

## Characteristics of dental morphology in the Xinjiang Uyghurs and correlation with the EDARV370A variant

TAN JingZe<sup>1,2†</sup>, PENG QianQian<sup>3†</sup>, LI JinXi<sup>3</sup>, GUAN YaQun<sup>4</sup>, ZHANG LiPing<sup>4</sup>,  
JIAO Yi<sup>4</sup>, YANG YaJun<sup>1,2</sup>, WANG SiJia<sup>3\*</sup> & JIN Li<sup>1,2,3\*</sup>

<sup>1</sup>MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institute of Biomedical Sciences, Fudan University, Shanghai 200433, China;

<sup>2</sup>China Medical City Institute of Health Sciences, Taizhou 225300, China;

<sup>3</sup>CAS-MPG Partner Institute for Computational Biology, Shanghai Institute for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China;

<sup>4</sup>Department of Biochemistry, Preclinical Medicine College, Xinjiang Medical University, Urumqi 830011, China

Received September 10, 2013; accepted March 5, 2014; published online April 17, 2014

Teeth are one of the most important materials for anthropological studies because they are likely to be preserved in ancient remains. While the frequencies of dental characteristics can provide clues to the phylogeny of populations, genetic studies at the individual level can further reveal the biological mechanisms and evolutionary context of dental characteristics. In this study, by analyzing 38 dental characteristics of 242 Xinjiang Uyghur individuals, we found that (i) the dental characteristics of the Uyghurs showed evidence of admixture between European and East Asian populations. The admixture proportions were in line with those previously reported in population genetic studies; (ii) the Xinjiang Uyghur dental characteristics formed three clusters in pairwise correlation analysis. One of the main clusters consisted of characteristics including incisor shoveling, double shoveling and mesial ridge; and (iii) all the characteristics in this cluster were significantly correlated with the genetic variant EDARV370A. The extracted composite phenotypic factor was also significantly associated with EDARV370A, which explained 18% of the total phenotypic variance. This indicated a pleiotropic effect, i.e., the same genetic factor affects a number of dental characteristics at the same time. Our results confirmed that EDARV370A, a genetic variant that first originated in East Asia about 30000 years ago, played an important role in incisor shoveling in East Asia. This finding suggested that incisor shoveling in modern humans in East Asia is likely to have appeared after the late Pleistocene.

### Xinjiang Uyghur, dental characteristics EDARV370A, incisor shoveling, association

**Citation:** Tan JZ, Peng QQ, Li JX, Guan YQ, Zhang LP, Jiao Y, Yang YJ, Wang SJ, Jin L. Characteristics of dental morphology in the Xinjiang Uyghur and correlation with the EDARV370A variant. *Sci China Life Sci*, 2014, 57: 510–518, doi: 10.1007/s11427-014-4654-x

Teeth are some of the best-preserved materials in ancient samples, and therefore are of great value in anthropological studies [1–4]. There is great diversity in dental morphology amongst different human populations. Such diversity is closely related to the origin and evolution of the populations. Using the dental characteristics of East Asian populations,

the American scholar Turner divided them into two types: Sundadonty and Sinodonty. These two types represent the main characteristics of Southeast Asian populations and Northeast Asian populations, respectively [5–7]. In general, the main differences between the Sundadonty and Sinodonty types are that the Sundadonty type is simple, primitive and generalized, while the Sinodonty type is strengthened, complex and specialized. In Southeast Asian populations, the prevalence of mandibular second molar cusp 4 is

†Contributed equally to this work

\*Corresponding author (email: lijin.fudan@gmail.com; wangsijia@picb.ac.cn)

high, while those of incisor shoveling and incisor double shoveling are low. In Northeast Asian populations, pronounced high-frequency characteristics include maxillary incisor shoveling, incisor double shoveling, incisor winging, incisor interruption grooves, incisor tuberculum dentale, and third molar degradation (narrow/shape/missing), as well as mandibular molar deflecting wrinkles and protostylid. These characteristics clearly show the trend of complexity and specialization in Sinodonty [5–7]. Interestingly, these dental morphological characteristics are much less sensitive to environmental factors than morphometric features, and are therefore considered to be influenced to a large extent by genetic factors [5–7].

By studying dental morphological characteristics at the population level (e.g., the frequencies of dental morphological characteristics among different populations), we can reveal the evolutionary relationships of the populations. However, population studies cannot unearth the genetic factors of dental characteristics. For example, which gene(s) actually contributes to the major dental characteristic of incisor shoveling in the Chinese population? Genetic studies are different from traditional anthropology studies. The former can probe into differences among both individuals and populations, while the latter only pays attention to population differences. Therefore, studying the mechanism and evolutionary process of incisor shoveling and related dental characteristics at the individual level will provide important evidence for the evolution of dental morphology in East Asian populations.

EDARV370A (T1540C, rs3827760) is a missense mutation in ectodysplasin A receptor gene that occurred 30000 years ago [8]. Kamberov and colleagues discovered that this mutation was related to multiple morphological characteristics related to ectodermal derivatives, including incisor shoveling [8]. Kimura and colleagues found that EDARV370A was associated with incisor shoveling in the Japanese and Korean populations [9,10]. However, the diversity of EDARV370A is very low in both East Asian and European populations. The ancestral allele 370V mutated to the derived allele 370A in East Asia about 30000 years ago [8]. The derived 370A then underwent strong positive selection in East Asian populations, and eventually its allele frequency reached up to 95% in the Han Chinese population. Conversely, the derived 370A is almost absent in European populations [11]. This unusual frequency distribution means that the power of association studies regarding EDARV370A tends to be insufficient in either East Asian or European populations. It is, however, appropriate to conduct association studies on EDARV370A in Eurasian admixed populations, as the allele frequencies of both ancestral 370V and derived 370A should be moderate in Eurasian admixed populations. In this work, we studied the dental morphological characteristics of Xinjiang Uyghurs, a typical Eurasian admixed population. By systematically exploring the correlations among multiple dental morpho-

logical characteristics and the effect of EDARV370A, we are able to shed light on the genetic mechanism and the evolution of incisor shoveling in East Asians.

## 1 Materials and methods

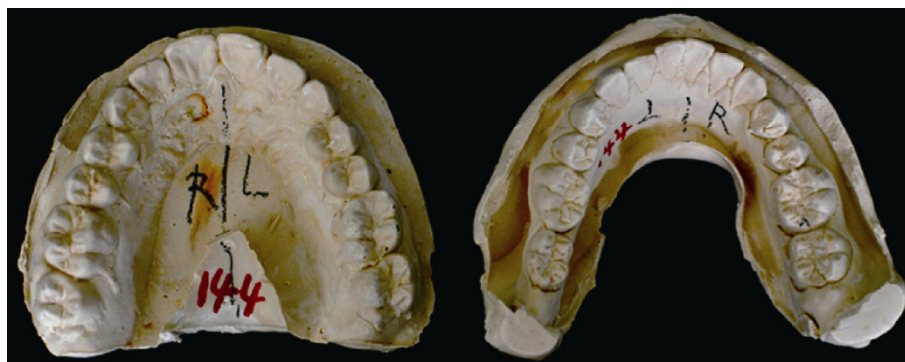
### 1.1 Samples

We studied 242 Uyghur students from Xinjiang Medical University. There were 93 males and 149 females in the sample, with a gender ratio of 1:1.6. The ages of the students ranged from 18 to 25 years old. We carried out questionnaire, dental casts and blood samples collection from the students. Family information for each participant was recorded in the questionnaire, including the student, two parents and four grandparents. Only students whose registered family members were all Uyghurs living in Xinjiang from birth were included in the sample. The research was conducted with the official written approval (written form) of the Ethics Committee of Fudan University, Shanghai, China. All the participants provided informed consent.

### 1.2 Observation of dental morphology

We used dental plaster models to observe the dental morphology of the sample participants. Dental plaster modeling is an established method for dental morphological observation, and has been used in several previous studies, e.g., Kimura et al. [10], Kanazawa et al. [12], Park et al. [9] and Kamberov et al. [8]. We made a dental plaster model for each participant. Example model pictures are shown in Figure 1. Observations of dental morphology used the criteria from the Arizona State University Dental Anthropology System [13,14]. We collected 132 phenotypes describing more than 28 dental characteristics. As the participants in our study were aged 18–25 years, tooth wear was less of a problem and did not affect our observations of dental morphology. On the few occasions when there were quality problems with the dental plaster models and we were unsure of the phenotype scores, we assigned missing data as the score.

We performed quality control on the data from dental morphological observation before data analysis. First, variables were deleted when the missing rate was greater than 20%. In the final dataset, the mean missing rate is less than 5%. Then we assigned dummy variables for winging, interruption grooves, cusp number and groove. We obtained 38 variables from dental morphological observation (Table 1). Finally, we used the individual count method to arrange the data [15]. The criteria for this method were (i) a dental characteristic was observed in bilateral teeth; (ii) if bilateral dental characteristics were not equal, one side was “positive” while the other side was “negative”, and the dental characteristic was recorded as positive, or if the present degrees of bilateral teeth were different then the higher degree



**Figure 1** Dental plaster model of a participant in our sample.

**Table 1** Recording method for dental morphological characteristics

Dental morphology characteristics	Variables	Tooth position <sup>a)</sup>	Record method <sup>b)</sup>
Winging	WUI	UI1	+(appear)/+, - <sup>c)</sup>
Shoveling	SUI	UI1-2	3-7/0-7 <sup>d)</sup>
Double shoveling	DSUI	UI1-2	2-6/0-6
Interruption grooves	IGUI	UI1-2	+(appear)/+, -
Tuberculumdentale	TDUI(C)	UI1-2, UC	1-6/0-6
Mesial ridge	CMRUC	UC	1-3/0-3
Distal accessory ridge	CDARUC	UC	2-5/0-5
Odontomes	OUP(M)	UP1-2, UM1	+(appear)/+, -
Hypocone	C4UM	UM1-3	2-5/0-5
Cusp 5	C5UM	UM1-3	1-5/0-5
Carabelli's trait	CCUM	UM1-3	2-7/0-7
Parastyle	PUM	UM1-3	1-5/0-5
Congenital absence	MSUM	UM3	-(missing)/+, -
Reduced/Peg	PSUM	UM3	+(reduced/peg shaped)/+, -
Lingual cusp variation	PVLP	LP1-2	1-9/0-9
Odontomes	OLP(M)	LP1-2, LM1	+(appear)/+, -
Cusp number	CNLM	LM1-3	3-7
Cusp 4	C4LM	LM1-3	+(appear)/+, -
Hypoconulid	C5LM	LM1-3	1-5/0-5
Cusp 6	C6LM	LM1-3	1-5/0-5
Cusp 7	C7LM	LM1-3	1-4/0-4
Protostylid	PLM	LM1-3	1-7/0-7
Deflecting wrinkle	DWLM	LM1-3	2-3/0-3
Y-Groove	GPYLM	LM1-3	+(appear)/+, -
+Groove	GP+LM	LM1-3	+(appear)/+, -
X-Groove	GPXLM	LM1-3	+(appear)/+, -
Congenital absence	MSLM	LM3	-(missing)/+, -
Reduced/Peg	PSLM	LM3	+(reduced/peg shaped)/+, -

a) UI1, maxillary central incisor; UI2, maxillary lateral incisor; UC, cuspid; U(L)P1-2, maxillary and mandible premolar 1-2; U(L)M1-3, maxillary and mandible molar 1-3. b) Prevalence grade of a characteristic/total grades. c) +, appear; -, disappear; + of Winging includes single, double and reverse types. d) Prevalence grades include 0-7, and 3-7 record as appear.

was recorded; (iii) if an individual only had teeth on one side, then the dental characteristics of that side were recorded; and (iv) when bilateral dental characteristics were equal, then either side was recorded. This method has advantages in individual genetic analysis. When bilateral dental characteristics were not equal, the "positive" side or the higher degree side was most related to the individual gene-

tic background. Additionally, the maximum frequency of dental characteristics in the population was recorded [15]. Our analysis found that there was no significant difference between male and female dental morphology, so we analyzed the data together. We also found that there was no significant difference between bilateral teeth, so we merged them and analyzed them together.

### 1.3 Genotyping

DNA was extracted from peripheral blood for each participant in this study. We used the DNA extraction kit GEN-Eray™ to extract DNA from peripheral blood. We then applied the SNaPshot technique to genotype the single nucleotide polymorphism EDARV370A (rs3827760) located at the exon region of the EDAR gene on chromosome 2.

### 1.4 Statistical analysis

First, we summarized the dental morphological characteristics after quality control and arrangement using the individual count method. We also summarized the distribution of dental morphological characteristics in the Xinjiang Uyghur population. We then performed an association study between EDARV370A and the dental morphological characteristics using a general linear model (GLM). The genotype was assigned to different genetic models: the additive model, the dominant model and the recessive model [16]. Taking the additive model as an example, it was assumed that derived allele 370A has an additive effect on dental morphological characteristics. The general linear model under the additive model was

$$y_i = \alpha + \beta_{ADD}G_i + \varepsilon_i, \quad (1)$$

where  $y_i$  indicates a dental morphological characteristic of an individual,  $\alpha$  indicates a fixed effect on a dental morphological characteristic,  $\beta_{ADD}$  indicates the effect of the genotype on a dental morphological characteristic,  $\varepsilon_i$  indicates the residual of the model and  $G_i$  indicates the genotype of the individual. The genotype was assigned as

$$\begin{cases} G_i = 0, & 0 \text{ copy } 370A, \\ G_i = 1, & 1 \text{ copy } 370A, \\ G_i = 2, & 2 \text{ copy } 370A. \end{cases} \quad (2)$$

To explore the relationships between multiple dental characteristics, we applied pairwise correlation analysis on dental morphological characteristics. Based on the correlation, the constitution of composite phenotypes was obtained. Then we applied the partial least square path model (PLSPM) to extract composite phenotypes from the dental morphological characteristics. We applied the bootstrap method to test the path coefficients in PLSPM [17,18]. Finally we obtained the optimal PLSPM that best fitted the data. PLSPM is a structural equation model based on variance maximization. This method can not only extract composite phenotypes from multiple complicated manifest variables (dental morphological characteristics in this study), but can also estimate path coefficients among composite phenotypes and manifest variables. In this study, we applied PLSPM to extract composite phenotypes from dental characteristics and examined the relation between EDARV370A and these composite phenotypes.

## 2 Results

### 2.1 Frequencies of dental characteristics in the Xinjiang Uyghur population

In this sample of the Xinjiang Uyghur population, the frequency of maxillary central incisor shoveling was 38.82%. This frequency is between that of European populations (2.2%–2.7%) [14] and East Asian populations (80.11%) [8]. This admixture estimation was in accordance with previous genome-wide results [19–21].

The prevalence of maxillary central incisor double shoveling, mandibular second molar Y-grooves, mandibular first molar deflecting wrinkles and maxillary central winging in the Xinjiang Uyghur population was similar to those in East Asian populations. The prevalence of mandibular second molar cusp 4 and first molar cusp 6 was similar to that in European populations (Table 2). Additionally, the prevalence of some dental characteristics was higher than that in East Asian and European populations, such as maxillary first molar Carabelli's trait and cuspid mesial ridge. The prevalence of Carabelli's trait in Bali, East Malay Archipelago and Chaoxian populations was high (about 50%) [12]. The prevalence of maxillary cuspid mesial ridge in West African, South African and Koisian populations was high [14].

In addition, we conducted principle component analysis (PCA) on the prevalence data of dental morphological characteristics in the Xinjiang Uyghur and Taizhou Han populations together with data published on African, European, and South and East Asian populations [12,14]. Eleven dental morphological characteristics of these populations were used in PCA, including maxillary central incisor shoveling (SUI1), incisor double shoveling (DSUI1), cuspid mesial ridge (CMRUC), Carabelli's trait (CCUM1), first molar cusp 5 (C5UM1), mandible second molar cusp 4 (C4LM2), second molar Y-groove (GPYLM2), first molar cusp 6 (C6LM1), first molar cusp 7 (C7LM1), first molar deflecting wrinkle (DWLM1), and maxillary and mandible first and second odontomes (OU-LP1-2). The PCA results showed that the dental morphology of the Xinjiang Uyghur population was between those of the European and East Asian populations (Figure 2). This indicated that the dental morphology in the Xinjiang Uyghur population shows admixture of European and Asian populations.

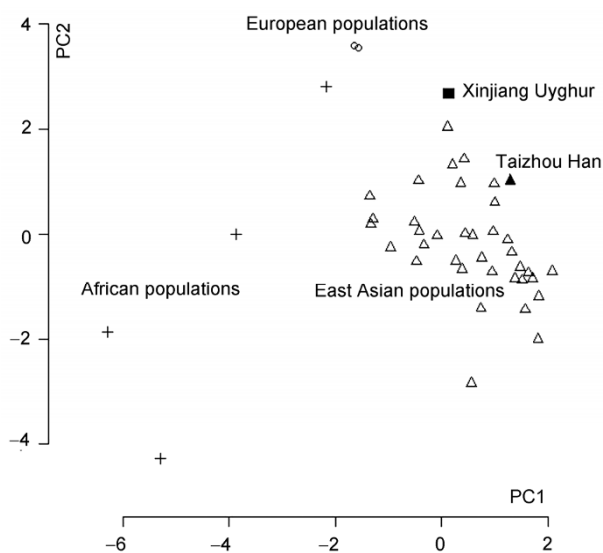
### 2.2 Pairwise correlation results of the dental morphological characteristics

The results of pairwise correlation analysis of dental morphological characteristics in the Xinjiang Uyghur population are shown in Figure 3. We found three major clusters; each cluster represented dental morphological characteristics that are highly correlated. The correlation coefficients in each cluster were generally greater than 0.4, while those

**Table 2** Frequencies of dental morphological characteristics in the Xinjiang Uyghur, European and East Asian populations

Dental morphology characteristics <sup>a)</sup>	West Europe <sup>[14]</sup>	North Europe <sup>[14]</sup>	China-Mongolia <sup>[14]</sup>	Japan <sup>[14]</sup>	Han <sup>[8]</sup>	Uyghur
SUI1	2.70%	2.20%	72.00%	66.00%	80.11%	38.82%
DSUI1	3.80%	5.00%	28.80%	19.50%	29.22%	21.90%
CMRUC	4.30%	0.00%	2.80%	3.00%	5.53%	10.33%
C5UM1	11.80%	26.40%	24.20%	19.70%	11.31%	7.17%
CCUM1	27.30%	18.10%	16.20%	14.90%	20.91%	66.53%
C4LM1	7.80%	10.00%	0.20%	0.30%	2.72%	5.43%
C4LM2	71.10%	84.40%	20.80%	13.60%	36.01%	70.39%
C6LM1	8.30%	16.90%	35.90%	42.70%	18.70%	8.14%
C7LM1	4.50%	5.00%	7.90%	5.70%	4.27%	8.26%
GPYLM2	27.20%	21.00%	7.60%	13.10%	5.43%	7.30%
OU-LP1-2	0.80%	0.00%	5.50%	5.00%	0.70%	0.83%
IGUI2	42.00%	30.00%	43.00%	44.50%	21.43%	31.76%
DWLM1	5.20%	16.00%	15.70%	14.90%	15.17%	14.67%
WUI1	7.20%	4.70%	24.50%	21.90%	31.96%	23.43%

a) The dental morphology characteristics have been observed in all the populations.



**Figure 2** Results of principle component analysis on the dental morphological characteristics of the Xinjiang Uyghur and other populations from around the world (the dataset is attached as Table S1 in Supporting Information).

outside the clusters were less than 0.4. The first cluster consisted of maxillary incisor shoveling (SUI1 and SUI2), incisor double shoveling (DSUI1 and DSUI2) and Carabelli's trait (CMRUC). The second cluster consisted of mandible second molar cusp number (CNLM2), cusp 4 (C4LM2), cusp 5 (C5LM2) and cusp 6 (C6LM2). The third cluster consisted of mandible first molar cusp number (CNLM1), cusp 4 (C4LM1), cusp 5 (C5LM1) and cusp 6 (C6LM1).

### 2.3 Association between EDARV370A and dental morphological characteristics

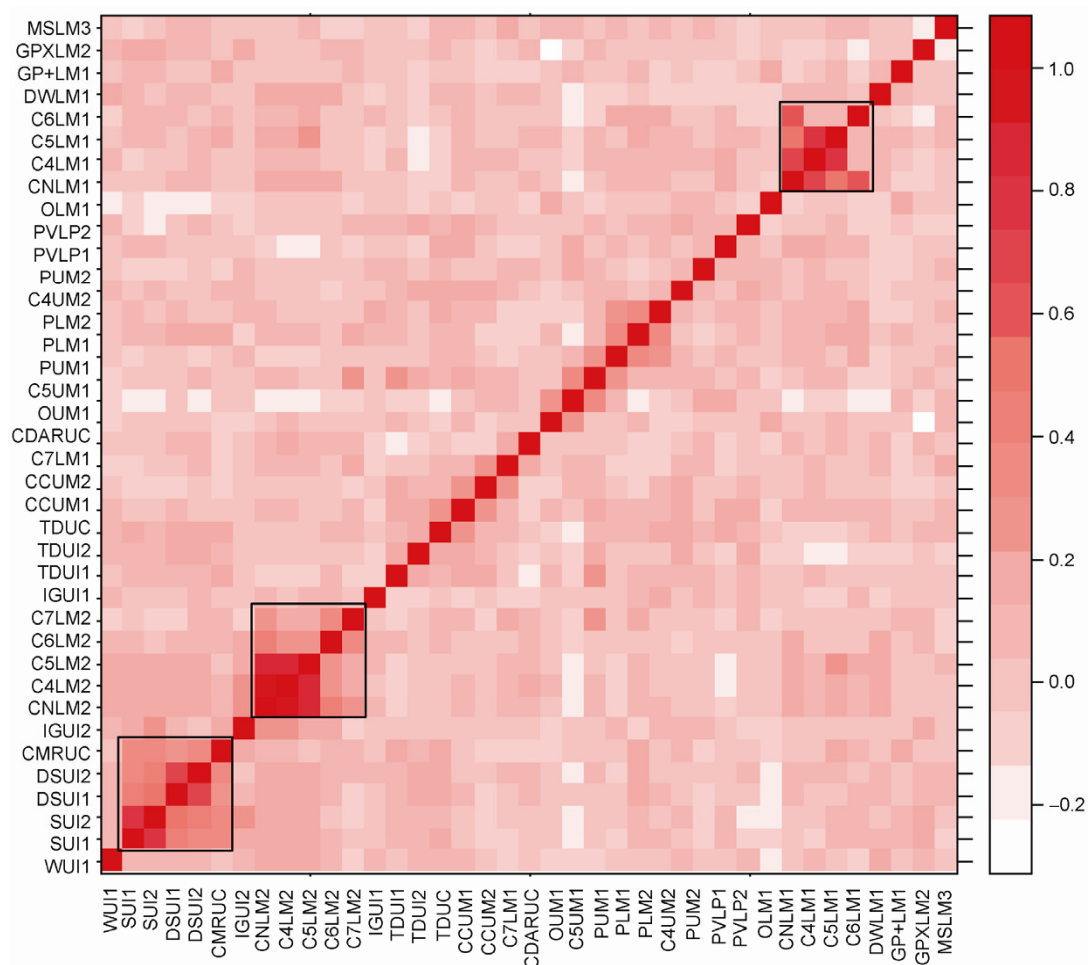
Finally, we successfully obtained the genotype results for

240 subjects. EDARV370A agreed with the Hardy-Weinberg equilibrium. The frequency of the ancestral allele 370V was 37.50%. The frequency of the derived allele 370A was 62.50%. The frequency of ancestral homozygotes 370V was 12.92%. The frequency of heterozygotes was 49.17%, and the frequency of derived homozygotes 370A was 37.92%.

Our results showed that EDARV370A was significantly associated with maxillary incisor shoveling (SUI1, SUI2), incisor double shoveling (DSUI1) and mandible second molar cusp 5 (C5LM2) in the Xinjiang Uyghur sample (Table 3). The effect of EDARV370A on maxillary incisor shoveling was even more significant under the additive model. In the GLM results (Table 3), the regression coefficients of EDARV370A on SUI1, SUI2, DSUI1 and C5LM2 were 0.76 ( $P$ -value= $2.28 \times 10^{-9}$ ), 0.89 ( $P$ -value= $1.15 \times 10^{-12}$ ), 0.49 ( $P$ -value= $1.75 \times 10^{-8}$ ) and 0.51 ( $P$ -value=0.0004), respectively. These results were still significant after Bonferroni correction ( $\alpha=0.001$ ). The influence of EDARV370A on CMRUC was significant (GLM regression coefficient 0.29,  $P$ -value=0.003) but not significant after Bonferroni correlation. The results of the association study for other genetic models (the dominant model and the recessive model) were similar to those of the additive model (Table S2 in Supporting Information).

### 2.4 PLSPM results

We applied PLSPM to the Xinjiang Uyghur dental morphological characteristics and extracted three composite phenotypes (Figure 4). The first composite phenotype (L1) reflected the dental morphological characteristics of maxillary incisor shoveling, incisor double shoveling and cuspid Carabelli's trait. The second composite phenotype (L2) re-



**Figure 3** Correlation among dental morphological characteristics in the Xinjiang Uyghur population. The figure shows the pairwise correlation coefficients of 38 dental morphological characteristics in the Xinjiang Uyghur population. The correlation coefficient is larger where the color is deeper (for example, the diagonal in the figure, shown in dark red, represents the correlation coefficient of the characteristic itself, which is always one). In each of the three clusters shown in black boxes, the dental characteristics are highly correlated with each other, and the correlation coefficients are generally larger than 0.4. Outside the black boxes, the correlation coefficients are less than 0.4.

flected the dental morphological characteristics of mandible second molar cusp number, cusp 4, cusp 5 and cusp 6. The third composite phenotype reflected the dental morphological characteristics of mandible first molar cusp number, cusp 4, cusp 5, cusp 6 and odontomes. The loading coefficients of each composite phenotype on their manifest variables (dental morphological characteristics) were very large (Figure 4). As shown in Figure 4, the loading coefficients of composite phenotype L1 on SUI1, SUI2, DSUI1, DSUI2 and CMRUC were larger than 0.5. We also applied the bootstrap method to test the relation between EDARV370A and the three composite phenotypes. We found that the relation between EDARV370A and composite phenotype L1 was significant (path coefficient 0.43). EDARV370A contributed more than 18% (bootstrap mean percentage 19%) of the variation in composite phenotype L1. There was no significant relation between EDARV370A and the other two composite phenotypes.

### 3 Discussion

#### 3.1 Dental morphological characteristics in the Xinjiang Uyghur population

The Xinjiang Uyghur population is a classical admixed population from European and East Asian populations. Previous studies showed that the two ancestral populations contributed almost equally to the Xinjiang Uyghur population in terms of genetics, and this equal contribution is similar at the individual level [19–21]. Such admixture in Uyghurs causes higher diversity in both genetic loci and physical characteristics than in the ancestral populations, which in turn brings higher statistical power in association analyses. Taking EDARV370A as an example, the frequencies of ancestral allele 370V and derived allele 370A in the Xinjiang Uyghur were 37.50% and 62.50%, respectively, and the frequencies of ancestral homozygotes, heterozygotes and derived homozygotes were 12.92%, 49.17% and

**Table 3** Results of the association study between EDARV370A and dental morphological characteristics in the Xinjiang Uyghur sample

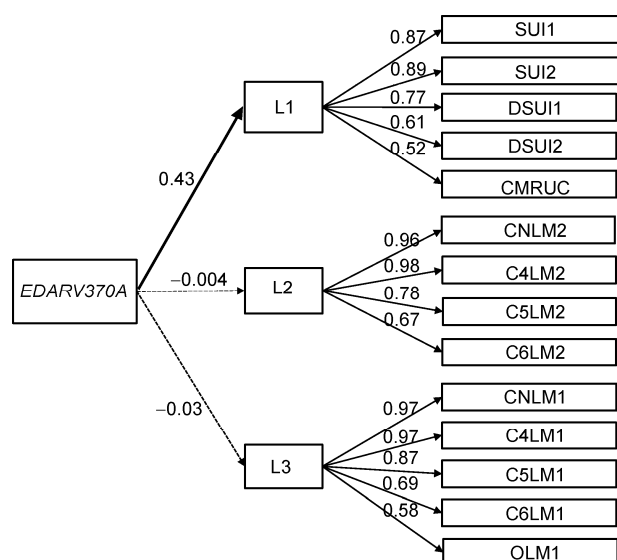
Dental morphology characteristic	Coefficient (SE)	<i>t</i> value	<i>P</i> -value	<i>R</i> <sup>2</sup>
WUI1	0.10±0.04	2.47	0.01	0.03
SUI1	0.76±0.12	6.22	2.28×10 <sup>-9</sup>	0.14
SUI2	0.89±0.12	7.52	1.15×10 <sup>-12</sup>	0.19
DSUI1	0.49±0.08	5.85	1.75×10 <sup>-8</sup>	0.13
DSUI2	0.23±0.08	2.75	0.01	0.03
CMRUC	0.29±0.10	3.00	3.00×10 <sup>-3</sup>	0.04
IGUI2	0.09±0.05	2.01	0.05	0.02
CNLM2	0.14±0.05	2.77	0.01	0.03
C4LM2	0.11±0.04	2.34	0.02	0.02
C5LM2	0.51±0.14	3.61	4.00×10 <sup>-4</sup>	0.05
C6LM2	0.03±0.03	0.95	0.34	0.00
C7LM2	0.00±0.03	-0.05	0.96	0.00
IGUI1	-0.02±0.02	-1.08	0.28	0.01
TDUI1	0.00±0.08	-0.01	0.99	0.00
TDUI2	-0.08±0.11	-0.72	0.47	0.00
TDUC	0.24±0.11	2.19	0.03	0.02
CCUM1	-0.05±0.04	-1.02	0.31	0.00
CCUM2	-0.04±0.04	-1.08	0.28	0.01
C7LM1	-0.15±0.09	-1.59	0.11	0.01
CDARUC	-0.05±0.04	-1.21	0.23	0.01
OUM1	0.00±0.02	0.20	0.84	0.00
C5UM1	-0.08±0.05	-1.46	0.14	0.01
PUM1	-0.01±0.04	-0.23	0.82	0.00
PLM1	0.26±0.12	2.21	0.03	0.02
PLM2	0.05±0.06	0.86	0.39	0.00
C4UM2	-0.02±0.05	-0.50	0.62	0.00
PUM2	-0.12±0.06	-1.82	0.07	0.01
PVLP1	0.02±0.04	0.50	0.62	0.00
PVLP2	-0.07±0.05	-1.39	0.17	0.01
OLM1	-0.06±0.04	-1.48	0.14	0.01
CNLM1	0.07±0.04	2.04	0.04	0.02
C4LM1	0.03±0.02	1.26	0.21	0.01
C5LM1	0.10±0.12	0.81	0.42	0.00
C6LM1	0.11±0.05	2.19	0.03	0.02
DWLM1	0.04±0.04	1.25	0.21	0.01
GP+LM1	0.04±0.06	0.77	0.44	0.00
GPXLM2	-0.06±0.05	-1.13	0.26	0.01
MSLM3	-0.05±0.05	-0.88	0.38	0.00

37.92%, respectively. However, in the Han Chinese population, the frequencies of 370V and 370A were 15% and 85%, respectively, and the frequencies of ancestral homozygotes, heterozygotes and derived homozygotes were 0.49%, 29.53% and 69.98%, respectively. The prevalence of incisor shoveling in Xinjiang Uyghur population was 38.82%, while that in Han Chinese population was up to 80.11%. Based on these data, we found that to achieve the same statistical power, the sample size in the Xinjiang Uyghur population and the Han Chinese population was around 1:24 (calculated using the Quanto program, <http://hydra.usc.edu/gxe>). Obviously, the Xinjiang Uyghur population is more suitable for an association study on EDARV370A and den-

tal morphological characteristics. Compared with the results of Kimura et al. [10], in which the sample size was similar to our study, the results of our study were more reliable.

### 3.2 Use of composite phenotypes in the association analysis

Our results showed different levels of correlations among dental morphological characteristics, indicating that these characteristics were not independent traits. There could be an underlying factor or factors, which may be genetic or environmental, influencing multiple dental morphological characteristics. We applied PLSPM to extract composite



**Figure 4** PLSPM results of EDARV370A and dental morphological characteristics in the Xinjiang Uyghur sample.

phenotypes from multiple dental morphological characteristics, and to estimate the relation between EDARV370A and composite phenotypes. Each composite phenotype reflected several dental morphological characteristics that may be influenced by the same factor. The association analysis for composite phenotypes was more powerful, because it reduced random effects compared with single phenotype association. In this study, we found that composite phenotype L1 mainly reflected the dental morphological characteristics SUI1, SUI2, DSUI1, DSUI2 and CMRUC. The effect of EDARV370A on composite phenotype L1 was significant (path coefficient 0.43,  $R^2=0.18$ ). From the single phenotype association, the contribution of EDARV370A to each dental morphological characteristic was SUI1: 14.18%, SUI2: 19.47%, DSUI1: 13.13%, DSUI2: 3.30%, CMRUC: 4.00%. The average contribution of EDARV370A to these dental morphological characteristics was only 10.82% for the single phenotype association. However, the contribution of EDARV370A to these dental morphological characteristics was up to 18%, as we found in the PLSPM results. Obviously, EDARV370A simultaneously influenced multiple dental morphological characteristics, such as SUI1, SUI2, DSUI1, DSUI2 and CMRUC.

### 3.3 EDARV370A and incisor shoveling

EDARV370A is a mutation that occurred 30000 years ago [8]. Previous studies in the Japanese population reported that EDARV370A contributed 18.9% of the variation in incisor shoveling [9,10]. In this study, we found that EDARV370A contributed 18%–19% of the variation in incisor shoveling (composite phenotype L1). In the Japa-

nese studies, the analysis was focused only on single dental traits, and 18.9% was the contribution of EDARV370A to the single trait of incisor shoveling. However, it is unclear whether the highly correlated dental morphological characteristics were simultaneously influenced by the same factor. In our study, we applied PLSPM to extract composite phenotypes. Each composite phenotype was composed of highly related dental morphological characteristics that might be influenced by the same factor. We used bootstrapping to test the association between EDARV370A and composite phenotypes. We found that the effect of EDARV370A on composite phenotype L1, composed of SUI1, SUI2, DSUI1, DSUI2 and CMRUC, was significant. This indicated that EDARV370A contributed to the development of SUI1, SUI2, DSUI1, DSUI2 and CMRUC. Incisor shoveling was an important characteristic in studying the origin and evolution of modern East Asians. It has been taken as crucial evidence for the continuous evolution theory in East Asia for a long time [1]. However, our results confirmed that EDARV370A, a genetic variant that first originated in East Asia about 30000 years ago, is an important gene causing incisor shoveling in East Asia. This finding suggests that incisor shoveling in modern humans in East Asia is likely to have appeared after the late Pleistocene.

*This work was supported by the National Natural Science Foundation of China (31271338, 31071102, 31071096, 31260263, 91331108, 31322030), the National Key Technology Research and Development Program (2011BAI09B00), the National High Technology Research and Development Program of China (2012AA021802), the Ministry of Health (201002007), the Key Project of Ministry of Education (311016), and the Philosophy and Social Science Foundation of Shanghai (2010BZH005).*

- 1 Wu XZ. The evolution of humankind in China. *Acta Anthropol Sin*, 1990, 9: 312–321
- 2 Wu XZ. Origin of modern humans of China viewed from cranio-dental characteristics of late *Homo sapiens* in China. *Acta Anthropol Sin*, 1998, 17: 276–282
- 3 Liu W, Zhu H. Non-metric traits of human teeth from Miaozigou neolithic site. *Acta Anthropol Sin*, 1995, 14: 8–20
- 4 Liu W. The dental morphology of neolithic humans in north China and its relationship with modern origin. *Acta Anthropol Sin*, 1995, 14: 360–380
- 5 Turner CG. Late Pleistocene and Holocene population history of East Asia based on dental variation. *Am J Phys Anthropol*, 1987, 73: 305–321
- 6 Turner CG. Teeth and prehistory in Asia. *Sci Am*, 1989, 260: 88–91, 94–86
- 7 Turner CG. Major features of Sundadonty and Sinodonty, including suggestions about East Asian microevolution, population history, and late Pleistocene relationships with Australian aboriginals. *Am J Phys Anthropol*, 1990, 82: 295–317
- 8 Kamberov YG, Wang S, Tan J, Gerbault P, Wark A, Tan L, Yang Y, Li S, Tang K, Chen H, Powell A, Itan Y, Fuller D, Lohmueller J, Mao J, Schachar A, Paymer M, Hostetter E, Byrne E, Burnett M, McMahon AP, Thomas MG, Lieberman DE, Jin L, Tabin CJ, Morgan BA, Sabeti PC. Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell*, 2013, 152: 691–702
- 9 Park JH, Yamaguchi T, Watanabe C, Kawaguchi A, Haneji K, Takeda M, Kim YI, Tomoyasu Y, Watanabe M, Oota H, Hanihara T,



- Ishida H, Maki K, Park SB, Kimura R. Effects of an Asian-specific nonsynonymous EDAR variant on multiple dental traits. *J Hum Genet*, 2012, 57: 508–514
- 10 Kimura R, Yamaguchi T, Takeda M, Kondo O, Toma T, Haneji K, Hanihara T, Matsukusa H, Kawamura S, Maki K, Osawa M, Ishida H, Oota H. A common variation in EDAR is a genetic determinant of shovel-shaped incisors. *Am J Hum Genet*, 2009, 85: 528–535
- 11 Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Sun W, Wang H, Wang Y, Xiong X, Xu L, Wayne MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermizakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Sham PC, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Johnson TA, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Yakub I, Birren BW, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Saeki K, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 2007, 449: 913–918
- 12 Kanazawa E, Satake T, Sasaki K, Matsuno M, Igarashi Y, Nakayama M, Roshan P, Takahashi M, Ito Y, Ito K, Guo DL, Ji XP, Ma J. Dental traits in five Chinese minorities in Yunnan Province. *Commun Contemp Anthropol*, 2009, 3: 77–84
- 13 Turner CG, Nichol CR, Scott GR. Scoring procedures for key morphological traits of the permanent dentition: the Arizona State University dental anthropology system. In: Kelley M, Larsen C, eds. *Advances in Dental Anthropology*. New York: Wiley-Liss, 1991. 13–32
- 14 Scott GR, Turner CG. *The Anthropology of Modern Human Teeth*. Cambridge: Cambridge University Press, 1997
- 15 Turner CG. Expression count: a method for calculating morphological dental trait frequencies by using adjustable weighting coefficients with standard ranked scales. *Am J Phys Anthropol*, 1985, 68: 263–267
- 16 Zhu J. *Genetic Model Analysis Methods (in Chinese)*. Beijing: China Agriculture Press, 1996
- 17 Haenlein M, Kaplan AM. A beginner's guide to partial least squares analysis. *Underst Statist*, 2004, 3: 283–297
- 18 Tenenhaus M, Vinzi VE, Chatelin Y-M, Lauro C. PLS path modeling. *Comput Statist Data Anal*, 2005, 48: 159–205
- 19 Xu S, Huang W, Qian J, Jin L. Analysis of genomic admixture in Uyghur and its implication in mapping strategy. *Am J Hum Genet*, 2008, 82: 883–894
- 20 Xu S, Jin L. A genome-wide analysis of admixture in Uyghurs and a high-density admixture map for disease-gene discovery. *Am J Hum Genet*, 2008, 83: 322–336
- 21 Xu S, Jin W, Jin L. Haplotype-sharing analysis showing Uyghurs are unlikely genetic donors. *Mol Biol Evol*, 2009, 26: 2197–2206

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Supporting Information

**Table S1** Frequencies of 10 dental morphological characteristics in the Xinjiang Uyghur and other populations in the world

**Table S2** Association results for EDARV370A and dental morphological characteristics for the dominant and recessive models

The supporting information is available online at [life.scichina.com](http://life.scichina.com) and [link.springer.com](http://link.springer.com). The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.