Published in final edited form as:
Dev Biol. 2011 July 15; 355(2): 175–182. doi:10.1016/j.ydbio.2011.02.033.

The FaceBase Consortium: A comprehensive program to facilitate craniofacial research

Harry Hochheiser[a,*], **Bruce J. Aronow**[b,c], **Kristin Artinger**[d], **Terri H. Beaty**[e], **James F. Brinkley**[f], **Yang Chai**[g], **David Clouthier**[d], **Michael L. Cunningham**[h], **Michael Dixon**[i,j], **Leah Rae Donahue**[k], **Scott E. Fraser**[l], **Benedikt Hallgrimsson**[m], **Junichi Iwata**[g], **Ophir Klein**[n], **Mary L. Marazita**[o,p,q,r], **Jeffrey C. Murray**[s], **Stephen Murray**[k], **Fernando Pardo-Manuel de Villena**[t], **John Postlethwait**[u], **Steven Potter**[v], **Linda Shapiro**[f], **Richard Spritz**[w], **Axel Visel**[x,y], **Seth M. Weinberg**[z], and **Paul A. Trainor**[aa,ab,**]

[a]Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh PA 15232, USA

[b]Division of Biomedical Informatics, Center for Computational Medicine, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, CHRF 8504, 3333 Burnet Ave Cincinnati, OH 45229, USA

[c]Division of Developmental Biology, Center for Computational Medicine, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, CHRF 8504, 3333 Burnet Ave Cincinnati, OH 45229, USA

[d]Department of Craniofacial Biology, University of Colorado Denver Anschutz Medical Campus, Aurora, CO 80045, USA

[e]Department of Epidemiology, Johns Hopkins University, 615 N. Wolfe Street, Baltimore, MD 21205, USA

[f]Department of Computer Science and Engineering, University of Washington, Box 352350 Seattle, WA 98195-2350, USA

[g]Center for Craniofacial Molecular Biology, Ostrow School of Dentistry, University of Southern California, 2250 Alcazar Street, CSA 103, Los Angeles, CA 90033, USA

[h]Seattle Children's Hospital, 4800 Sand Point Way NE, Seattle, WA 98105, USA

[i]Faculty of Medical and Human Sciences, Manchester Academic Health Sciences Center, University of Manchester, Oxford Road, Manchester, M13 9PT, England, UK

[j]Faculty of Life Sciences, Michael Smith Building, University of Manchester, Oxford Road, Manchester, M13 9PT, England, UK

[k]Jackson Laboratory, 600 Main St., Bar Harbor, ME 04609, USA

[l]Biological Imaging Center Beckman Institute 133, M/C 139-74 California Institute of Technology Pasadena, CA 91125, USA

[m]Dept. of Cell Biology and Anatomy and the Alberta Children's Hospital Research Institute, University of Calgary, 3330 Hospital Dr. NW, Calgary, Alberta T2N 4N1, Canada

[n]Department of Orofacial Sciences and Program in Craniofacial and Mesenchymal Biology, University of California San Francisco, San Francisco, USA

[*]Correspondence to: H. Hochheiser, Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA 15232, USA.
[**]Correspondence to: P.A. Trainor, Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, USA. pat@stowers.org (P.A. Trainor).

[o]Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA

[p]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15219, USA

[q]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15219, USA

[r]Clinical and Translational Science Institute, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA

[s]Department of Pediatrics, University of Iowa, Iowa City, IA 52242, USA

[t]Department of Genetics The University of North Carolina at Chapel Hill120 Mason Farm Road Room 5046 Genetic Medicine Building Campus Box 7264, Chapel Hill, NC 27599-7264, USA

[u]Institute of Neuroscience, University of Oregon, Eugene, OR 97403, USA

[v]Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229, USA

[w]Human Medical Genetics Program, University of Colorado School of Medicine, Aurora, CO 80045, USA

[x]Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[y]U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA

[z]Center for Craniofacial and Dental Genetics Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Bridgeside Point Building 100 Technology Drive, Suite 500 Pittsburgh, PA 15219, USA

[aa]Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, USA

[ab]Department of Anatomy and Cell Biology, University of Kansas Medical Center, Kansas City, KS 66160, USA

Abstract

The FaceBase Consortium consists of ten interlinked research and technology projects whose goal is to generate craniofacial research data and technology for use by the research community through a central data management and integrated bioinformatics hub. Funded by the National Institute of Dental and Craniofacial Research (NIDCR) and currently focused on studying the development of the middle region of the face, the Consortium will produce comprehensive datasets of global gene expression patterns, regulatory elements and sequencing; will generate anatomical and molecular atlases; will provide human normative facial data and other phenotypes; conduct follow up studies of a completed genome-wide association study; generate independent data on the genetics of craniofacial development, build repositories of animal models and of human samples and data for community access and analysis; and will develop software tools and animal models for analyzing and functionally testing and integrating these data. The FaceBase website (http://www.facebase.org) will serve as a web home for these efforts, providing interactive tools for exploring these datasets, together with discussion forums and other services to support and foster collaboration within the craniofacial research community.

Keywords

Craniofacial development; Cleft lip and palate; Human genetics; Animal models; Database; Morphometrics

Motivation to create the FaceBase Consortium

Among the lessons learned from the growth of “omics” research is the demonstration that collections of large publicly available, thoroughly annotated and integrated gene- and bioinformatics datasets can dramatically facilitate focused, hypothesis-driven research and open new areas of systems-level research. The availability of genome sequences in readily accessible user interfaces (e.g.; http://genome.ucsc.edu/; (Kent et al., 2002)), has fundamentally changed approaches to molecular biology and genetic research. Analyses that used to require months of benchwork in the pre-genomic era can now be performed *in silico* in a fraction of that time. Likewise, access to carefully annotated repositories of published phenotype, developmental and expression data (e.g.; http://www.informatics.jax.org/genes.shtml;(Bult et al., 2008); http://emouseatlas.org/emage/; (Richardson et al., 2010)) can enable systems-wide integrative studies to identify functional gene networks that underlie the etiology and pathogenesis of human disease. The NIDCR has sought to accelerate research in craniofacial developmental biology and further understanding of craniofacial and dental disorders by establishing the FaceBase Consortium.

Ten research and technology projects currently focused on the mid-face plus a data management and coordination hub were selected from a pool of peer-reviewed applications submitted in response to a Request for Applications (RFA) issued in 2008. Projects were selected based on merit as well as their potential for synergism with other projects and the integrative goals of the central hub. For example, studies of craniofacial development in model systems may identify candidate genes for human disorders, and findings from human genome-wide association studies (GWAS) may lead to the development of important new animal models (Dixon et al., 2011).

Other constraints were also considered in establishing this Consortium. With a goal of incorporating many different types of bioinformatics data about specific developmental processes and disorders, the FaceBase projects could not cover the entire panoply of structures subsumed under “dental and craniofacial” and still have enough in common for the desired data integration with one another. The decision was therefore made to restrict the first iteration of FaceBase to studies investigating the “mid-facial” region (e.g., nose, upper lip, palate, etc.), because it is the anatomic site of several common craniofacial birth defects such as cleft lip/palate and frontonasal dysplasia (Naidich et al., 2003). Clefts involving the midface, in particular those of the lip and the palate, are among the most common birth defects and create life-long challenges for affected individuals, including repeated surgeries, as well as speech, and dental problems and long-term health issues such as increased risks for cancer and delayed cognitive development (Bille et al., 2005; Conrad et al., 2009). Thus, research into the development of the mid-face has relevance to the mission of the NIDCR and addresses significant health burdens.

As summarized in the next section, the FaceBase Consortium researchers studying mid-face development utilize diverse organisms including fish, mice and humans. Investigators employ a wide variety of experimental approaches such as GWAS, detailed phenotyping, transcriptional profiling, and conditional knockout mice; and/or are working on technological advances to assist in the FaceBase goals.

FaceBase Consortium projects

Consortium efforts can be broadly classified into three categories: (1) five animal model projects (primarily mouse and zebrafish) examining the role of various developmental and functional elements in the midfacial region; (2) three human projects, two investigating the genetic determinants of facial morphometrics and structural birth defects; (one of which also

includes mouse models for gene discovery) and one utilizing sequencing approaches to extend GWAS results in CL/P; and (3) two technology projects developing software and animal models to support the FaceBase Consortium goals and craniofacial research community. Importantly, the aims of several projects overlap considerably across these domains. The data management and coordination hub project serves as the focal point for developing Consortium policies and the centralized software and hardware infrastructure needed to support the FaceBase data resource and web portal and for housing the human biorepository. An overview of the various projects and key investigators is given in Table 1, with a graphical representation in Fig. 1.

Animal model projects

Functional analysis of neural crest and palate: imaging craniofacial development

As craniofacial development involves tissue movements and cell rearrangements at scales difficult to capture fully in conventional histological sections, techniques combining genetic data with detailed imaging are needed to provide more detailed understanding of genetic influences and processes influencing craniofacial development. This research project will refine and deploy a set of advanced tools for the imaging of tissue structure, gene expression domains and cellular dynamics during craniofacial development. Volumetric imaging tools will be used to create accurate 3D atlases that can be digitally dissected to permit tissue interactions and cellular events to be better understood in the forming faces of normal, mutant and perturbed mouse embryos. Molecular agents, optimized for imaging intact tissues, will be employed to create atlases of the molecular correlates in these embryos. Finally, intravital imaging tools will be used to study cell and tissue interactions as they take place, offering a view into the dynamic events that occur during craniofacial development. These data will be assembled into atlases that will be linked to other FaceBase data to provide unprecedented tools for exploring the cellular, tissue and molecular correlates of craniofacial development.

Genome-wide atlas of craniofacial transcriptional enhancers

Accumulating evidence from GWAS indicates that sequence variation in non-coding regions strongly contributes to a variety of clinical disorders including orofacial clefting (Birnbaum et al., 2009; Visel et al., 2009b; Beaty et al., 2010). These variants may impact disease by affecting functional properties of distant-acting transcriptional enhancers. However, very few isolated examples of such regulatory variation have been identified. This is in large part due to the fact that the genomic location and function of the vast majority of enhancers in the human genome remains unknown. This FaceBase project will use integrated genomic and transgenic mouse strategies to discover enhancers involved in face and palate development, and to characterize their activities. Specifically, a ChIP-seq approach (Visel et al., 2009a) will be employed to identify genome-wide sets of enhancers active in mouse face and palate tissues at embryonic stages relevant for orofacial clefting. A transgenic mouse enhancer screen will be utilized to validate and characterize subsets of these enhancer predictors in detail by determining their *in vivo* activity. Furthermore, disease-associated variants from GWAS and other human genetic studies that map to craniofacial enhancers will be interrogated to determine how these variants affect *in vivo* enhancer activity. The genomic and *in vivo* datasets, as well as molecular reagents developed through these experiments will be made available to other researchers through the FaceBase website.

Global gene expression atlas of craniofacial development

High-resolution genome-wide views of temporal, spatial, and tissue-and cell-specific gene expression patterning can provide mechanistic insights and facilitate hypotheses concerning normal and abnormal developmental mechanisms. This project will use both laser capture

microdissection and fluorescence activated cell sorting (FACS), together with microarrays and next generation sequencing, to create a global gene expression atlas of craniofacial development. A combination of morphologic, lectin staining, and transgenic GFP expression features will be used to identify specific compartments and lineages including the neural folds, the epidermal ectoderm, neural crest, paraxial mesoderm, nasal placodes and pits, lateral and medial facial eminences, maxilla and mandibular processes, signaling centers, and the palatal shelves. Laser capture will allow purification of discrete structures, while transgenic GFP/FACS will isolate specific cell types. In addition, to better define gene expression heterogeneity within individual cell types, we will perform extensive single cell analyses. Whole genome microarrays will provide global, sensitive and quantitative measures of gene expression levels. Next-generation RNA-seq will provide a digital readout of gene expression levels, cross-validate microarray data, provide additional valuable information concerning alternative processing, and detect expression of genes not well represented on arrays. Using integrative bioinformatics analysis approaches, gene expression patterns that are reflective of variously differentiating components will be linked to gene networks implicated by specific or shared structural, functional, and interactome features of the hundreds of genes already known to play individual roles in craniofacial development. This will provide an important framework for integrating other craniofacial projects.

Identification of miRNAs involved in midfacial development and clefting

The regulation of biological processes occurs through intricate and continuous refinement of gene expression. MicroRNAs (miRNAs) are small non-coding RNAs implicated as a mechanism for controlling gene expression in a wide variety of developmental processes (Bernstein et al., 2003). Previous results have shown that miR-140-mediated regulation of platelet-derived growth factor a (Pdgfa) is required for normal migration of a subset of neural crest cells towards the oral ectoderm, where they later take part in palate development (Eberhart et al., 2008). This current project will identify other miRNAs involved in vertebrate midface development, and determine their function in this process. The temporal and spatial expression patterns of miRNAs in the developing mouse maxillary/frontonasal prominences between E10.5 and E14.5 will be characterized using massively parallel miRNA sequencing (miRNA-seq). In situ expression patterns of identified miRNAs will then be compared in mouse and zebrafish embryos to define those that show a conserved pattern of expression. Finally, gain-and loss-of-function analysis in zebrafish will be used to determine the function of individual miRNAs. Bioinformatic interrogation of putative miRNA targets and comparisons to gene expression atlases will further elucidate the genetic networks that regulate craniofacial development.

Functional genomics, image analysis and rescue of cleft

The analysis of mutant animal models has significantly improved our understanding of the genetic causes of cleft palate. Human linkage studies have also shown that genetic mutations are a major contributing factor in the etiology of cleft palate. For example, mutations in transforming growth factor-β (Tgf-β) signaling can cause cleft palate in both mice and humans (Nawshad et al., 2004). In mice, loss of Tgf-β signaling in cranial neural crest (CNC) cells (*Tgfbr2*$^{fl/fl}$;*Wnt1-Cre*) results in complete cleft palate whereas loss of Tgf-β in midline epithelial cells (*Tgfbr2*$^{fl/fl}$;*K14-Cre*) results in submucous cleft (Ito et al., 2003; Xu et al., 2006). These two animal models represent two common types of cleft palate in humans. To improve the utility of mutant animal models for investigating genetic causes of cleft palate in humans, this project will develop a cleft palate classification system to greatly facilitate the organization of data and assist the coordination between mouse and human cleft palate research. This classification system will serve to standardize vocabulary and phenotypic descriptions for all researchers to communicate effectively. Furthermore, in collaboration with other projects of the FaceBase Consortium, global and specific gene

expression profiling analysis in *Tgfbr* mutant animal models will be used to continuously generate candidate genes critical for CNC cell fate determination during palatogenesis. In parallel, sophisticated imaging analysis (microMRI and microCT) will be used to build a comprehensive database for investigation of the regulatory mechanisms of palatogenesis. Dissecting distinct signaling pathways and identifying the point(s) of intersection where multiple signaling pathways converge will aid in developing therapeutic strategies to prevent cleft palate and/or restore palate formation.

Human projects

Genetic determinants of orofacial shape and relationship to cleft lip/palate

Approximately 70% of all cleft lip and/or cleft palate occur as sporadic and isolated abnormalities (Stanier and Moore, 2004). Such “non-syndromic” orofacial clefts act as complex traits, involving multiple genetic and environmental risk factors. There is considerable evidence that orofacial malformations can occur at the extremes of the normal ranges of phenotypic variation of midfacial size and shape. Therefore, genes which control normal orofacial size and shape could have important roles in the occurrence of orofacial clefts. To identify such genes, this project will undertake detailed morphometric analysis of midfacial shape differences in informative mouse strains, as well as in select human populations. Combining these studies with genetic analyses will identify genes controlling midfacial morphometries. Specific inbred strains of mice have heritable differences in measurable parameters of facial shape. This project takes advantage of a valuable new resource, the mouse “Collaborative Cross” (CC; (Churchill et al., 2004)) to correlate heritable differences in facial shape among the eight founder strains of the CC, along with select Recombinant Inbred lines and Recombinant Intercross (RIX), with detailed genetic mapping data for these mice. It will also generate the largest repository of mouse microCT scan data that can be used for multiple future genetic and morphometric studies. This approach will enable identification of quantitative trait loci (QTLs) that underlie these morphometric differences. The mouse studies will be complemented by a similar analysis of humans, studying specific populations with different susceptibilities to orofacial clefts. These comparative studies will identify genes that underlie midfacial shape in humans. Together, these studies should provide a basis for understanding the relationship between human facial morphogenesis and susceptibility to orofacial clefts, and for initiating studies of the functions of these genes in animal models relevant to human orofacial development.

3D analysis of normal facial variation: data repository and genetics

Although ample evidence exists that facial appearance and structure are highly heritable, there is little information regarding how variation in specific genes relates to the diversity of facial forms evident in our species. With the advent of affordable, non-invasive 3D surface imaging technology, it is now possible to capture detailed quantitative information about the face in a large number of individuals (Kau et al., 2010). The combination of state-of-the-art 3D imaging with advances in high-throughput genotyping provides an unparalleled opportunity to map the genetic determinants of normal facial variation. An improved understanding of the relationship between genotype and facial phenotype may help illuminate the factors influencing sensitivity to common craniofacial anomalies, particularly orofacial clefts, which are among the most prevalent birth defects in humans. This project will construct a normative repository of 3D facial and genetic data and utilize this data repository to identify genes that influence normal midfacial variation. This will involve collecting 3D facial surface images and DNA samples on 3500 healthy Caucasian individuals (age 3–40) drawn from the general population. Quantitative facial measures will be extracted from the 3D images and all DNA samples will be genotyped for genome-wide SNP markers. All of the 3D images, quantitative measures, and genotype data will be

available to outside investigators through the FaceBase repository. The project will focus on identifying SNPs associated with variation in midfacial morphology, including those facial features relevant to orofacial cleft predisposition. Salient measures of midfacial morphology will be derived from 3D facial surface images, and a genome-wide association approach will then be employed to identify polymorphisms that influence quantitative variation in the facial features of interest.

Oral clefts: moving from genome-wide studies toward functional genomics

The FaceBase Consortium provides a timely opportunity for follow-on GWAS of oral clefts (Beaty et al., 2010). A systematic analysis of intensity data for single nucleotide polymorphic (SNPs) markers and monomorphic probes in regions of known copy number variants (CNV) available from our genome wide association study will be performed to identify genes that influence risk through structural variation. CNV markers will then be used in a second genome wide test for linkage and association under a case-parent trio design, which may identify additional genes of interest. The case-parent trio design offers a unique opportunity to identify de novo CNVs, and this information will be combined with evidence from transmitted CNVs to identify influential genes. In collaboration with other researchers in FaceBase, we have the opportunity to undertake high-throughput sequencing (HTS) studies of specific genes and chromosomal regions identified in our GWAS to identify both rare and common variants that may play a causal role in the etiology of oral clefts. The genes/ regions identified in our case-parent trio design will be the first area of HTS studies, but we will go on to further conduct whole exome HTS studies using affected pairs of relatives drawn from multiplex families. Finding from these sequencing studies should identify new candidates for functional studies in animal models through collaboration with other FaceBase projects. We are particularly interested in genes that show some evidence of gene–environment interaction in human data, because animal models offer the opportunity to further explore the combined effects of genes and environmental risk factors.

Technology projects

Shape-based retrieval of 3D craniofacial data

As shape is a critical factor in the classification of most craniofacial disorders, computational tools for analyzing 3D shape are essential to better describe individual conditions as well as understand their pathogenesis. Quantitative shape descriptors allow for reproducible shape description, while similarity-based shape retrieval allows comparisons to be made between individuals or populations. Researchers studying disorders of craniofacial anatomy have access to a number of 3D imaging tools, including computed tomography, magnetic resonance imaging, and 3D surface scans (Robb, 1999). This project will (1) develop software producing quantitative representations of craniofacial anatomy to assist in the study of mid-face hypoplasia and cleft lip and palate; (2) develop tools for quantifying the similarity of craniofacial data between two individuals, between an individual and an average over a selected population, or between two populations; (3) develop mechanisms for organization and retrieval of multimodality 3D craniofacial data based on their quantitative representations; (4) design and implement a prototype system for Craniofacial Information Retrieval (CIR) that incorporates quantification, organization, and retrieval; and (5) utilize the CIR System on 3D craniofacial data, such as that generated by other FaceBase projects. The design of these tools and a pilot system will lead to a general methodology that is both immediately applicable to studies of mid-face hypoplasia, cleft lip and cleft palate, and also scalable and modifiable to all craniofacial abnormalities.

Genetic tools and resources for orofacial clefting research

The mouse is a powerful genetic model for understanding developmental mechanisms and the etiology and pathogenesis of human syndromes. As future progress requires an increasingly sophisticated set of genetic models and tools, this project will generate new Cre recombinase driver strains for orofacial clefting research and serve as a mouse repository for these and other relevant strains for the community as a whole. Four specific BAC transgenic Cre drivers, (*Lhx7/8-Cre; deltaNp63-CreER^{T2}*, *Krt6a-Cre*, *Tbx22-CreER^{T2}*) and three conserved enhancer driven Cre strains (*IRF6(MCS-9.7)-CreER^{T2}*, *HCES-546-CreER^{T2}*, *HCES-809-CreER^{T2}*) were chosen by a working group of experts in the field and are in progress. These strains were chosen to fill gaps in the existing repertoire of Cre strains, and to provide tools to interrogate the mechanisms of orofacial development in detail. Eight additional constructs will be selected based on data generated in other FaceBase projects centered on global gene expression analyses and identification of transcriptional enhancers.

The *Lhx7/8-Cre* driver will target the mesenchymal compartment directly adjacent to mid-facial fusion events, facilitating investigations of mesenchymal–epithelial interactions prior to and during mid-face fusion. Expression of *Lhx7/8* is first observed at E9.5 in the first branchial arch and by E10.5 is strongly expressed in the maxillary and mandibular components of the arch, proximal to the cleft that separates them (Grigoriou et al., 1998). Because expression in the mesenchyme of the medial nasal prominence only is noted just prior to fusion with the maxillary prominence, a standard Cre driver is the best approach in this case.

During secondary palate development, the deltaN alpha isoform of *Trp63* is expressed throughout the basal epithelial cell layer of the oral ectoderm, but is excluded from the periderm cell layer (Thomason et al., 2008). The *deltaNp63-CreER^{T2}* driver line will allow investigators to specifically target the basal epithelial cell layer in the oral cavity without affecting the periderm and mesenchyme. Because expression of deltaNp63 occurs prior to stratification of the periderm, an inducible strategy is required to produce a Cre line with the desired specificity.

The *Krt6a-Cre* driver will serve to specifically target the periderm cell layer during palate shelf fusion. Expression of mouse *Krt6a* begins at approximately E14.5 in the periderm cell layer and persists until the periderm disappears before birth (Mazzalupo and Coulombe, 2001; Wong et al., 2000). Because expression is activated just prior to the fusion events in the secondary palate, a constitutive Cre driver is best suited to target the periderm during palate fusion.

Tbx22 is primarily expressed in the mid to posterior palate, but excluded from the anterior (Bush et al., 2002; Liu et al., 2008; Welsh et al., 2007). The development of the mammalian secondary palate appears to be differentially regulated along the anterior–posterior axis. Both proliferation rates and gene expression patterns are variable along this axis, and this patterning appears critical for normal development and fusion of the secondary palate (Hilliard et al., 2005; Zhang et al., 2002). Thus, the *Tbx22-Cre* driver strain will allow users to interrogate the function of anterior-posterior gene patterning during secondary palate development.

The first of the enhancer drivers, *IRF6(MCS-9.7)-CreER^{T2}*, is based on a critical enhancer element of the *IRF6* gene (Rahimov et al., 2008). The element drives expression throughout the epithelium covering the first and second branchial arches and is specifically enriched at E11.5 at the sites of fusion between the lateral and medial nasal prominences and maxillary prominence. The expression of the HCES-809 enhancer is prominent in the distal portion of the first branchial arch, the rostral aspect of the second arch, and the caudal aspect of the

frontonasal prominence. Expression is also noted in the cranial nerve and dorsal root ganglia of the trunk. Finally, the HCES-546 enhancer is expressed at E11.5 throughout the frontonasal prominence, the limb buds and variably in the eye. Notably, expression is excluded from the branchial arches, including the maxillary prominence, enabling the user to distinguish between these domains in the developing upper lip.

All Cre strains will be carefully characterized and made available to the community through The FaceBase Mouse Repository at the Jackson Laboratory (JAX), which will reside within the Genetic Resource Science (GRS) Repository and will be managed by GRS personnel with the same high standards as for other JAX mouse resources. The mission of the FaceBase Mouse Repository is to provide orofacial clefting mouse models and tool strains of known genetic background and high health status quickly to requesting investigators.

The FaceBase data management and coordination hub

The FaceBase Management and Coordination Hub provides administrative and management functions for FaceBase, the development of software for the FaceBase data repositories and FaceBase research web portal, and support for other related activities such as the FaceBase biorepository. The FaceBase web portal will allow access to the FaceBase data and sample repository — storing and providing access to the diverse types of data generated by Consortium members and including data and samples contributed by craniofacial researchers not directly supported by the FaceBase initiative. It will also serve as a community forum for craniofacial researchers and educational outreach.

The FaceBase Hub is built on the Drupal (http://www.drupal.org) open-source content management system. Drupal's core system and extensive modular library will provide a basis for rapidly deploying community functionality including user accounts, email fora, and file management. Additional modules and web service interfaces will be developed to provide tools for managing and using FaceBase data, with an eye towards general application to other translational bioinformatics data portals.

With multiple teams of investigators generating data from at least three organisms (mice, humans, and zebrafish) and numerous mutant phenotypes under study, effective data exchange, cross referencing and integration require shared vocabularies — FaceBase researchers must literally learn to talk to one another. Terminologies for phenotypes (Robinson et al., 2008) and anatomy for mice (Burger et al., 2004), humans (Rosse and Mejino, 2003), and zebrafish (www.zfin.org) will play an important role. Terminology alignment efforts mapping between these organisms are underway, along with preliminary extensions of existing anatomic and phenotypic ontologies (Robinson and Mundlos, 2010) as needed to provide the detail needed for discussion of craniofacial abnormalities such as orofacial clefting.

Consortium efforts will generate potentially individually identifiable human data that must be protected and handled appropriately. Effective use of these data will require systems informatics that helps interested researchers both identify data of interest and request access to sensitive data, without compromising privacy. Policies and procedures for handling such data access have been developed, with likely implications for system architecture, administrative tools, and end-user interfaces. Similarly, sample access to a collection of cases with craniofacial malformations and annotated clinical and family data will also be housed in secure storage sites and be de-identified so that contributors cannot be identified. These samples and data will be available to users with legitimate scientific interests.

Effective use of data developed by the FaceBase Consortium also requires consideration of the cognitive and analytic challenges faced by biologists in using complex data sets to

inform their research. User interfaces in support of biological inquiry must provide appropriate contextual information, multiple scale views, and interactive querying and reporting facilities to support higher-level analytic processes and modeling (Mirel, 2009). Volumetric datasets such as microCT scans and optical projection tomography data pose significant challenges in terms of data storage, access and searchability, as well as visualization. This is particularly true for web-based interfaces, in which transfer of the data itself is very slow and local manipulation requires specialized software and computational power. Building on recent research in interactive search and discovery methods (Chau et al., 2008), we will develop interfaces to support interactive data exploration and help scientists uncover unanticipated links between data of different types collected from multiple organisms.

FaceBase and the wider research community

FaceBase is intended and designed to be a resource for the wider craniofacial research community, and not merely the FaceBase Consortium. To achieve this goal, the Consortium is pursuing several directions. The first will be the availability of freely browsable content such as gene expression and anatomic atlases, unified and detailed terminology for describing craniofacial development, averaged morphometric data and images, and shared resources such as animal models and protocols. A craniofacial Gene Wiki maintained by volunteers will serve to bring together the latest information about genes known to play a role in craniofacial development and disorders (Huss et al., 2010, 2008). The second will be the availability of controlled access datasets, for example, microarray and RNA-seq data, that can be independently analyzed and mined by outside researchers. The Consortium has established controlled data access procedures to assess the appropriateness of requests from the wider research community, especially datasets containing potentially identifiable information from individual human subjects (e.g. individual level facial 3D images). A third will be to use the FaceBase website as a central portal to accept datasets generated outside Consortium projects, and to disseminate these data through FaceBase. In situations where relevant data are already being disseminated through other sites, FaceBase will link to those sites such as dbGaP for GWAS. In situations where those data are not available by other means, the Consortium will consider providing them via the FaceBase website. Procedures and criteria have been established to ensure that appropriate consents/permission has been obtained before making these data publicly available. The biorepository function is also recruiting contributions from the clinical community. Their engagement will also provide a community for input into the relevance of the scientific findings and for future translational studies.

Lastly, the forums available through the FaceBase web portal will serve as a community resource for exchanging ideas, for discussing new discoveries, for highlighting emerging opportunities and technologies, and for sharing research approaches and methods relevant to the craniofacial research community. We hope this effort will facilitate the short term goal of data production and integration as well as the long term goal of translational and clinical application of this knowledge for the prevention, treatment and management of craniofacial birth defects. The role that such an academically oriented effort can also play in informing and teaching the non-expert lay community will also be explored in the future.

Acknowledgments

The FaceBase Consortium thanks the many investigators and staff members for working collaboratively and untiringly toward fulfilling the goals of the program. In particular, we appreciate the vision and support of Dr. Lawrence A. Tabak, former Director of the National Institute of Dental and Craniofacial Research (NIDCR), in enabling this effort. We thank the NIDCR Program Staff for their advice in implementing FaceBase and suggestions for this manuscript (Program Official: Steven Scholnick; Project Scientists: Emily Harris, Nadya Lumelsky, Lillian Shum). We are very grateful to Carolyn Randolph for editorial assistance in preparing this paper.

Projects comprising the FaceBase Consortium are supported by grants from the NIDCR, U01-DE20049 (S.S.P., B.A., P.A.T), U01-DE20050 (L.S., M.L.C., J.F.B), U01-DE20052 (L.R.B.D., S.A.M.), U01-DE20054 (R.A.S.), U01-DE020057 (J.C.M., M.L.M., H.H. M.D.), U01-DE20060 (A.V.), U01-DE20063 (S.E.F.), U01-DE20065 (Y.C., J.I.), U01-DE20073 (T.H.B.), U01-DE20076 (D.E.C., K.B.A., J.H.P.) and U01-DE20078 (S.M.W., M.L.M. M.L.C).

References

Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, Redett RA, Raymond G, Schwender H, Jin SC, Cooper ME, Dunnwald M, Mansilla MA, Leslie E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M, Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL, Molloy AM, Kirke PN, Scott JM, Arcos-Burgos M, Scott AF. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet. 2010; 42:525–529. [PubMed: 20436469]

Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KV, Hannon GJ. Dicer is essential for mouse development. Nat Genet. 2003; 35:215–217. [PubMed: 14528307]

Bille C, Winther JF, Bautz A, Murray JC, Olsen J, Christensen K. Cancer risk in persons with oral cleft—a population-based study of 8093 cases. Am J Epidemiol. 2005; 161:1047–1055. [PubMed: 15901625]

Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freudenberg J, Lauster C, Schmidt G, Scheer M, Braumann B, Berge SJ, Reich RH, Schiefke F, Hemprich A, Potzsch S, Steegers-Theunissen RP, Potzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nothen MM, Mangold E. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet. 2009; 41:473–477. [PubMed: 19270707]

Bult CJ, Eppig JT, Kadin JA, Richardson JE, Blake JA. The Mouse Genome Database (MGD): mouse biology and model systems. Nucleic Acids Res. 2008; 36:D724–D728. [PubMed: 18158299]

Burger A, Davidson D, Baldock R. Formalization of mouse embryo anatomy. Bioinformatics. 2004; 20:259–267. [PubMed: 14734318]

Bush JO, Lan Y, Maltby KM, Jiang R. Isolation and developmental expression analysis of Tbx22, the mouse homolog of the human X-linked cleft palate gene. Dev Dyn. 2002; 225:322–326. [PubMed: 12412015]

Chau, D.; Myers, B.; Faulring, A. What to do When Search Fails: Finding Information by Association. Proceeding of the twenty-sixth annual SIGCHI conference on Human factorsin computing systems; ACM, Florence, Italy. 2008. p. 999-1008.

Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, Beavis WD, Belknap JK, Bennett B, Berrettini W, Bleich A, Bogue M, Broman KW, Buck KJ, Buckler E, Burmeister M, Chesler EJ, Cheverud JM, Clapcote S, Cook MN, Cox RD, Crabbe JC, Crusio WE, Darvasi A, Deschepper CF, Doerge RW, Farber CR, Forejt J, Gaile D, Garlow SJ, Geiger H, Gershenfeld H, Gordon T, Gu J, Gu W, de Haan G, Hayes NL, Heller C, Himmelbauer H, Hitzemann R, Hunter K, Hsu HC, Iraqi FA, Ivandic B, Jacob HJ, Jansen RC, Jepsen KJ, Johnson DK, Johnson TE, Kempermann G, Kendziorski C, Kotb M, Kooy RF, Llamas B, Lammert F, Lassalle JM, Lowenstein PR, Lu L, Lusis A, Manly KF, Marcucio R, Matthews D, Medrano JF, Miller DR, Mittleman G, Mock BA, Mogil JS, Montagutelli X, Morahan G, Morris DG, Mott R, Nadeau JH, Nagase H, Nowakowski RS, O'Hara BF, Osadchuk AV, Page GP, Paigen B, Paigen K, Palmer AA, Pan HJ, Peltonen-Palotie L, Peirce J, Pomp D, Pravenec M, Prows DR, Qi Z, Reeves RH, Roder J, Rosen GD, Schadt EE, Schalkwyk LC, Seltzer Z, Shimomura K, Shou S, Sillanpaa MJ, Siracusa LD, Snoeck HW, Spearow JL, Svenson K, Tarantino LM, Threadgill D, Toth LA, Valdar W, de Villena FP, Warden C, Whatley S, Williams RW, Wiltshire T, Yi N, Zhang D, Zhang M, Zou F. The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat Genet. 2004; 36:1133–1137. [PubMed: 15514660]

Conrad AL, Richman L, Nopoulos P, Dailey S. Neuropsychological functioning in children with non-syndromic cleft of the lip and/or palate. Child Neuropsychol. 2009:1–14.

Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 2011; 12:167–178. [PubMed: 21331089]

Eberhart JK, He X, Swartz ME, Yan YL, Song H, Boling TC, Kunerth AK, Walker MB, Kimmel CB, Postlethwait JH. MicroRNA Mirn140 modulates Pdgf signaling during palatogenesis. Nat Genet. 2008; 40:290–298. [PubMed: 18264099]

Grigoriou M, Tucker AS, Sharpe PT, Pachnis V. Expression and regulation of Lhx6 and Lhx7, a novel subfamily of LIM homeodomain encoding genes, suggests a role in mammalian head development. Development. 1998; 125:2063–2074. [PubMed: 9570771]

Hilliard SA, Yu L, Gu S, Zhang Z, Chen YP. Regional regulation of palatal growth and patterning along the anterior–posterior axis in mice. J Anat. 2005; 207:655–667. [PubMed: 16313398]

Huss JW III, Orozco C, Goodale J, Wu C, Batalov S, Vickers TJ, Valafar F, Su AI. A gene wiki for community annotation of gene function. PLoS Biol. 2008; 6:e175. [PubMed: 18613750]

Huss JW III, Lindenbaum P, Martone M, Roberts D, Pizarro A, Valafar F, Hogenesch JB, Su AI. The Gene Wiki: community intelligence applied to human gene annotation. Nucleic Acids Res. 2010; 38:D633–D639. [PubMed: 19755503]

Ito Y, Yeo JY, Chytil A, Han J, Bringas P Jr, Nakajima A, Shuler CF, Moses HL, Chai Y. Conditional inactivation of Tgfbr2 in cranial neural crest causes cleft palate and calvaria defects. Development. 2003; 130:5269–5280. [PubMed: 12975342]

Kau CH, Richmond S, Zhurov A, Ovsenik M, Tawfik W, Borbely P, English JD. Use of 3-dimensional surface acquisition to study facial morphology in 5 populations. Am J Orthod Dentofacial Orthop. 2010; 137:S56, e1–9. discussion S56–7. [PubMed: 20381762]

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res. 2002; 12:996–1006. [PubMed: 12045153]

Liu W, Lan Y, Pauws E, Meester-Smoor MA, Stanier P, Zwarthoff EC, Jiang R. The Mn1 transcription factor acts upstream of Tbx22 and preferentially regulates posterior palate growth in mice. Development. 2008; 135:3959–3968. [PubMed: 18948418]

Mazzalupo S, Coulombe PA. A reporter transgene based on a human keratin 6 gene promoter is specifically expressed in the periderm of mouse embryos. Mech Dev. 2001; 100:65–69. [PubMed: 11118885]

Mirel B. Supporting cognition in systems biology analysis: findings on users' processes and design implications. J Biomed Discovery Collab. 2009; 4:2.

Naidich, TP.; Blaser, SI.; Bauer, BS.; Armstrong, DC.; McLone, DG.; Zimmerman, RA. Embryology and Congenital Lesions of the Midface. In: Som, PM.; Curtin, HD., editors. Head and Neck Imaging. 4. Mosby, New York: 2003.

Nawshad A, LaGamba D, Hay ED. Transforming growth factor beta (TGFbeta) signalling in palatal growth, apoptosis and epithelial mesenchymal transformation (EMT). Arch Oral Biol. 2004; 49:675–689. [PubMed: 15275855]

Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, Domann FE, Govil M, Christensen K, Bille C, Melbye M, Jugessur A, Lie RT, Wilcox AJ, Fitzpatrick DR, Green ED, Mossey PA, Little J, Steegers-Theunissen RP, Pennacchio LA, Schutte BC, Murray JC. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. Nat Genet. 2008; 40:1341–1347. [PubMed: 18836445]

Richardson L, Venkataraman S, Stevenson P, Yang Y, Burton N, Rao J, Fisher M, Baldock RA, Davidson DR, Christiansen JH. EMAGE mouse embryo spatial gene expression database: 2010 update. Nucleic Acids Res. 2010; 38:D703–D709. [PubMed: 19767607]

Robb RA. 3-D visualization in biomedical applications. Annu Rev Biomed Eng. 1999; 1:377–399. [PubMed: 11701494]

Robinson PN, Kohler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. Am J Hum Genet. 2008; 83:610–615. [PubMed: 18950739]

Robinson PN, Mundlos S. The human phenotype ontology. Clin Genet. 2010; 77:525–534. [PubMed: 20412080]

Rosse C, Mejino JL Jr. A reference ontology for biomedical informatics: the foundational model of anatomy. J Biomed Inform. 2003; 36:478–500. [PubMed: 14759820]

Stanier P, Moore GE. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. Hum Mol Genet. 2004; 13(Spec No 1):R73–R81. [PubMed: 14722155]

Thomason HA, Dixon MJ, Dixon J. Facial clefting in Tp63 deficient mice results from altered Bmp4, Fgf8 and Shh signaling. Dev Biol. 2008; 321:273–282. [PubMed: 18634775]

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009a; 457:854–858. [PubMed: 19212405]

Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. Nature. 2009b; 461:199–205. [PubMed: 19741700]

Welsh IC, Hagge-Greenberg A, O'Brien TP. A dosage-dependent role for Spry2 in growth and patterning during palate development. Mech Dev. 2007; 124:746–761. [PubMed: 17693063]

Wong P, Colucci-Guyon E, Takahashi K, Gu C, Babinet C, Coulombe PA. Introducing a null mutation in the mouse K6alpha and K6beta genes reveals their essential structural role in the oral mucosa. J Cell Biol. 2000; 150:921–928. [PubMed: 10953016]

Xu X, Han J, Ito Y, Bringas P Jr, Urata MM, Chai Y. Cell autonomous requirement for Tgfbr2 in the disappearance of medial edge epithelium during palatal fusion. Dev Biol. 2006; 297:238–248. [PubMed: 16780827]

Zhang Z, Song Y, Zhao X, Zhang X, Fermin C, Chen Y. Rescue of cleft palate in Msx1-deficient mice by transgenic Bmp4 reveals a network of BMP and Shh signaling in the regulation of mammalian palatogenesis. Development. 2002; 129:4135–4146. [PubMed: 12163415]

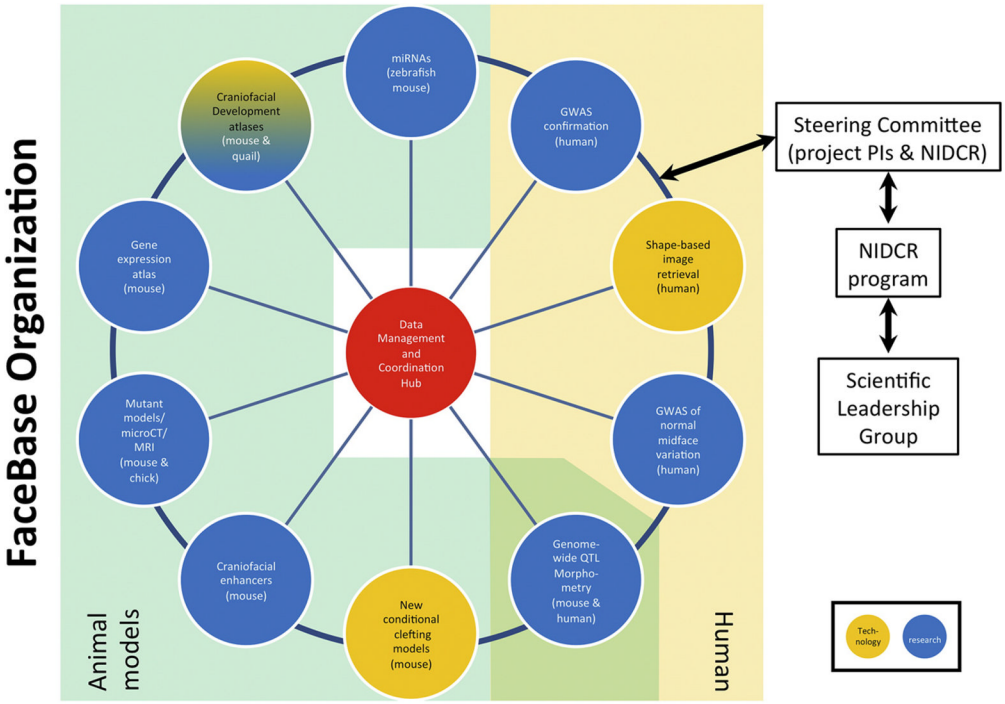

Fig. 1.

The FaceBase Consortium includes five projects using a variety of techniques to investigate craniofacial development, morphometry, and genetics in model organisms (mouse, zebrafish, chick, and quail); three projects investigating human genetics and morphometry, and two technology projects developing new mouse strains and algorithms for shape-based image retrieval. The Data Management and Coordination Hub is responsible for coordinating the various efforts and developing a web portal which will host the data and support collaboration. A steering committee consisting of the project PIs and program staff provide oversight, with the guidance of program officers and an external scientific leadership group.

Dev Biol. Author manuscript; available in PMC 2012 September 12.

Table 1

Summary of key personnel, organisms and data to be generated for each FaceBase project.

Project type	Project	PIs, co-is and key personnel	Organisms	Developmental stage	Data types
FaceBase hub					
	FaceBase management and coordination hub	Mary Marazita Jeffrey Murray Harry Hochheiser Michael Becich Michael Cunningham Michael Dixon David FitzPatrick Jill Helms Robert Kuhn Alan Scott Andrew Su			
Animal models					
	Functional analysis of neural crest and palate	Scott Fraser Marianne Bronner Yang Chai	Mouse Chick	Embryos	3D high-resolution MRI, gene expression
	Genome-wide atlas of craniofacial transcriptional enhancers	Axel Visel David FitzPatrick	Mouse	Embryos	ChIP-Seq, RNA-Seq, transgenic mouse data, Optical Projection Tomography (OPT) imaging
	Global gene expression atlas of craniofacial development	Steve Potter Bruce Aronow Paul Trainor	Mouse	Embryos	LCM/FACS microarray
	Identification of miRNAs involved in midfacial development and clefting	David Clouthier Kristin Artinger John Postlethwait	Mouse Zebrafish	Embryos	miRNA
	Research on functional genomics, image analysis and rescue of cleft palate	Yang Chai Scott Fraser Joseph Hacia Pedro Sanchez	Mouse	Embryos	Cleft palate models, microarray, microCT and microMRI images
Human					
	Genetic determinants of orofacial shape and relationship to cleft lip/palate	Richard Spritz Benedikt Hallgrimsson Ophir Klein Fernando Pardo-Manuel de Villena	Mouse Human	Embryonic/adult ages 3–12	3D mouse and human facial images, microCT, morphometrics, genotypes, GWAS
	3D analysis of normal facial variation	Seth Weinberg Mary Marazita Michael L. Cunningham Carrie Heike Jacqueline Hecht Chung How Kau Karen T. Cuenco	Human	Ages 3–40	3D human facial images, morphometrics, genotypes, GWAS

Project type	Project	PIs, co-is and key personnel	Organisms	Developmental stage	Data types
	Oral clefts: moving from genome-wide studies toward functional genomics	Terri Beaty Kung Yee Liang Alan F. Scott Hongkai Ji Ingo Ruczinski Jacqueline Hetmanski Tao Wu	Human	Varied	Genotypes, GWA, copy-number variation, pedigrees
Technology					
	Shape-based retrieval of 3D craniofacial data	Linda Shapiro Michael Cunningham James Brinkley	Human	Varied	Algorithms and software for 3D facial images
	Genetic tools and resources for orofacial clefting research	Leah Rae Donahue Stephen Murray	Mouse	Varied	Cre recombinase driver strains