Upon arrival, the mice were acclimated to a reverse light/dark cycle (light cycle from 10:00 PM to 10:00 AM) for 1 week before treatment began. Each dam was randomly assigned to one of three treatment conditions: a) alcohol exposure [ALC] via consumption of an alcohol-containing liquid diet (PMI® Purina Micro-stabilized Alcohol Rodent Liquid Diet LD 101A, Purinamills, Richmond, Indiana) both pre-mating (2 weeks prior to pregnancy) and during pregnancy from embryonic days (E) 7-E16, b) pair-fed [PF], using PMI® liquid diet with maltose-dextrin isocalorically substituted for alcohol with daily volumes matched to alcohol-consuming dams, and c) CHOW, *ad libitum* mouse chow (Teklad Global 18% protein extruded rodent diet, 2018sX, Harlan, Indianapolis, IN) and water throughout the experiment. Pre-mating treatments were started at 19 ± 2 weeks of age. The alcohol liquid diets contained alcohol concentrations of 2.1% v/v on the initial 2 days followed by 4.2% v/v afterward. All liquid diet was sweetened with 5% w/v sucrose, as per the supplier's instructions. The liquid diets were administered using a 35-mL drinking tube (Dyets Inc., NY). All animals were given chow/water diets 4 days before mating procedures to lower the risk of miscarriage.

For timed pregnancies, 2 females were placed with 1 male for a 2-hour restricted period beginning at the start of the dark cycle (10:00 AM to 12:00 noon) as previously described (Anthony et al., 2010). Animals were checked after each mating for sperm plugs and identified as pregnant at detection of a sperm plug (embryonic day 0, E0). All animals were mated daily (until plugs were detected) over a period of 3 weeks. If no plug was observed during this period, animals were eliminated from the study.

The day of birth was designated as P0. Mouse pups were all fostered to surrogate mothers under chow feeding conditions to avoid potential nursing negligence and maternal factors from treated dams. To align the potential delayed birth of the ALC group, the fertilization

age was used with reference to postnatal birth age (e.g. E26 as P7). Maternal body weight and diet consumption were measured daily.

MicroCT Imaging

The MicroCT employed in this study was an EVS-R9 system (Enhanced Vision Systems Corp., London, Ontario, Canada) which operated at 50 kVp and 1 mA maximum cube current. In the standard mode of operation, the distance from detector to the source was 315 mm and source to object distance was 250 mm. The field of view in the long direction (cross section) was ~ 80 mm and in the short direction (longitudinal direction) was ~ 40 mm, corresponding to either 40-micron or 80-micron voxels with 2 × 2 or 4 × 4 binning in the detector. The image acquisition was performed *in vivo* with animals anesthetized with Isoflurane. The MicroCT scan yielded volumetric data for the entire mouse with 40 × 40 × 40 μm^3 voxel spacing. Prior to the subject scanning, a calibration scan for CT number accuracy was performed by scanning a phantom target containing air, water, and SB3 – a human cortical bone mineral-equivalent material. The phantom target's measured CT values were used to validate the scanner's linearity and to calibrate the CT numbers in the mouse data. During the scanning period of greater than 12 months, the scanner remained stable with less than 2% change of CT values observed with respect to the CT number of SB3, at a CT value of ~ 3200 HU (Hounsfield Unit).

Sample Inclusion

All delivered pups were culled to 6 per litter. This study included pups from 29 litters. Due to the limited scan capacity of the CT scanner per day, random sampling within a litter was performed when needed, which yielded variability in litter size. As a result, the total sample used in this MicroCT study included 30 ALC pups (from 7 litters), 44 PF pups (from 11 litters), and 49 CHOW pups (from 11 litters). MicroCT scans were successfully acquired for all pups at P7 (n = 123), and for most pups at P21 (n = 115, excluding 2 pups due to animal death and 6 pups due to scanner failure). A quality check (QC) of the MicroCT scans was performed to remove samples with severe motion artifacts. Table 1 shows the cross tabulations of the QC results and treatment groups for the P7 and P21 sample sets. A Pearson chi-square test revealed no significant relationship between QC results and treatment groups. Table 2(a-d) summarizes the total experimental sample (n = 123), the sample set with successfully extracted P7 measures (n = 75), the set with P21 measures (n = 78), and the set with both P7 and P21 measures (n = 56), respectively.

Image Analysis

Amira Software (Visage Imaging, Inc., San Diego, CA) was used for landmark labeling and bone segmentation (i.e., identifying individual bones from images). After 3D reconstruction of the mouse scan, bony tissues were segmented (i.e., identified) by a global threshold that was determined with a combination of visual inspection of the 2D images and the analysis of histograms. The bone tissues could be visually perceived in all 2D slice images. A threshold for bone tissues was set at 300 HU (Hounsfield Unit) after careful tests for all P7 and P21 pups. This threshold led to a clean cut of all bone tissues from the soft tissues and the surrounding ringing artifacts (spurious signals near sharp transitions in a scanned specimen). The segmented results were then re-examined in a slice-by-slice comparison with the original CT images to eliminate the residual artifact contribution to the image. Amira Software was used to visualize the craniofacial bone structures (Figure 1) as well as to extract the following 3D measurements by an experienced technician: 1) Twenty craniofacial anatomic landmarks were manually defined on the cranium (Figure 1), and 2) the entire skull (including cranium and mandible) as well as the frontal, parietal, occipital, mandible, and nasal bones were manually segmented and their volumes were calculated based on the segmentation results. Figure 2 shows a typical skull segmentation result with 5

individual bones labeled in different colors. Of note — all the bones were successfully segmented at both P7 and P21 except the nasal structure. For P7 pups, the nasal structure was not detectable in their MicroCT scans. For P21 pups, the nasal structure was successfully segmented in 70 of 78 MicroCT scans, and was not completely imaged in the remaining 8 scans. Thus the nasal measures were available on only 70 P21 pups.

Since the landmarks were well distributed on the skull and face, their cubic centroid size was calculated and used as a proxy measure for the total head size. The centroid size (Bookstein, 1991) was the size measure used to scale a configuration of landmarks and was defined as the square root of the sum of squared distances of a set of landmarks from the centroid. Among the 20 landmarks shown in Figure 1, 3 of them (left and right endpoints for bony mandible, and bregma) were not identifiable on some P7 samples; therefore only the remaining 17 landmarks were used to calculate the cubic centroid size as an estimate of the head size.

The head circumference was measured using the MxView software (Philips, Amsterdam) as the outer perimeter of the skull enclosure at the level of the nasion, where the Frankfort plane (http://en.wikipedia.org/wiki/Frankfurt_plane) was used to standardize this process (Figure 2).

Statistical Analysis

SAS 9.2 (SAS Institute Inc., Cary, NC) was used in statistical analyses for examining the relationships among the following variables: volumes of the total skull bone and 5 individual bones (frontal, parietal, occipital, mandible, nasal), head size, head circumference, sex, and treatment group (ALC, PF, CHOW).

Pearson chi-square tests were performed on the number of males and females in the 3 prenatal treatment groups at each age, to assess potential distribution differences for all the pups and the pups whose measures were successfully extracted at P7, at P21, or at both P7 and P21 (Table 2[b-d]). Pearson correlation analyses were conducted to examine the pairwise correlation among all 8 craniofacial measures, which included the 5 individual bones listed earlier, along with the skull volume, the head size, and the head circumference (Supplemental Table 1).

To examine the gender effect, a general linear model was performed for each measure at each stage (P7 or P21), where correlated measures due to litter effect were accounted for by incorporating litter as a nested factor within the gender group. Bonferroni-corrected alpha level was used as the significance threshold.

To examine the effect of prenatal treatment, a general linear model was performed for each measure (litter modeled as a nested factor within treatment). Bonferroni-corrected alpha level was used as the significance threshold. For any measure meeting the Bonferroni-corrected alpha level on the statistical test, the pair-wise comparisons among the 3 treatment groups within the measure were performed and multiple comparisons were corrected by the Sidak method. The following analytical procedure was applied to each of 4 scenarios. The *first scenario* examined treatment effects on all bone measures, head size, and head circumference at each stage. The *second scenario* examined treatment effects on all bone measures at each stage while covarying for gender and head size. The *third scenario* examined treatment effects on all bone measures at each stage while covarying for gender and head circumference. The *fourth scenario* examined the treatment effect on each of the 5 P21 bone measures while covarying for the corresponding P7 measure. In this scenario, the P7 variance was removed to concentrate on the growth amount between P7 and P21; this ANCOVA method was statistically more powerful than a simple ANOVA applied to P21-P7

(Vickers, 2001). In the first 3 scenarios, the P7 sample (n = 75, Table 2b) and the P21 sample (n = 78, Table 2c) were tested. In the fourth scenario, the longitudinal sample (n = 56, Table 2d) was tested.

Discriminant Analysis

Univariate and multivariate discriminant analyses of the craniofacial bone measures were examined for 3 pairwise comparisons (case vs. control): ALC vs. CHOW, ALC vs. PF, and PF vs. CHOW. Univariate Receiver Operating Characteristic (ROC) analysis was implemented using Matlab (The MathWorks, Inc., Natick, MA) and performed on each of the 5 bone measures. Area under the ROC curve (AUC) was reported. Multivariate Support Vector Machine (SVM) analysis using either linear or sigmoid function as its kernel (Cortes and Vapnik, 1995) was performed on each possible combination of 2 or 3 bone measures. Classification performance was calculated based on leave-one-out cross-validation results, in which one observation was omitted and the classification function was estimated based on the remaining n-1 observations. The data from the omitted observation were used in the estimated classification function and the resulting discriminative value was recorded. This was repeated for all observations. For each comparison, the set of predictors that yielded the best AUC was reported. Both univariate and multivariate analyses were performed on the following bone measures: 1) P7 measures, 2) P21 measures, and 3) longitudinal growth measures (P21-P7). The best overall classification accuracy for percent of all animals correctly classified on the ROC curve, its corresponding sensitivity for percent of cases correctly classified, and its corresponding specificity for percent of controls correctly classified, were also reported for the best AUC case in each comparison.

Results

Sample Sets

All delivered pups (n = 174) were culled to 6 per litter. After random sampling within litters (depending on the scan capacity of the CT scanner), 123 pups were included in this MicroCT study. All 123 pups were alive at P7 and 121 pups were alive at P21. Our sample inclusion data are shown in Table 1. There were no significant differences on the treatment distribution before and after the quality check (Pearson chi square: $p = .18$ for P7 and $p = .14$ for P21). The treatment distribution of males and females for the total sample is shown in Table 2a and the distributions for the 3 sample sets used in our analyses are shown in Table 2(b-d). There were no significant differences for any sample set [Pearson chi square: Total (n = 123), $p = .981$; P7 (n = 75), $p = .341$; P21 (n = 78), $p = .649$; both P7/P21 (n = 56), $p = .250$].

Blood Alcohol Concentration

A committed set of treated dams (n = 8) was used to collect tail vein blood for blood alcohol concentration (BAC) analysis. Samples were collected during pre-pregnancy treatment on days 3 and 7 and during pregnancy treatment on days 3 and 10. All BACs were taken at 2 time periods, at 2 hours and at 4 hours into the dark cycle (12:00 noon and 2:00 PM). With 24-hour self-administration of liquid diets, it is difficult to identify with certainty the peak BACs of individual mice, but they typically occur during the first half of the dark cycle. Samples at 3 hours and 4 hours after lights-off reflect the most likely onset of ethanol liquid diet consumption during the first several hours of the dark cycle, but may not represent peak BAC values, which may occur at later time points in the dark cycle. Average BAC measures were 63 ± 6 mg/dL in the pre-pregnancy period and 21 ± 6 mg/dL in the pregnancy period. The complete details were previously reported and are available in Anthony et al., 2010.

Craniofacial Measures and Correlation Analysis

A sample skull segmented from a MicroCT scan is shown in Figure 1, where manually labeled landmarks are also displayed. The head size was estimated as the cubic centroid size of 17 landmarks, as previously described. An example segmentation result for the frontal, parietal, occipital, mandible, and nasal bones is shown in Figure 2. Head circumference, measured as the outer perimeter of the skull enclosure at the level of the nasion, is also shown in Figure 2. For a visual comparison of the segmentation results, one ALC subject and one CHOW subject at both P7 and P21 are shown in Figure 3.

Volume was calculated for the entire skull (Figure 1) as well as for each bone (Figure 2) based on the corresponding segmentation result. In total, 8 craniofacial measures were collected in our study: head size, head circumference, and 6 bone volumes including the total skull, frontal, parietal, occipital, mandible, and nasal volumes.

Supplemental Table 1 displays the pairwise comparison results of the Pearson correlation analysis among all of these measures. The correlation coefficients ranged from r = 0.483 (between total skull and head circumference) to r = 0.905 (between parietal and occipital bones) at P7, and ranged from r = 0.406 (between parietal and nasal bones) to r = 0.862 (between total skull and head size) at P21. All the correlations were significant with $p < .0001$.

Gender Effect

The craniofacial measures for males and females at both P7 and P21 are shown in Figure 4. No significant gender effect was detected on any of these measures based on the Bonferroni-corrected α level ($\alpha = 0.0071$ at P7, $\alpha = 0.003$ at P21). Two of the measures at P21, occipital bone ($p = .036$) and nasal bone ($p = .0386$), demonstrated an effect (female < male) meeting the nominal significance level of $p < .05$ without correcting for multiple comparisons.

Treatment Effect

Without including any covariate, treatment effects on craniofacial measures at both P7 and P21 are shown in Figure 5. Bonferroni-corrected α levels were $\alpha = 0.0071$ at P7 and $\alpha = 0.0063$ at P21. At P7, significant overall treatment effects were observed on skull bone ($p = .0014$), parietal bone ($p = .0025$), head size ($p = .0002$), and head circumference ($p = .0002$). Further pairwise group comparisons (see Table 3a) showed significant differences on skull (ALC < PF with $p = .0025$, CHOW < PF with $p = .0187$), parietal (ALC < CHOW with $p = .0023$), head size (ALC < CHOW with $p = .0001$, ALC < PF with $p = .0255$), and head circumference (ALC < CHOW with $p = .0002$, PF < CHOW with $p = .0191$). At P21, significant overall treatment effects were observed on skull, frontal, parietal, and occipital bones (all $p < .0001$), mandible ($p = .003$), head size ($p < .0001$), and head circumference ($p < .0001$). Further pairwise group comparisons (see Table 3a) showed significant differences on skull (ALC < CHOW with $p < .0001$, ALC < PF with $p < .0001$), frontal (ALC < CHOW with $p < .0001$, ALC < PF with $p = .0011$), parietal (ALC < CHOW with $p < .0001$, ALC < PF with $p < .0001$), occipital (ALC < CHOW with $p < .0001$, ALC < PF with $p < .0001$), mandible (ALC < CHOW with $p = .0043$, ALC < PF with $p = .0094$), head size (ALC < CHOW with $p < .0001$, ALC < PF with $p < .0001$), and head circumference (ALC < CHOW with $p < .0001$, ALC < PF with $p = .001$).

After covarying for gender and head size, treatment effects on craniofacial bone measures are shown in Figure 6. Bonferroni-corrected α levels were $\alpha = 0.01$ at P7 and $\alpha = 0.0083$ at P21. At P7, overall treatment effects were significant for skull ($p < .0001$) and mandible ($p = .0009$). Further pairwise group comparison (see Table 3b) showed significant differences

for skull (ALC > CHOW with $p = .0314$, PF > CHOW with $p < .0001$), and mandible (ALC > CHOW with $p = .0061$, PF > CHOW with $p = .0023$). At P21, overall treatment effects were significant on skull ($p < .0001$), parietal ($p < .0001$) and occipital ($p = .0005$). Further pairwise group comparisons (see Table 3b) showed significant differences on skull (ALC < CHOW with $p < .0001$, ALC < PF with $p = .0001$), parietal (ALC < CHOW with $p = .0003$, ALC < PF with $p = .0001$), and occipital (ALC < CHOW with $p = .0005$).

After covarying for gender and head circumference, treatment effects on craniofacial bone measures are shown in Figure 7. Bonferroni-corrected α levels were $\alpha = 0.01$ at P7 and $\alpha = 0.0083$ at P21. At P7, overall treatment effects were significant on skull ($p < .0001$) and mandible ($p = .0016$). Further pairwise group comparison (see Table 3b) showed significant differences on skull (ALC < PF with $p = .0198$, PF > CHOW with $p < .0001$), and mandible (ALC > CHOW with $p = .0307$, PF > CHOW with $p = .0015$). At P21, overall treatment effects were significant on skull, parietal, and occipital (all $p <$group comparisons (see Table 3b) showed significant differences on total skull (ALC < CHOW with $p < .0001$, ALC < PF with $p < .0001$), parietal (ALC < CHOW with $p < .0001$, ALC < PF with $p < .0001$), and occipital (ALC < CHOW with $p = .0001$, ALC < PF with $p = .0049$).

After covarying for the corresponding P7 measure (to focus on the growth occurring between P7 and P21), treatment effect on each of the 5 P21 bone measures is shown in Figure 8. Bonferroni-corrected α level was $\alpha = 0.01$. Overall treatment effects were significant on total skull ($p < .0001$), frontal ($p = .0002$), parietal ($p = .0048$), and occipital ($p = .0017$). Further pairwise group comparisons (see Table 3c) showed significant differences on total skull (ALC < PF < CHOW with $p \leq .0002$), frontal (ALC < CHOW with $p = .0001$, ALC < PF with $p = .01$), parietal (ALC < CHOW with $p = .0097$, ALC < PF with $p = .004$), and occipital (ALC < CHOW with $p = .0018$).

Discriminant Analyses

Area under ROC curve (AUC) results from the univariate ROC analyses are shown in Table 4a. The change of the skull bone volume between P7 and P21 was the best predictor to distinguish ALC and CHOW with AUC of 0.914 and best overall accuracy of 88.9% (sensitivity 55.6%, specificity 100%). The skull bone at P21 was the best predictor to classify ALC and PF with AUC of 0.922 and best overall accuracy of 85.7% (sensitivity 64.7%, specificity 100%). The skull bone volume at P7 was the best predictor to distinguish the PF and CHOW groups with AUC of 0.67 and best overall accuracy of 72.1% (sensitivity 86.2%, specificity 59.4%, PF as case, CHOW as control).

AUC results from multivariate SVM analyses are shown in Table 4b. The best model for classifying ALC and CHOW was a sigmoid-kernel SVM using the volume changes (P21-P7) of the occipital and skull bones as predictors: It achieved an AUC of 0.967 and the best overall accuracy on the ROC curve was 94.4% (sensitivity 77.8%, specificity 100%). The best model for classifying ALC and PF was a sigmoid-kernel SVM using the parietal, mandible, and skull volumes at P21 as predictors: It achieved an AUC of 0.995 and the best overall accuracy on the curve was 97.6% (sensitivity 94.1%, specificity 100%). The best model for classifying PF and CHOW was a linear-kernel SVM using the parietal and skull volumes at P7: It achieved an AUC of 0.805 and the best overall accuracy on the curve was 77% (sensitivity 69%, specificity 84.4%, PF as case, CHOW as control).

Discussion

The alcohol treatment, relative to the chow controls (ALC vs. CHOW) as well as to the isocaloric controls (ALC vs. PF), affected the head size, head circumference, and all the tested bones to variable degrees except the nasal bone. There was a growth trend of ALC <

CHOW and ALC < PF (Figure 5, Table 3a). Our discriminant analysis distinguished between ALC and CHOW (AUC = 0.967), between ALC and PF (AUC = 0.995), and between PF and CHOW (AUC = 0.805). The distinction among the 3 treatments was also found in body weight at embryonic (E) day 17 (Anthony et al., 2010), and facial dimension (craniometry) measurements at E17 (Anthony et al., 2010) and at P7 and P21 (Liang et al., 2011) under similar treatment conditions. The embryo body weights and craniometry in C57BL/6J mice were altered in ALC as compared to PF and CHOW. The alterations in craniometry were also detectable at P7 and P21 (Liang et al., 2011).

The head size (measured by the cubic centroid size of landmarks distributed through the entire craniofacial bone shown in Figure 1) and the head circumference (Figure 2) were both consistent distinguishing markers between ALC and CHOW, and between ALC and PF. This observation held at both P7 and P21 (Figure 5) and the retardation of the head size growth was exacerbated from P7 to P21. These findings suggest that the bony head size and head circumference were affected by alcohol throughout all ages disregarding dietary difference. The head size in this study was determined by 17 manually defined craniofacial landmarks and would be technically demanding in practice if it were to be used for human analysis. However, head circumference potentially could serve as a practical alternative measure for evaluating the alcohol effects in the mouse. In addition, because a smaller circumference was an important criterion for diagnosis for FAS and Partial FAS in humans (Hoyme et al., 2005), head circumference could be an important parameter to analyze the effects of dose and timing of alcohol exposure, as well as suboptimal nutritional conditions contributing to variability of FASD. A further examination of the head circumference in relation to alcohol and nutrition effects has been done in a parallel study (Liang et al., 2011).

The head size and circumference were significantly affected by prenatal alcohol exposure, and all bone measures were highly correlated with both the head size and circumference (Supplemental Table 1). Without controlling for either head size or circumference, skull and parietal bones differed significantly at P7, and skull, frontal, parietal, occipital, and mandible bones differed significantly at P21 (Figure 5). In addition, after adjusting each P21 measure by removing the corresponding P7 variance, each adjusted P21 volume (i.e., measure of growth occurring between P7 and P21) of skull, frontal, parietal, and occipital bones differed significantly among treatment groups (Figure 8, Table 3c). This demonstrated that, following the global growth pattern of head size or circumference, the retardation of individual bone growth was similarly exacerbated from P7 to P21.

With the above observations, any tested P21 craniofacial measure except nasal could sufficiently serve as a differential marker to distinguish ALC from CHOW or PF. However, to have a better mechanistic understanding, one might also be interested in learning whether there were additional effects of prenatal alcohol exposure on individual bones after controlling for the global effects on head size or circumference as well as the gender effect. There was a consistent trend regarding gender difference on craniofacial measures (Figure 4): male < female at P7, and male > female at P21 (except mandible). Although none of these differences met the Bonferroni-corrected significance level, the gender effects on occipital and nasal bones at P21 were nominally significant (uncorrected $p < .05$). Thus, gender was still included as a covariate in our ANCOVA analyses.

Two sets of ANCOVA analyses on bone volumes were performed: one with head size and gender as covariates, one with head circumference and gender as covariates. The results were extremely similar (Figures 6 and 7, Table 3b). At P7, skull and mandible were affected by dietary disparity, where CHOW was less developed than ALC and PF. At P21, skull, parietal, and occipital bones were affected by alcohol relative to both control groups (ALC < PF, ALC < CHOW), and the 2 control groups did not differ. This suggests that P21 may be a

better time to detect the alcohol effect, since the nutritional effect between PF and CHOW is not pronounced. Further investigation is warranted to determine whether the continued alcohol effects on bone between P7 and P21 are solely due to alcohol, or might also be partly due to nutrition.

Significantly, the neurocranial bones (upper skull: occipital, parietal, and frontal) were affected by the alcohol treatment to a greater degree than the viscerocranial bones (facial skeleton), which were either differentially affected (mandible), or not affected (nasal). In particular, while covarying for the head size, the adjusted mandible volume was increased in the ALC group, relative to CHOW, at an early stage (P7) and no difference was observed for the 3 neurocranial bones. This indicated that, at P7, the alcohol treatment affected the head growth more significantly than the mandible growth, and the retardation of the neurocranial bone growth was at a similar level to that of the head growth. On the other hand, while covarying for the P7 measure, the treatment affected only the 3 neurocranial bones at P21 but not the mandible.

Thus the retardation of the growth was exacerbated from P7 to P21 on the neurocranial bones but not the mandible. The mandible, unlike the rest of the face, is not derived directly from neural crest cells (Huang and Saint-Jeannet, 2004). It is unique in that it is the first cranial bone to form and originates from the first branchial arch derivatives through intramembranous condensation around a cartilaginous (Meckle's cartilage) model (Moore et al., 2011). A general trend is observed that the alcohol effect decreases in order from posterior bones (occipital, parietal, frontal) to anterior bones (nasal, mandible). Given this, further investigation is warranted to examine the diagnostic potential of neurocranial dysmorphology in addition to that of viscerocranial (facial) dysmorphology.

By visualizing the segmented bones, we observed deformity and missing pieces in the ALC group compared to CHOW (Figure 3), and compared to PF (not shown). Detailed morphometric analysis may identify structural changes beyond simple volumetric analysis and thus warrants further investigation.

The above analyses indicated that craniofacial bones had different discriminative powers for detecting the effects of alcohol relative to either the normal CHOW group or the isocaloric control PF group. Our univariate discriminant analyses demonstrated that the skull was the best predictor to differentiate the treatment groups: 1) skull bone change (P21-P7) classified ALC and CHOW with AUC of 0.914 and accuracy of 88.9% (sensitivity 55.6%, specificity 100%), 2) skull at P21 differentiated ALC and PF with AUC of 0.922 and accuracy of 85.7% (sensitivity 64.7%, specificity 100%), and 3) skull at P7 classified PF and CHOW with AUC of 0.67 and accuracy of 72.1% (sensitivity 86.2%, specificity 59.4%). Our multivariate discriminant analyses, by involving more than one predictor, demonstrated further improved performance: 1) volume changes of occipital bone and skull (P21-P7) classified ALC and CHOW with AUC of 0.967 and accuracy of 94.4% (sensitivity 77.8%, specificity 100%), 2) parietal bone, mandible, and skull at P21 classified ALC and PF with AUC of 0.995 and accuracy of 97.6% (sensitivity 94.1%, specificity 100%), and 3) parietal bone and skull volumes at P7 classified PF and CHOW with AUC of 0.805 and accuracy of 77% (sensitivity 69%, specificity 84.4%).

In the process of alcohol administration, a liquid diet with PMI® was used in the ALC and PF groups, which was isocaloric in its design (in content and in intake volume). Due to the limited intake of the liquid diet by the ALC (and PF) groups, and/or due to differences in dietary content relative to chow, there was a nutritional disparity between the liquid diet groups (ALC and PF) and the CHOW group. Based on comparisons of amount of diet consumed across the groups, the PMI® diet had low amounts of micronutrients and protein

compared to the chow diet (Teklad diet 2018SX Global 18% Protein Rodent Diet, Harlan, Indianapolis, Indiana). Despite this nutritional disparity, the PF group given liquid diet did not differ significantly from the CHOW group for most of the craniofacial measures except skull and mandible at P7 and change in skull growth between P7 and P21. Using both skull and parietal bone at P7, discriminant analysis could classify CHOW and PF with cross validation accuracy of 77% (sensitivity 69%, specificity 84.4%). On the other hand, the P21 measures indicated that the alcohol effect carried through later life while the initial PF effect diminished (i.e., no significant difference between PF and CHOW at P21, see Table 3a and 3b). This lasting effect would be a useful diagnosis for FASD, while the stage-dependent transient effect would be used to identify other environmental insults (e.g., nutrition disparity). The lack of information about body size and weight of pups is one limitation of this work. Future experimental replication and validation studies should take into account this information and remove potential confounding factors.

Furthermore, facial measurements analogous to anthropometry demonstrated a synergistic effect of alcohol and dietary disparity (Liang et al., 2011). A related study on craniofacial bone growth is ongoing, focusing on examining the relationship between restricted caloric intake and corresponding calcium consumption. Since our observation with MicroCT indicated that the alcohol-exposed pups had a less mineralized neurocranium at P21, we suspect that alcohol exposure might reduce calcium consumption, deposit or absorption. For example, a relevant prenatal ethanol study on rats demonstrated that the alcohol-induced delay in development differed for weight and skeletal ossification (Simpson et al., 2005), and sensitivity to alcohol appeared to be greatest in bones that were more mineralized and underwent a greater proportion of their development *in utero*.

In sum, these results validated our hypothesis that prenatal alcohol exposure significantly reduced craniofacial bone development after birth. A suboptimal nutrition condition associated with the liquid diet procedure might have magnified the effect of prenatal alcohol exposure on some bones (Figure 5). Given that all the tested bones were highly correlated to the head size and circumference, the retardation of the craniofacial bone growth was coupled with that of the head size/circumference growth and was exacerbated from P7 to P21. In addition, we found that the neurocranium was more sensitive to alcohol than the viscerocranium. It warrants further investigation to delineate effects of quantity and frequency of prenatal alcohol exposure on craniofacial bones for diagnosis of FASD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by NIH grant UO1 AA14819 to F.C. Zhou. This UO1 is a part of the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD) Consortium. L. Shen, Y. Liang, C. Goodlett, and R. Ward are co-investigators. H. Ai is a PhD candidate who assisted with creating MicroCT scans and craniofacial bone segmentation and landmark labeling. X. Ren is a PhD candidate who assisted with the statistical analyses, and C.B. Anthony is a collaborator who assisted with the animal drinking paradigm. All or part of this work was done in conjunction with the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD), which is funded by grants from the National Institute on Alcohol and Alcohol Abuse (NIAAA). E. Riley is the principal investigator of the Administrative Core of the CIFASD.

References

Anthony B, Vinci-Booher S, Wetherill L, Ward R, Goodlett C, Zhou FC. Alcohol-induced facial dysmorphology in C57BL/6 mouse models of fetal alcohol spectrum disorder. Alcohol. 2010; 44:659–671. [PubMed: 20570474]

Astley SJ, Clarren SK. Measuring the facial phenotype of individuals with prenatal alcohol exposure: correlations with brain dysfunction. Alcohol Alcohol. 2001; 36:147–159. [PubMed: 11259212]

Bookstein, FL. Morphometric tools for landmark data: geometry and biology. New York: Cambridge University Press; 1991.

Chen SY, Sulik KK. Free radicals and ethanol-induced cytotoxicity in neural crest cells. Alcohol Clin Exp Res. 1996; 20:1071–1076. [PubMed: 8892529]

Chen Y, Ozturk NC, Ni L, Goodlett C, Zhou FC. Strain differences in developmental vulnerability to alcohol exposure via embryo culture in mice. Alcohol Clin Exp Res. 2011; 35:1293–1304. [PubMed: 21410487]

Cortes C, Vapnik V. Support-Vector Networks. Mach Learn. 1995; 20:273–297.

Debelak KA, Smith SM. Avian genetic background modulates the neural crest apoptosis induced by ethanol exposure. Alcohol Clin Exp Res. 2000; 24:307–314. [PubMed: 10776667]

Goodlett CR, Horn KH, Zhou FC. Alcohol teratogenesis: mechanisms of damage and strategies for intervention. Exp Biol Med (Maywood). 2005; 230:394–406. [PubMed: 15956769]

Hassler JA, Moran DJ. The effects of ethanol on embryonic actin: a possible role in teratogenesis. Experientia. 1986; 42:575–577. [PubMed: 3519273]

Hernandez-Guerrero JC, Ledesma-Montes C, Loyola-Rodriguez JP. Effects of maternal ethanol intake on second alcoholic generation murine skull and mandibular size. Arch Med Res. 1998; 29:297–302. [PubMed: 9887546]

Hoyme HE, May PA, Kalberg WO, Kodituwakku P, Gossage JP, Trujillo PM, Buckley DG, Miller JH, Aragon AS, Khaole N, et al. A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 institute of medicine criteria. Pediatrics. 2005; 115:39–47. [PubMed: 15629980]

Huang X, Saint-Jeannet JP. Induction of the neural crest and the opportunities of life on the edge. Dev Biol. 2004; 275:1–11. [PubMed: 15464568]

Kaminen-Ahola N, Ahola A, Maga M, Mallitt KA, Fahey P, Cox TC, Whitelaw E, Chong S. Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. PLoS Genet. 2010; 6:e1000811. [PubMed: 20084100]

Liang Y, Ai H, Anthony B, Wetherill L, Ward R, Zhou FC. CIFASD. Postnatal Craniofacial Bone Dysmorphology and Growth Pattern Affected by Prenatal Alcohol Exposure and Nutrition. Alcohol Clin Exp Res. 2011; 35(172A Supplement)

May PA, Gossage JP. Estimating the prevalence of fetal alcohol syndrome. A summary. Alcohol Res Health. 2001; 25:159–167. [PubMed: 11810953]

May PA, Gossage JP, Kalberg WO, Robinson LK, Buckley D, Manning M, Hoyme HE. Prevalence and epidemiologic characteristics of FASD from various research methods with an emphasis on recent in-school studies. Dev Disabil Res Rev. 2009; 15:176–192. [PubMed: 19731384]

Metscher B. MicroCT for comparative morphology: simple staining methods allow high-contrast 3D imaging of diverse non-mineralized animal tissues. BMC Physiol. 2009; 9:11. [PubMed: 19545439]

Moore, KL.; Persaud, TVN.; Torchia, MG. The Developing Human: Clinically Oriented Embryology. Elsevier/Saunders; 2011.

Neues F, Epple M. X-ray microcomputer tomography for the study of biomineralized endo- and exoskeletons of animals. Chem Rev. 2008; 108:4734–4741. [PubMed: 18754688]

Parsons TE, Kristensen E, Hornung L, Diewert VM, Boyd SK, German RZ, Hallgrímsson B. Phenotypic variability and craniofacial dysmorphology: increased shape variance in a mouse model for cleft lip. J Anat. 2008; 212:135–143. [PubMed: 18093101]

Robin NH, Zackai EH. Unusual craniofacial dysmorphia due to prenatal alcohol and cocaine exposure. Teratology. 1994; 50:160–164. [PubMed: 7801303]

Rosett HL. A clinical perspective of the Fetal Alcohol Syndrome. Alcohol Clin Exp Res. 1980; 4:119–122. [PubMed: 6990814]

Rovasio RA, Battiato NL. Role of early migratory neural crest cells in developmental anomalies induced by ethanol. Int J Dev Biol. 1995; 39:421–422. [PubMed: 7669554]

Rovasio RA, Battiato NL. Ethanol induces morphological and dynamic changes on in vivo and in vitro neural crest cells. Alcohol Clin Exp Res. 2002; 26:1286–1298. [PubMed: 12198407]

Sampson PD, Streissguth AP, Bookstein FL, Little RE, Clarren SK, Dehaene P, Hanson JW, Graham JM Jr. Incidence of fetal alcohol syndrome and prevalence of alcohol-related neurodevelopmental disorder. Teratology. 1997; 56:317–326. [PubMed: 9451756]

Schmidt E, Parsons T, Jamniczky H, Gitelman J, Trpkov C, Boughner J, Logan CC, Sensen C, Hallgrimsson B. Micro-computed tomography-based phenotypic approaches in embryology: procedural artifacts on assessments of embryonic craniofacial growth and development. BMC Dev Biol. 2010; 10:18. [PubMed: 20163731]

Sensen C, Hallgrimsson B. Micro-computed tomography-based phenotypic approaches in embryology: procedural artifacts on assessments of embryonic craniofacial growth and development. BMC Dev Biol. 2010; 10:18. [PubMed: 20163731]

Simpson ME, Duggal S, Keiver K. Prenatal ethanol exposure has differential effects on fetal growth and skeletal ossification. Bone. 2005; 36:521–532. [PubMed: 15777686]

Sulik KK. Craniofacial defects from genetic and teratogen-induced deficiencies in presomite embryos. Birth Defects Orig Artic Ser. 1984; 20:79–98. [PubMed: 6542431]

Sulik KK, Cook CS, Webster WS. Teratogens and craniofacial malformations: relationships to cell death. Development. 1988; 103(Suppl):213–231. [PubMed: 3074910]

Sulik KK, Johnston MC. Sequence of developmental alterations following acute ethanol exposure in mice: craniofacial features of the fetal alcohol syndrome. Am J Anat. 1983; 166:257–269. [PubMed: 6846205]

Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. BMC Med Res Methodol. 2001; 1:6. [PubMed: 11459516]

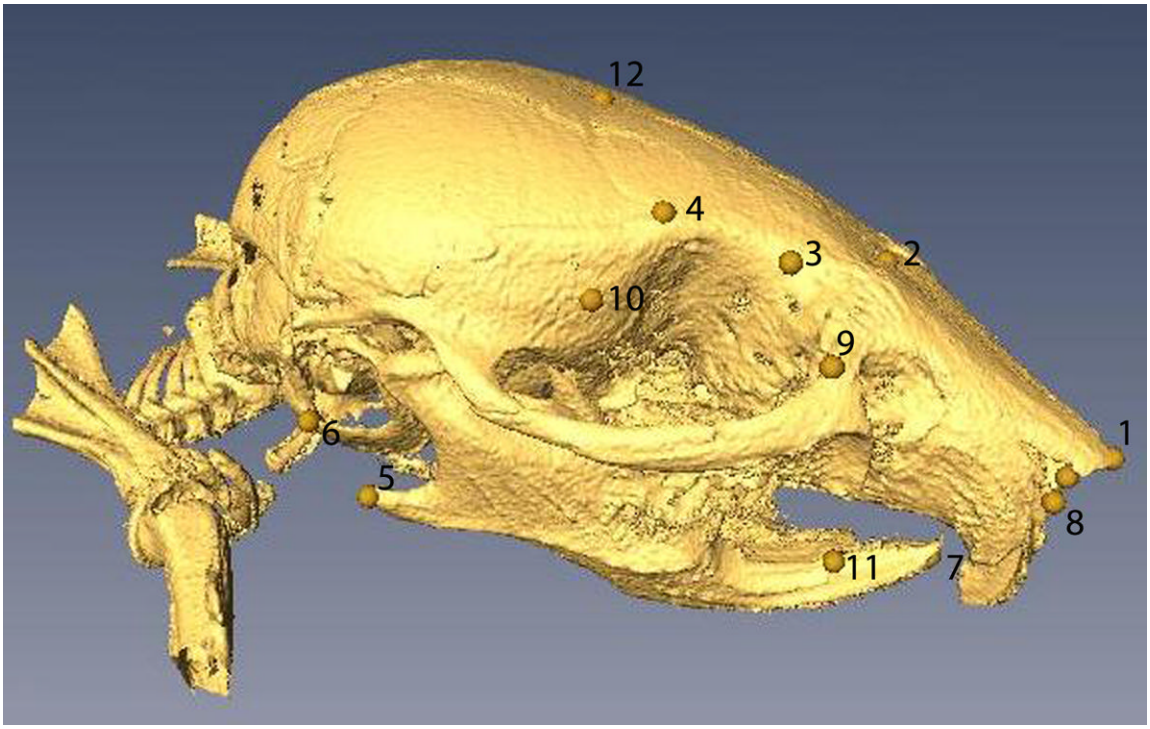

Figure 1.
Landmarks manually selected on the craniofacial bone surfaces: 1) forward center tip of nasal bone, 2) nasion, 3) left and right endpoints for defining inner canthal width, 4) left and right endpoints for defining minimal frontal width, 5) left and right endpoints for defining bigonial width, 6) left and right endpoints for defining bitragal width, 7) forward tip of mandible, 8) left and right endpoints for defining nasal length, 9) left and right interior points of orbital socket, 10) left and right exterior points of orbital socket, 11) left and right endpoints for bony mandible, and 12) bregma. Note that landmarks 3, 4, 5, 6, 8, 9, 10 and 11 are bilateral and only the ones on the right side of the skull are visible in the picture. There are 20 landmarks in total.

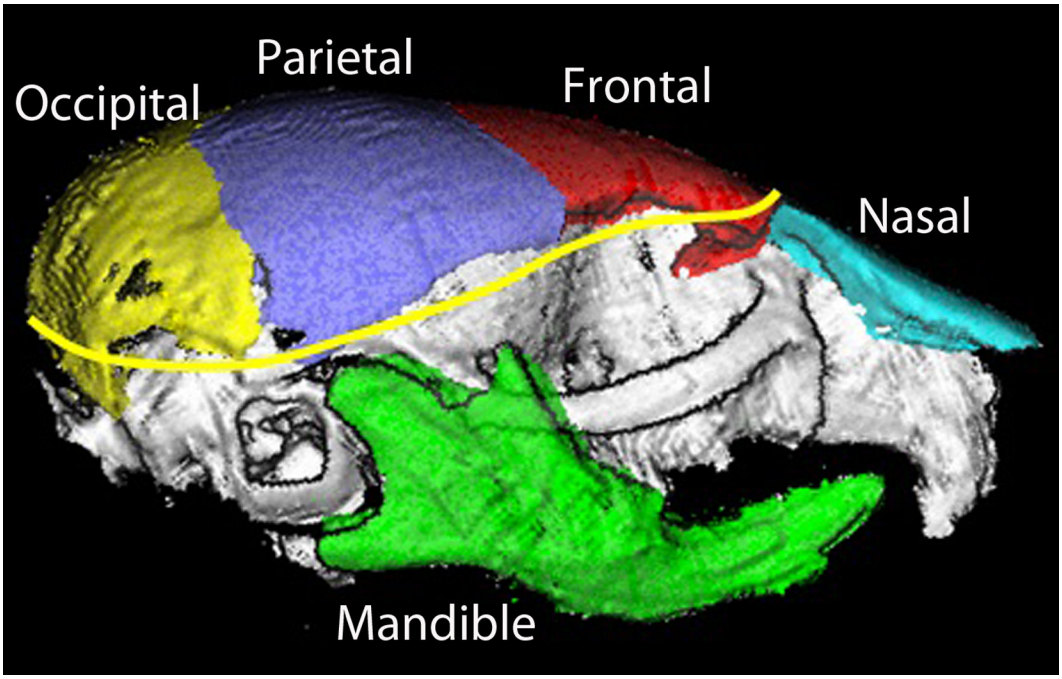

Figure 2.
Sample segmentation results of frontal, parietal, occipital, mandible, and nasal bones are shown on the whole skull. Head circumference, measured as the outer perimeter of the skull enclosure at the level of nasion, is shown as the yellow line.

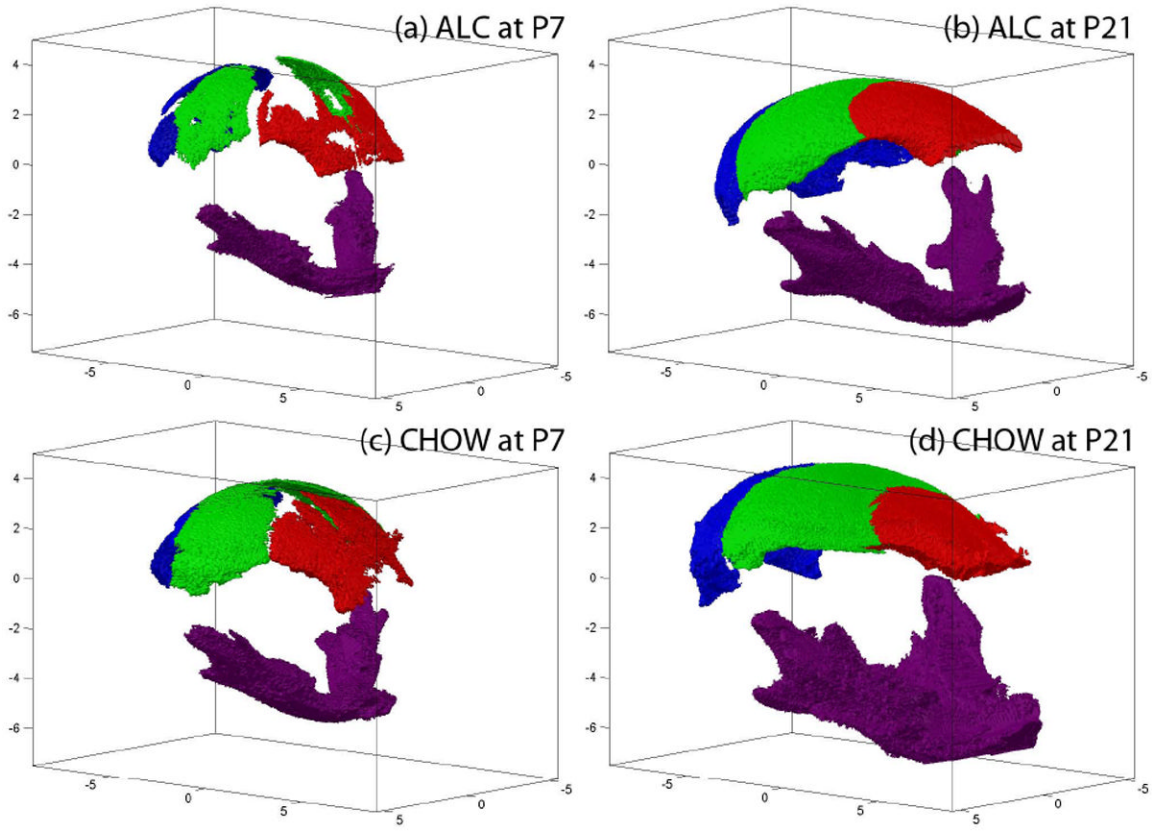

Figure 3.
Segmented frontal (red), parietal (green), occipital (blue), and mandible (purple) bones of example ALC and CHOW subjects are shown at both P7 and P21 stages.

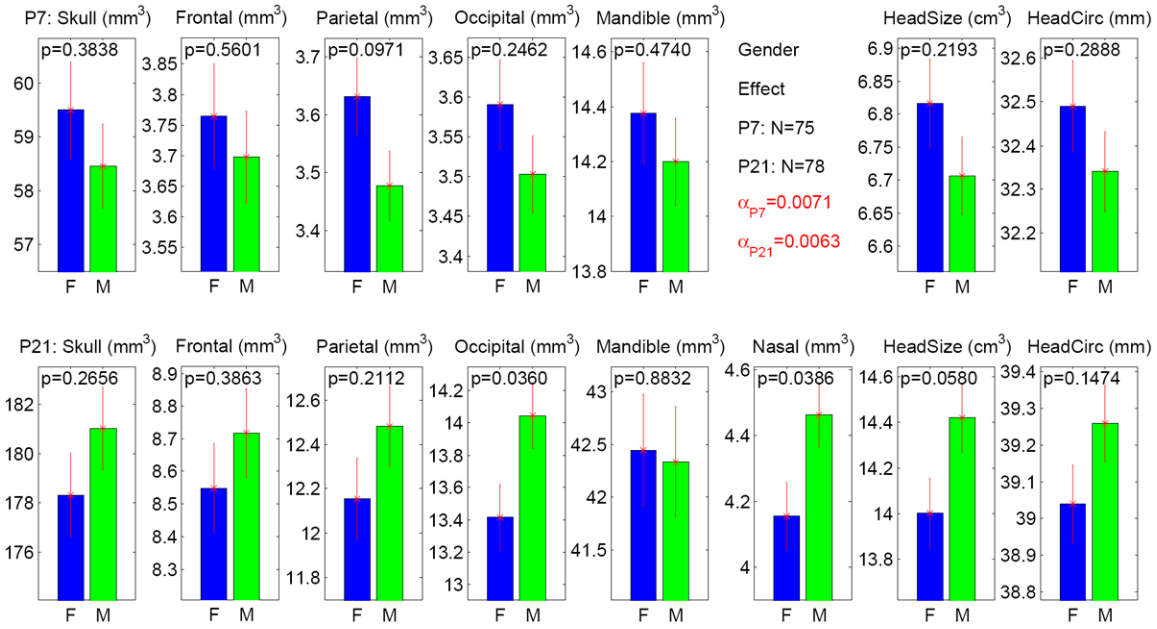

Figure 4.

Gender effect on 8 measures at both P7 and P21 stages. The y axis indicates the adjusted measure after accounting for litter effect as a nested factor. The α level is adjusted to 0.0071 (= 0.05/7) for P7 results (7 tests in total), and adjusted to 0.0063 (= 0.05/8) for P21 results (8 tests in total). None of the *p* values meets the adjusted α level. The standard error is shown as the error bar.

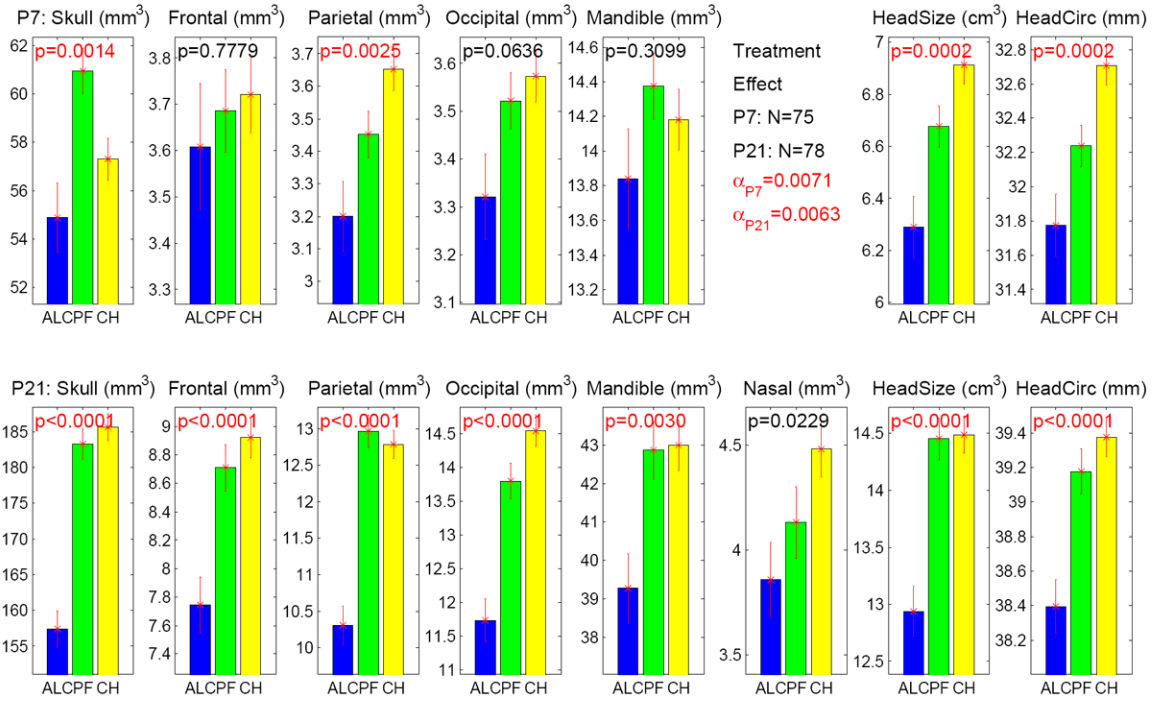

Figure 5.

Treatment effect on 8 measures at both P7 and P21 stages. The y axis indicates the adjusted measure after accounting for litter effect as a nested factor. The α level is adjusted to 0.0071 (= 0.05/7) for P7 results (7 tests in total), and adjusted to 0.0063 (= 0.05/8) for P21 results (8 tests in total). Significant *p* values, which meet the adjusted α, are colored in red. The standard error is shown as the error bar.

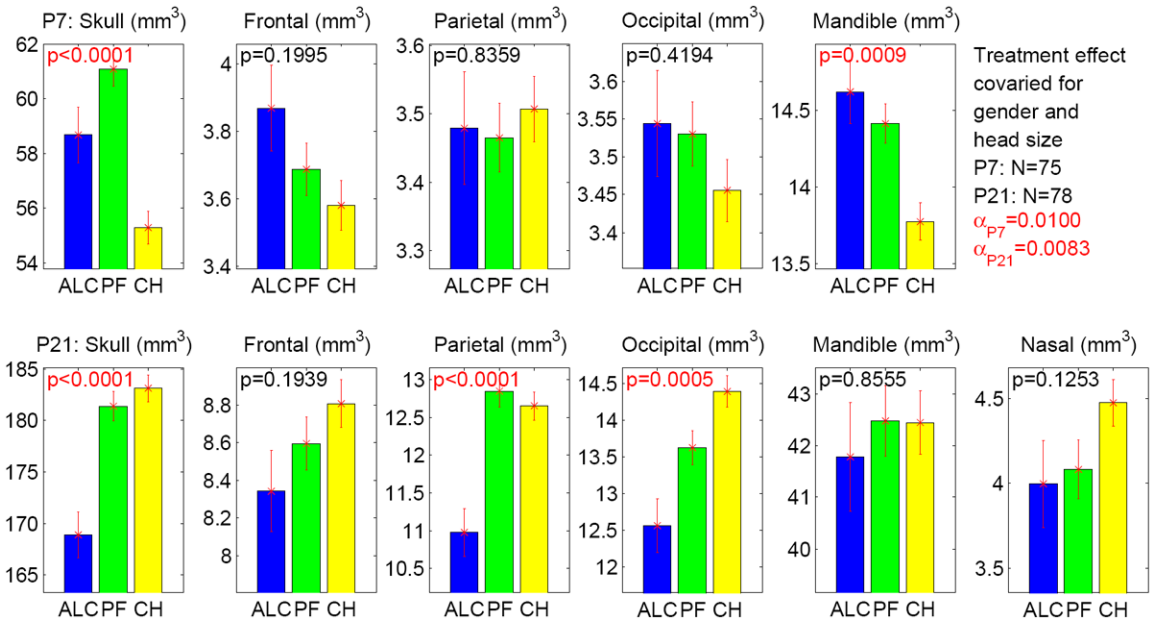

Figure 6.

Treatment effect on 6 bone volumes at P7 and P21 stages while controlling for gender and head size. The y axis indicates the adjusted measure after accounting for litter effect as a nested factor and regressing out the effects of gender and head size. The α level is adjusted to 0.01 (= 0.05/5) for P7 results (5 tests in total), and adjusted to 0.0083 (= 0.05/6) for P21 results (6 tests in total). Significant *p* values, which meet the adjusted α, are colored in red. The standard error is shown as the error bar.

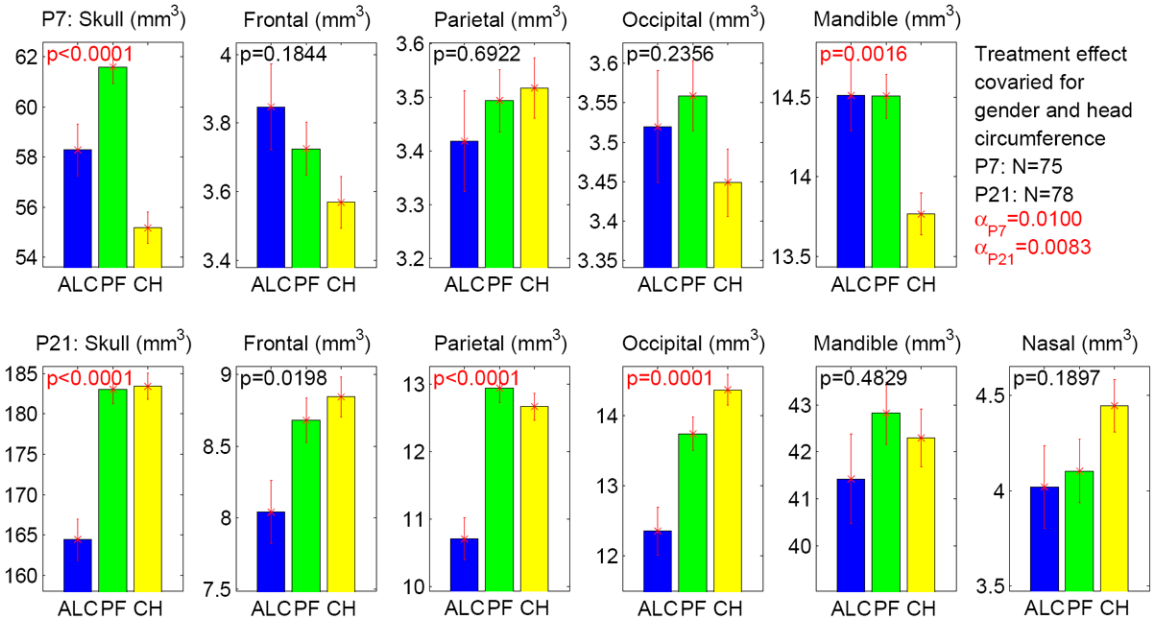

Figure 7.

Treatment effect on 6 bone volumes at P7 and P21 stages while controlling for gender and head circumference. The y axis indicates the adjusted measure after accounting for litter effect as a nested factor and regressing out the effects of gender and head circumference. The α level is adjusted to 0.01 (= 0.05/5) for P7 results (5 tests in total), and adjusted to 0.0083 (= 0.05/6) for P21 results (6 tests in total). Significant *p* values, which meet the adjusted α, are colored in red. The standard error is shown as the error bar.

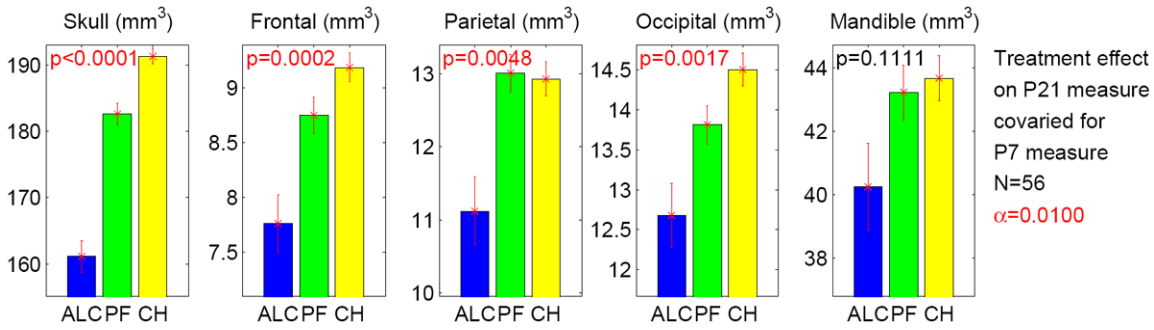

Figure 8.

Treatment effect on each of 5 P21 measures covaried for the corresponding P7 measure. The y axis indicates the adjusted P21 measure after regressing out the effect of the P7 measure. The α level is adjusted to 0.01 (= 0.05/5), given that there are 5 tests in total. Significant *p* values, which meet the adjusted α, are colored in red. The standard error is shown as the error bar.

Table 1

MicroCT scans were successfully acquired for all the samples at P7 (N=123), and for most samples at P21 (N=115, excluding 2 samples due to animal death and 6 samples due to scanner failure). Quality check (QC) was performed to remove samples with severe motion artifacts. This table shows the QC result by group crosstabs for P7 and P21 sample sets. Chi-square test indicated no significant relationship for QC results between treatment groups.

Pass quality check	P7 (N=123)				P21 (N=115)			
	ALC	PF	CHOW	Total	ALC	PF	CHOW	Total
No	16	15	17	48	6	19	12	37
Yes	14	29	32	75	17	25	36	78
Total	30	44	49	123	23	44	48	115
Chi-square (p)	0.18				0.14			

Table 2

Group-by-sex crosstabs for 4 sample sets: a) All the samples (N = 123), b) samples whose P7 measures were successfully extracted (N = 75), c) samples whose P21 measures were successfully extracted (N = 78), and d) samples whose P7 and P21 measures were both successfully extracted (N = 56). Pearson chi-square significance is shown for each sample set.

(a) Total sample				**(b) Sample with P7 measures**			
	Female	**Male**	**Total**		**Female**	**Male**	**Total**
ALC	15	15	30	ALC	6	8	14
PF	23	21	44	PF	10	19	29
CHOW	25	24	49	CHOW	17	15	32
Total	63	60	123	Total	33	42	75
Chi-square significance (*p*)			.981	Chi-square significance (*p*)			.341
(c) Sample with P21 measures				(d) Sample with P7 and P21 measures			
	Female	Male	Total		Female	Male	Total
ALC	8	9	17	ALC	4	5	9
PF	11	14	25	PF	7	13	20
CHOW	20	16	36	CHOW	16	11	27
Total	39	39	78	Total	27	29	56
Chi-square significance (*p*)			.649	Chi-square significance (*p*)			.250

Table 3

Significance (corrected p values by the Sidak method) of pairwise group difference: (ab) Treatment effects on P7 volumes (N=75) and P21 volumes (N=78) without (a) and with (b) covariates. (c) Treatment effect on the P21 measure covaried for the P7 measure. Bones without any significant pairwise group difference (p>0.05) are not shown in the table. Significant results (p<0.05) are highlighted in bold.

(a) Pairwise treatment effect on P7 and P21 measures (no covariates)				
Exp	**Bone**	**No Covariates**		
		ALC vs CHOW	**ALC vs PF**	**PF vs CHOW**
P7 (N=75)	Skull	0.3922	**0.0025**	**0.0187**
	Parietal	**0.0023**	0.1606	0.1244
	HeadSize	**0.0001**	**0.0255**	0.0900
	HeadCirc	**0.0002**	0.1152	**0.0191**
P21 (N=78)	Skull	**0.0000**	**0.0000**	0.7823
	Frontal	**0.0000**	**0.0011**	0.7007
	Parietal	**0.0000**	**0.0000**	0.9082
	Occipital	**0.0000**	**0.0000**	0.1117
	Mandible	**0.0043**	**0.0094**	0.9989
	HeadSize	**0.0000**	**0.0000**	0.9987
	HeadCirc	**0.0000**	**0.0010**	0.5897

(b) Pairwise treatment effect on P7 and P21 measures (with covariates)							
Exp	Bone	Covaried for sex and headsize			Covaried for sex and headcirc		
		ALC vs CHOW	ALC vs PF	PF vs CHOW	ALC vs CHOW	ALC vs PF	PF vs CHOW
P7 (N=75)	Skull	**0.0314**	0.1234	**0.0000**	0.0633	**0.0198**	**0.0000**
	Mandible	**0.0061**	0.7702	**0.0023**	**0.0307**	1.0000	**0.0015**
P21 (N=78)	Skull	**0.0000**	**0.0001**	0.7501	**0.0000**	**0.0000**	0.9978
	Frontal	0.2382	0.7254	0.5906	**0.0177**	0.0613	0.8217
	Parietal	**0.0003**	**0.0001**	0.8618	**0.0000**	**0.0000**	0.7351
	Occipital	**0.0005**	0.0661	0.0531	**0.0001**	**0.0049**	0.1619

(c) Pairwise treatment effect on the P21 measure covaried for the P7 measure		
Exp	Bone	Covaried for the P7 measure

		ALC vs CHOW	ALC vs PF	PF vs CHOW
P21 (N=56)	Skull	**0.0000**	**0.0000**	**0.0002**
	Frontal	**0.0001**	**0.0100**	0.1355
	Parietal	**0.0097**	**0.0040**	0.9936
	Occipital	**0.0018**	0.0697	0.0891

Table 4

(a) Results of univariate ROC analysis: Areas under ROC curve (AUCs) are shown. (b) Best results of multivariate SVM analysis using either linear or sigmoid function as its kernel: AUCs estimated from leave-one-out cross-validation trials are shown. (a-b) Best AUC results in each comparison are highlighted in bold.

(a) Univariate ROC analysis: AUC results				
	Bone	**ALC vs CHOW**	**ALC vs PF**	**PF vs CHOW**
P7	**Skull**	0.558	0.666	**0.67**
	Frontal	0.522	0.5	0.512
	Parietal	0.54	0.512	0.56
	Occipital	0.549	0.571	0.502
	Mandible	0.507	0.505	0.503
P21	**Skull**	0.852	**0.922**	0.558
	Frontal	0.801	0.746	0.616
	Parietal	0.902	0.92	0.527
	Occipital	0.864	0.838	0.663
	Mandible	0.694	0.692	0.528
	Nasal	0.708	0.587	0.604
P21-P7	**Skull**	**0.914**	0.817	0.678
	Frontal	0.794	0.661	0.613
	Parietal	0.885	0.917	0.589
	Occipital	0.811	0.783	0.593
	Mandible	0.621	0.65	0.561

(b) Multivariate SVM analysis using linear or sigmoid function as kernel: Best AUC results							
	Kernel Function	ALC vs CHOW		ALC vs PF		PF vs CHOW	
		Bones*	AUC	Bones*	AUC	Bones*	AUC
P7	Linear	P,M	0.629	F,M,S	0.751	**P,S**	**0.805**
	Sigmoid	F,O	0.616	O,S	0.621	M,S	0.742
P21	Linear	P,M	0.889	P,M,S	0.934	F,P,O	0.688
	Sigmoid	M,O	0.912	**P,M,S**	**0.995**	P,O,S	0.732
P21-P7	Linear	F,M,S	0.918	P,M	0.917	P,M,S	0.759

Sigmoid	**O,S**	**0.967**	M,P	0.928	O,P	0.761

* Bones used in the SVM model: Selected from Frontal (F), Parietal (P), Occipital (O), Mandible (M), and Skull (S).