



Comment

What's in a name? Comment on “Surname distribution in population genetics and in statistical physics”

Susanna Manrubia

Centro de Astrobiología (INTA-CSIC), Ctra. de Ajalvir km. 4, 28850 Torrejón de Ardoz, Madrid, Spain

Received 18 July 2013; accepted 25 July 2013

Available online 30 July 2013

Communicated by L. Peliti

The practice of naming human individuals is probably as old as spoken language. In ancient, small human groups, the given name may have been enough to distinguish one individual from another and facilitate social gossip—thus social cohesion [1]. As groups grew larger and social organization became more complex, the addition of family names to given names was a convenient way to signal social status or to protect a legacy. The use of family names became prominent in different cultures at different times. However, despite significant differences in the historical acquisition of this habit and in the fidelity of inheritance of family names, the current distribution of surnames presents remarkable similarities all over the world, to the point of having systematically attracted the attention of scientists from quantitative disciplines in the last century and a half. In a comprehensive revision [2], Paolo Rossi goes over the many models, mostly similar in spirit, that have attempted to capture the universal features of surname distributions.

Common patterns in surname abundance arise from a few simple dominating mechanisms (demography, vertical inheritance, and occasional introduction of new surnames) which form the core of most modelling attempts. Nowadays, however, the field is ripe enough to shift attention from the commonalities to the dissimilarities. The abundance of surnames in cultures with a small genealogical depth, for instance, still keeps memory of the historical situation when the inheritance of surnames was adopted. An example is Japan, with a genealogical depth of 4–5 generations (patrilineal inheritance of surnames began in year 1870) that yields a relatively narrow distribution of surname abundances [3], the large number of different surnames notwithstanding (over 290 000). The opposite situation is that of China and other neighbouring countries, which have the largest genealogical depth and a very high fidelity of transmission due to its writing system [4]. As a consequence, there are as few as 7000 different surnames at present, with a distribution of abundances depleted in rare surnames and clearly departing from a power-law. To compensate for this lack of diversity, individuals in the Chinese culture are distinguished thanks to a remarkable variety of given names [5].

The introduction of the spatial structure of surnames abundance speaks for inbreeding habits and reveals migratory phenomena thanks to its relation with the Y chromosome. However, few studies have related patterns of surname abundance and genetic markers. There seems to be a tacit agreement that genetic signals should be more reliable than cultural traits. Still, in groups where their cultural identity has been preserved in the face of foreign contributions,

DOI of original article: <http://dx.doi.org/10.1016/j.plrev.2013.06.005>.

E-mail address: scmanrubia@cab.inta-csic.es.

it cannot be discarded that surnames conserve information on the origin of the population better than its genetic counterpart. An analogous case is that of Madagascar, whose aboriginal language belongs to the Austronesian linguistic family (thus disclosing the origin of its initial settlers) but where the genetic composition of the population is Austronesian and Bantu in equal proportion [6]. The steadily increasing amount of high-quality data on cultural and genetic traits of human groups demand a rise in the efforts to analyse surname patterns with the aim of exploring further questions in fields such as history and anthropology, as acknowledged in [2]. Perhaps the use of an ensemble of surnames to characterize a population might allow the derivation of “surname phylogenies” (in the spirit of linguistic phylogenies) that could be correlated with biological or other independent historical information. Modern human movements could be better quantified in the light of that information, similar to how language phylogenies expose migratory waves [7]. Further, more than one surname can be usually assigned to a given individual, simply using information on some of her previous generations. This might be a way to identify second and higher order contributions to the composition of human groups. As it is happening in genomics and linguistics, this procedure should eventually unveil a network of human relationships beyond the monoparental tree that the use of a single surname implies. There is really much in a name, and we should not miss the opportunity to find it out.

References

- [1] Dunbar R. *Grooming, gossip, and the evolution of language*. Harvard University Press; 1998.
- [2] Rossi P. Surname distribution in population genetics and in statistical physics. *Phys Life Rev* 2013;10(4):395–415 [in this issue].
- [3] Manrubia SC, Zanette DH. At the boundary between biological and cultural evolution: the origin of surname distributions. *J Theor Biol* 2002;216:461–77.
- [4] Manrubia SC, Derrida B, Zanette DH. Genealogy in the era of genomics. *Am Sci* 2003;91:158–65.
- [5] Zanette DH, Manrubia SC. Multiplicative processes in social phenomena. In: Blasius B, Kurths J, Stone L, editors. *Spatiotemporal patterns, networks and synchronization in biological, ecological and social systems*. Lecture notes in complex systems, vol. 7. World Scientific; 2007. p. 129–58.
- [6] Diamond J. *Guns, germs, and steel: the fates of human societies*. W.W. Norton & Company; 1997.
- [7] Gray RD, Drummond AJ, Greenhill SJ. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 2009;323:479–83.