

SHORT COMMUNICATION

Present Y chromosomes reveal the ancestry of Emperor CAO Cao of 1800 years ago

Chuanhao Wang¹, Shi Yan^{1,2}, Zheng Hou¹, Wenqing Fu¹, Momiao Xiong¹, Sheng Han¹, Li Jin^{1,2} and Hui Li¹

Emperor CAO Cao (155AD–220AD) is one of the most famous persons in Chinese history that had changed the history of East Asia. He claimed to be a descendant of Marquis CAO Can and therefore was of aristocratic ancestry. However, this claim has been suspected for around 1800 years. Here, we collected some present clans with full records of 70–100 generations claimed to be descendants of CAO Cao or CAO Can, and validated them by comparing their Y chromosomes. Haplotype O2-M268 is the only one that is enriched significantly in the Emperor's claimed descendant clans ($P=9.323\times 10^{-5}$, odds ratio=12.72) and, therefore, is most likely to be that of the Emperor. Moreover, our analysis showed that the Y chromosome haplotype of the Emperor is different from that of the Marquis (Haplotype O3-002611). Therefore, Emperor CAO Cao's claim was not supported by genetic evidence. This study offers a successful showcase of the utility of genetics in studying the ancient history.

Journal of Human Genetics (2012) 57, 216–218; doi:10.1038/jhg.2011.147; published online 22 December 2011

Keywords: China history; genetic history; molecular anthropology; Y chromosome haplogroup

The human Y chromosome can be used to reveal the paternal relationships among historical individuals by studying their present descendants, because the most part of the Y chromosome is passed along the male lineage.¹ Successful cases include the determination of paternity of the descendant of Thomas Jefferson² and the inference of Y chromosome haplotype of Jewish priests.³ When combined with the stemma records, Y chromosome can be used to study the relationships of the individuals in even more ancient history. Recording the stemmas has been a religious tradition of Han Chinese, and some stemmas link the contemporary individuals to their ancestors over 3000 years ago, although their authenticity requires careful validation. To explore the utility of the stemmas for evolution studies, we recently studied the Y chromosomes of the clans who claimed to be descendants of Emperor CAO Cao (155AD–220AD) as an example. Moreover, we examined the relationship of Emperor CAO with his claimed aristocratic ancestry.⁴

Emperor CAO is probably one of the most famous persons in the East Asia, because of the popularity of a novel, *Romance of the Three Kingdoms*. Emperor CAO claimed that his clan was derived from Marquis CAO Can (?–190BC), the second Prime Minister of Han Dynasty, and further to the ancient Dukes of CAO (eleventh century BC to 487 BC), and therefore was of aristocratic ancestry. To be of aristocratic ancestry was of great importance for justifying political privileges during early time of China. As Emperor CAO's grandfather was a eunuch, the Emperor's father should have been adopted from his grandfather's own clan (Marquis CAO's clan) according to the strict patriliate custom.⁵ However, Emperor CAO's opponents alleged that Emperor CAO's father was adopted from beggarmdom and was of humble ancestry. This controversy has lasted for nearly 1800 years.

Here, we typed 100 Y chromosome single-nucleotide polymorphisms (Supplementary Table 1) as listed in the latest Y-chromosome phylogenetic tree⁶ on 280 individuals of 79 Cao clans or clan clusters from different locations throughout the China (Figure 1), and 446 individuals of different clans with other surnames. A clan cluster may consist of several simplex clans if they carry different Y-chromosome haplotypes. Thus, we studied overall 111 simplex clans of CAO (Supplementary Table 2). According to their stemma records, 15 of the CAO clans claimed to be descendants of Emperor CAO. These 15 clans distributed in different provinces and never knew the existence of each other. Their Y chromosomes comprise six haplotypes (Table 1). Only one of these six haplotypes can be Emperor CAO's type. The other haplotypes found in the claimed clans might be introduced by other sources such as adoption, acceding to mother's surname, nonpaternities, and so on.⁷ Here we need to recognize the most probable Emperor CAO's haplotype by examining the haplotype distribution among the clan groups.

We classified the clans into three groups: claimed clans of Emperor CAO, unclaimed CAO clans (or claimed to be other origin) and the general population with other surnames. We pairwise examined the frequency difference of each haplotype among the three groups.^{8,9} There is no significant difference between the unclaimed CAO clans and the general population, indicating that the surname CAO may have multiple founders, similar to other surnames.¹⁰ Thus, these two groups were combined into one (referred to as the reference group hereafter) in the subsequent analysis. Interestingly, haplotype O2-M268 is the only one that is enriched significantly in the claimed clans, meanwhile, rare in the reference group (Fisher's exact test,

¹MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences and Department of History, Fudan University, Shanghai, China and ²CAS-MPG Partner Institute for Computational Biology, SIBS, CAS, Shanghai, China

Correspondence: Professor H Li, School of Life Sciences, Fudan University, Shanghai, China.

E-mail: LiHui.Fudan@gmail.com

Received 17 August 2011; revised 10 November 2011; accepted 14 November 2011; published online 22 December 2011

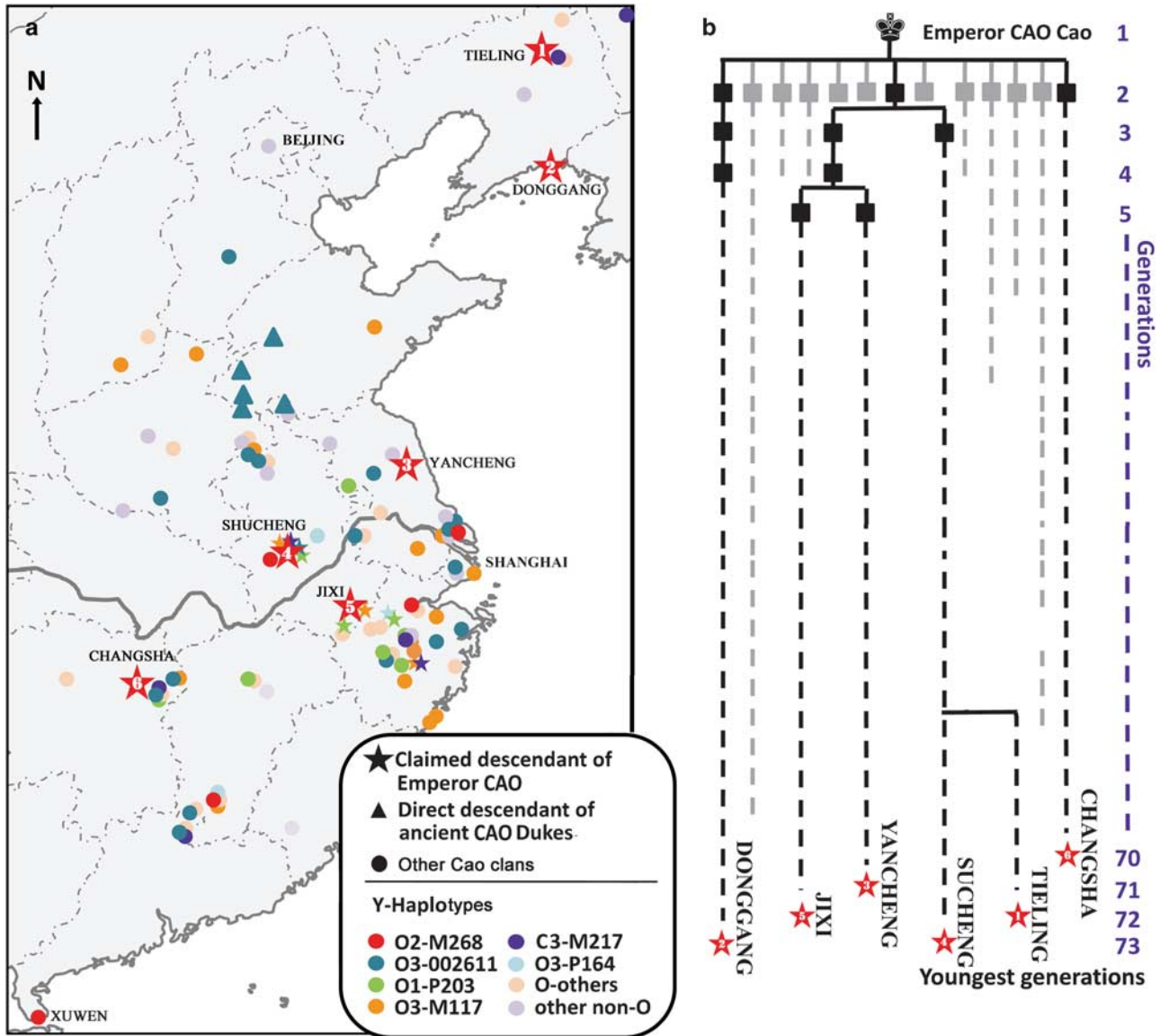


Figure 1 Geographical focuses and Y-chromosome genotypes of the Cao clans (a) and genealogical relationship among the six related clans (b).

Table 1 Fisher's exact tests between different clan groups

Haplotypes	Number of Clans in the clan groups			P-value for pairwise comparisons		OR between claimed group and reference group (95% CI)
	Claimed group	Unclaimed group	General populations	Unclaimed group vs general group ^a	Claimed group vs reference group ^b	
O2-M268	6	5	22	0.801	$9.32 \times 10^{-5*}$	12.72 (4.22–38.32)
O3-002611	1	21	79	0.384	0.952	0.32 (0.04–2.43)
O1-P203	2	6	65	0.030	0.607	1.02 (0.23–4.62)
O3-M117	3	15	67	0.876	0.408	1.40 (0.39–5.08)
C3-M217	2	5	25	1.000	0.211	2.63 (0.57–12.17)
O3-P164	1	2	13	1.000	0.358	2.51 (0.31–20.34)
Others	0	42	175	0.423	NA	0

Abbreviations: CI, confidence interval; OR, odds ratio.

Reference group is combined by unclaimed group and general populations.

^aTwo-tail test.

^bRight-tail test.

* $P < 0.01$.

$P=9.32\times 10^{-5}$), suggesting that it is the most likely haplotype of Emperor CAO (Table 1). The probability of a haplotype to be ancestor's type is not just its frequency in the case group, but need to be estimated from the frequency difference between the case group and the reference group. Here, we used the odds ratio (OR), which provide an estimate (with confident interval) for relationship between two binary variables, to estimate the probability.¹¹ OR between the two groups is 12.72 for haplotype O2-M268 compared with the reference group.

Haplotype O3-002611 reaches the highest frequency among the other CAO clans, and is the only haplotype appears in all five clans who claimed to be direct descendants of the ancient CAO Dukes and CAO Can. Here, different clan groups were classified to examine the probable haplotype of the CAO Dukes: five clans of haplotype O3-002611 and no clans of other haplotypes for the claimed group, 79 clans of O3-002611 and 458 clans of other haplotypes for the reference group. Thus, O3-002611 is most likely the haplotype of the ancient CAO Dukes (Fisher's exact test, $P=1.968\times 10^{-4}$). Taken the two results together, Emperor CAO was unlikely the descendant of CAO Can and the CAO Dukes, and his claim of aristocratic ancestry is therefore not supported by genetic evidence.

Our results, given the data, also established the authenticity of the stemma records of Emperor CAOs' descendants with O2-M268 and those of CAO Cans' descendants with O3-002611. Connected by the ancestors of 70–100 generations spanning over 2000 years, these samples are of great value for studying the Y-chromosome evolution. This study offers a successful showcase of the utility of genetics in studying the ancient history.

ACKNOWLEDGEMENTS

We thank all the volunteers and samplers including Dongdong Kang, Shuyuan Li, Kang Hu, and so on, and some leaders of the Cao clans (Weidong, Youping, Zuyi, Jinwei, Dianhua, Jianhe, and so on) for making this investigation possible. SH and HL were partly supported by the Fudan Social Science Facilitative Funding, Shanghai Commission of Education Research Innovation Key Project (11zz04), and Shanghai Professional Development Funding (2010001). CW was supported by Fudan Innovation Foundation for Graduate Students.

- 1 Sykes, B. & Irven, C. Surnames and the Y chromosome. *Am. J. Hum. Genet.* **66**, 1417–1419 (2000).
- 2 Foster, E. A., Jobling, M. A., Taylor, P. G., Donnelly, P., Knijff de, P., Mieremet, R. *et al.* Jefferson fathered slave's last child. *Nature* **396**, 27–28 (1998).
- 3 Skorecki, K., Skorecki, K., Selig, S., Blazer, S., Bradman, R., Bradman, N. *et al.* Y chromosomes of Jewish priests. *Nature* **385**, 32 (1997).
- 4 Fang, A. *The Chronicle of the Three Kingdoms*. vol. I (Harvard University Press, Cambridge, 1952).
- 5 Han, S. Historical background for the DNA investigation of the CÁO Cáo Clan. *Commun. Contemp. Anthropol.* **4**, 46–52 (2010).
- 6 Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L. & Hammer, M. F. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* **18**, 830–838 (2008).
- 7 King, T. E. & Jobling, M. A. What's in a name? Y chromosomes, surnames and the genetic genealogy revolution. *Trends Genet.* **25**, 351–360 (2009).
- 8 Lehmann, E. L. & Romano, J. P. *Testing Statistical Hypotheses* 3rd edn. (Springer, New York, 2005).
- 9 Weir, B. S. *Genetic Data Analysis Methods for Discrete Population Genetic Data* (Sinauer Associates Inc. Publishers, Sunderland, 1990).
- 10 King, T. E., Ballereau, S. J., Schürer, K. E. & Jobling, M. A. Genetic signatures of coancestry within surnames. *Curr. Biol.* **16**, 384–388 (2006).
- 11 Bland, J. M. & Altman, D. G. The odds ratio. *Brit. Med. J.* **320**, 1468 (2000).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)