Seeing Faces and History through Human Genome Sequences

CAS/MPG Partner Group on the Human Functional Genetic Variations

Shanghai-Leipzig, 2011.2.1–2016.1.31

Prof. Dr. TANG Kun (middle) with his cooperator, Prof. Dr. Mark Stoneking (right) and colleague, Prof. Dr. LI Haipeng (left).

Photos/Images: By courtesy of Prof. TANG Kun.

The scientific goals of the Partner Group are to utilize large scale genomic polymorphism data to make fine inferences about human demographic history and various forms of natural selection in the human genome; and to understand the genetic basis and evolutionary mechanisms underlying the common variations in human facial morphology, both within and between populations.

Reported by Group Leader TANG Kun

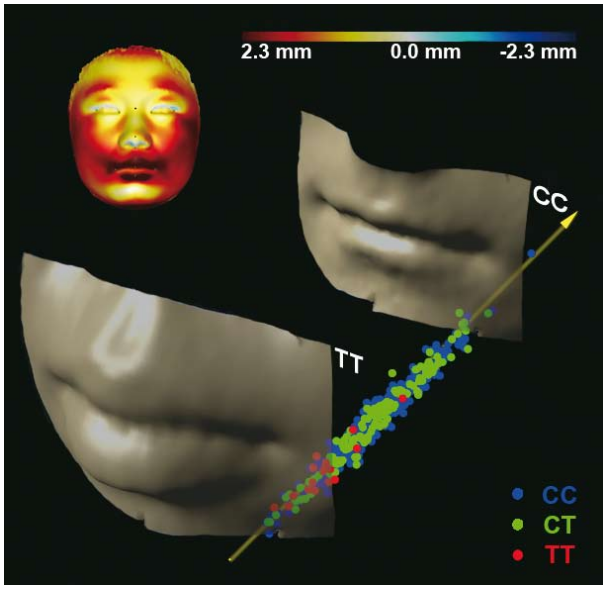

Figure 1: Detecting genetic association of common facial variation with high density 3D image registration: SNP rs642961, a genetic risk factor of NSCL/P, strongly predisposes the mouth shape in healthy Han Chinese females. Top: pixel-wise distance as heat-plot between the average facial shapes of wild types (CC) and mutants (TT). Bottom: mouth image projections on a discriminant hyper-line, where TT differ most from the other two genotypes. Middle: the modulation of rs642961 TT on mouth shape was simulated as 3D model (front) via linear transformation along the hyper-line, compared to the model simulating the opposite effects (behind). See Peng *et al*. 2013

ONGOING RESEARCH ACTIVITIES

Section 1: Demographics and Natural Selection in Recent Human Evolution

1. Demographic inference based on analysis of IBD/IBS

Demographics inference has always been an essential research question in population genetics and is crucial for reliable detection of signals of natural selection. We carried out a series of studies to find sufficient statistics that possesses simple mathematical relationships with the demographics, so that demographics can be inferred with high resolution and power. One set of novel statistics we developed is based on the haplotype extension, which can be used to infer demographic parameters of relatively complex models (Theunert *et al*, 2012) (In collaboration with the group led by Prof. Dr. Stoneking).

2. Studies of natural selection in human genome

For natural selection, we are interested in two fundamental questions: First, how did the genomic signals of natural selection fit into the big picture of recent human history; second, how do the footprints of ancient selection affect people's life today. Following these directions, several projects were carried out, including characterization of the global patterns of genetic diversity and signals of natural selection for human ADME genes (In collaboration with the group led by Prof. Dr. Stoneking), detection of signals of recent co-evolution in human populations as well as functional characterizing the genetic variants that carry strong evidence of recent positive selection.

3. The fine atlas of natural selection in human genome

To identify and accurately characterize the genomic loci of natural selection is a core problem in population genetics. Collaborating with Dr. Rostislav Matveev (MPI-EVA), we try to establish an integrated computational framework that addresses all the major questions in one shot, including controlling the effects of non-equilibrium demographics, detecting the loci of various natural selection scenarios (positive, negative and balancing selection), as well as the consequential parameter estimation, such as the selection times and selection coefficients for the recent positive selection events. The main idea is to first re-construct the coalescent trees across the whole genome, based on the pairwise coalescent estimation via PSMC approach (Li, H., & Durbin, R. (2011). *Nature*, 475(7357), 493–496). Using the tree data, several likelihood tests were developed to collectively assign all genome fragments to modes of neutral, negative, balancing or positive selection, and simultaneously estimate the key selection parameters. Interestingly, for the first time we are able to estimate the selection times for all the positive selection events in the genome. The temporal resolution of these selection events will shed light on the question as to what kinds of adaptation events corresponded to agriculture expansion (~ 10,000 years ago), migration and ice-age (20,000 ~ 30,000 years ago).

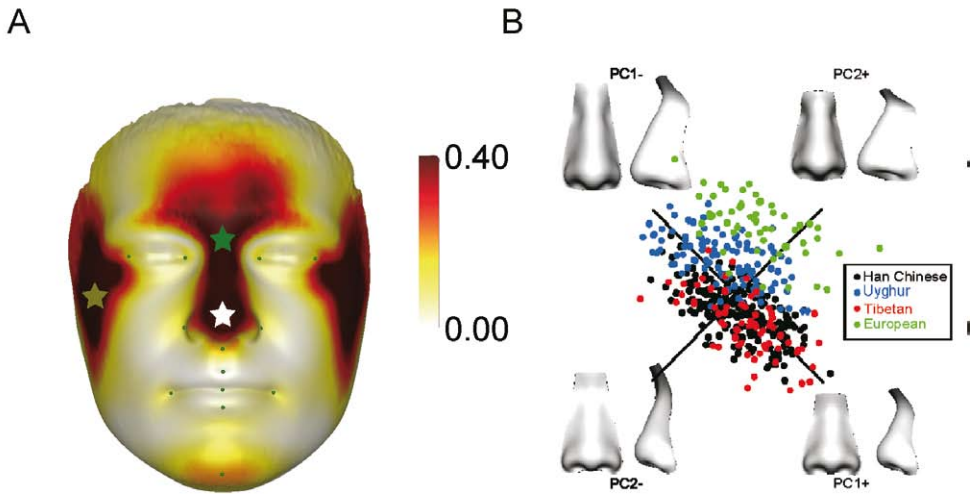

Figure 2: A. the point-wise Qst between Han Chinese and Europeans. The highest values on nasal tip, central eyebrows and cheeks (marked in stars) are 0.48, 0.46 and 0.58 respectively. B. the nose shape divergence on PC modes. PC2 almost completely separates Europeans (green) from Han Chinese (black). The pair-wise Qst on PC2 is 0.79.

Section 2: Genetic Basis and Evolutionary Mechanisms Underlying the Common Variations of Human Facial Morphology

The human face plays an essential role in everyday life. It hosts the most important sensory organs and acts as the central interface for expression, mutual identification, and communication. We would like to develop new technologies and do genetic studies to promote the understanding of the genetic bases and evolution of human faces.

1. Automatic landmark annotation and dense correspondence registration for 3D human facial images

We implemented a fully automatic and highly efficient 3D face image registration platform that features automatic 3D landmark registration and high density registration. Furthermore, this method is superior in accuracy and efficiency. With this new tool, we are able to carry out fully quantitative analyses on thousands of 3D facial images of different genders, ethnicities and ages.

2. Detecting genetic association of common human facial morphological variation

In this study, we applied the dense 3D face registration mentioned above to detect genetic associations. Different phenotype data schemes were evaluated in detecting the potential association of facial shape variations. The dense registration scheme showed slightly higher statistical power and stronger advantage in the fine inference of shape changes and 3D face modeling. One genetic variant, the rs642961-T in the gene IRF6, a known risk factor of non-syndromic cleft lips/palates, was strongly associated with more protrusive lips in the healthy female Han Chinese (see the figures below). Our method opens the possibility to systematically scan the subtle impacts of genetic variants on facial morphology (Figure 1) (Peng *et al*., 2013).

3. Variation and signatures of selection on the human face

Divergence of human facial shape across the globe has been a long-debated question. Previous studies based on skulls found low levels of divergence, therefore supported neutral drift as the major evolutionary force in the recent evolution of human facial morphology. In this study we analyzed high-resolution 3D images of soft-tissue facial surface in four Eurasian populations: Han Chinese, Tibetans, Uyghur and Europeans. The dense 3D face registration method allowed the facial shape diversity to be examined at unprecedented resolution. We found that noses, cheekbones and eyebrows exhibit strong signals of divergence (Qst estimates: 0.3~0.8) between Europeans and Han Chinese (see Figure 2). The highest divergence rate approaches that of skin pigmentation (~0.8). These results therefore suggest that some facial features likely underwent strong selection, either local adaptation or sexual selection or both. (In collaboration with Prof. Dr. Stoneking's group).

4. Genome-wide association studies of common facial morphological changes within and between populations

We are using pooling-based GWAS

INFO

Group Leader: Prof. Dr. TANG Kun
Founding Date: Feb 2011
Max Planck Institute for Evolutionary Anthropology
Prof. Dr. Svante Pääbo
Prof. Dr. Mark Stoneking
CAS-MPG Partner Institute of Computational Biology
Prof. Dr. JIN Li
Scientific Advisory Board Members
Prof. Dr. Mark Stoneking (MPI-EVA)
Prof. Dr. JIN Li (Fudan University)
Prof. Dr. Manfred Kayser (Erasmus MC)
Internet Addresses
http://www.picb.ac.cn
Phone: +86-21-54920277
tangkun@picb.ac.cn

STAFF

GUO Jing (Research Assistant)
LIU Chen (Technician)
CHEN Nianhao (Technician)

STUDENTS

PENG Shouneng
ZHOU Hang
HU Sile
QIAN Wei
QIAO Lu
GUO Jianya (part time 2011~2012)

technology to identify genetic loci associated with different facial shape variations. The main procedures include identification of highly variable/diverse features, classification of feature groups, subsequent DNA pooling based GWAS and the final re-validation of the candidate association signals using single SNP typing. For example, we observed strong divergence in nose shapes between Han Chinese and Europeans as described in study 3. To identify the genetic variants underlying such differences, we choose Uyghur, the admixed population between Europeans and East Asians. Uyghur noses were classified into two groups of European-like and Han-Chinese-like using the PC mode that distinguishes the European and Han noses the best. Pooling based GWAS was carried out on the two groups and several candidate loci were identified and partially re-validated in different groups.

FUTURE PERSPECTIVE

Direction 1: Demographic Inference and Fine Characterization of Natural Selection

The rapidly growing sequence data of human genome may provide unprecedented details about human recent demographic history and selection events. We have demonstrated in the Section I, study 3 that coalescent trees can be reliably re-constructed based on sequencing data. The coalescent trees essentially hold the maximum information one can achieve from the genetic data about the past. The well-defined statistical models of coalescent suggest that tests can be designed with moderate ease to examine all kinds of hypotheses, such as admixture, ancient admixture, migration and disease evolution, *etc*. We will try to explore these questions based on our new framework of coalescent reconstruction.

Direction 2: The Genetic Bases and Evolution of Human Facial Morphological Diversity

Our current studies indicate that methods of high-dimensional data decomposition are critical to the successful detection of complex and subtle facial shape differences. Therefore we would put strong efforts into the development of new methods suitable for the high dense face data analyses. On the other hand, if given more funding and resources, we would like to carry out large scale GWAS studies on the common variations of facial morphology in Han Chinese. Furthermore, the study of population structure and divergence on face can be extended to include more global populations. This would give further evidence for selection of facial shapes across the human populations. One particularly intriguing question is whether sexual selection has strongly shaped facial shape divergence. We would try to explore this question with genetic and probably psychological assays in the future. ◂

LATEST PUBLICATIONS

Published since 2011

- Theunert, C.*; Tang, K.*; Lachmann, M.; Hu, S. and Stoneking, M. (*co-first author) (2012) Inferring the history of population size change from genome-wide SNP data. *Mol Biol Evol*. 29(12): 3653–67.
- Li, J.; Zhang, L.; Zhou, H.; Stoneking, M. & Tang, K. (2011) Global patterns of genetic diversity and signals of natural selection for human ADME genes. *Hum Mol Genet*., 20(3): 528–40.
- Kamberov, YG., et al. (2013) Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell* 152.4: 691–702.
- Guo, J., Mei, X., & Tang, K. (2013) Automatic landmark annotation and dense correspondence registration for 3D human facial images. *BMC Bioinformatics*. 14:232.
- Peng, S.*; Tan, J.*; Hu, S.; Zhou, H.; Guo, J.; Jin, L.; & Tang, K.# (2013) Detecting genetic association of common human facial morphological variation using high density 3D image registration. *PLoS Computational Biology*, 9.12: e1003375. (* co-first author, # corresponding author.)
- Guo J.; Tan, J; Yang Y.; Zhou H.; Hu S.; Xu S.; Jin L.; Stoneking M.; Tang, K.# (# corresponding author). Variation and signatures of selection on the human face. *Journal of Human Evolution*, (accepted)

AWARDS

- 2011: Dr. TANG Kun received the SA-SIBS young scientist award
- 2013: GUO Jing received the 2013 SMBE undergraduate travel award
- 2013: ZHOU Hang received the 2013 SMBE graduate travel award

COOPERATION

- With Prof. Mark Stoneking from Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, *on the study of global evidence of positive selection and global human diversity*
- With Prof. KONG Xiangyin from the Institute of Health, SIBS, CAS, Shanghai, *on the association study of HIV resistance phenotypes*
- With Prof. TANG Hua, Stanford University, *on Genetics of human facial morphology*

EXCHANGING ACTIVITIES BETWEEN THE PARTNERING GROUPS

Promoted by the partner group fund, there have been frequent exchanging visits between the two partnering groups, namely the group led by Dr. TANG Kun (TK group) from Shanghai and the one by Prof. Mark Stoneking (MS group) from Leipzig, as briefly summarized below:

<u>From TK group to MS group:</u>

- (Aug & Dec 2011) Dr. TANG Kun visited MS group on the collaboration projects of demographic inference and fine atlas of natural selection;
- (Sep–Dec 2011) Ph.D. students HU Sile and ZHOU Hang visited MS group on the exchange program with the topic of natural selection and demographic inference from human genome sequence data;
- (Jun 2012) Dr. TANG Kun visited MS group on the collaboration project of genetic basis of human facial variations;
- (Jan–Mar 2013) Ph.D. student QIAN Wei visited MS group on the exchange program with the topic of recent co-evolution and population structure analyses.

<u>From MS group to TK group:</u>

- (Aug. 2011) Prof. Mark Stoneking visited TK group on the collaboration project of signals of natural selection in human genome, and gave a seminar talk;
- (Sep. 2011) Ph.D. student Christoph Theunert visited TK group on the collaboration project of demographic inference using IBD/IBS;
- (Oct. 2012) Ph.D. student Christoph Theunert visited TK group on the collaboration project of demographic inference using MCMC technology.