

19. Genealogical and Genetic Ancestry for k Loci

Supervisor: Jotun Hein

The genealogical ancestors to an individual are the individuals that can be reached by a string of operators mother/father to that individual, like motherfatherfather. Within that set are the individuals, that has genetic material that was transmitted to the present individual. Genetic ancestry is in general given more attention as that is a component in genetic analysis and mapping, while ancestors that does not carry ancestral material is irrelevant in this respect. However, it is of principal interest and with the surge of possibility of inference of familial relationships, then it is worth pursuing these questions.

Chang, Derrida, Møhle, Kammerle, Sagitov and others have investigated models of genealogical ancestry. Matsen and Evans (2008 - ME08) investigated a simple model that added genetic ancestry from a single locus to these models and many interesting results were obtained.

ME08 combined a model of choosing parents with a model of choosing ancestors to two alleles within the chosen parents.

ME08 had a series of interesting results. Some of which are illustrated in the figure below. Firstly (their Figure 2), how many alleles do an ancestor (t generations ago) contribute to the present population as a function of how many genealogical descents it has? Secondly, when a most recent common ancestor (MRCA) appears, how many appears with it? ME08 results were very puzzling at first, in that this number had a very non-monotonic relationship to the population size. Thirdly, how are the number of ancestral alleles distributed in the MRCA-generation. ME08 used simulation techniques that easily carries over to a k -locus model. [However, their more technical probability theoretical results might not]

This project proposes to repeat the Matsen-Evans analysis (ME08), but with k loci. Following would be interesting to do:

1 locus model extended investigations: ME08 figure 2 only shows averages and was done for 1 set of parameters. It could be of interest to investigate for different parameters and how varying individual realizations of the process were. In Figure 2 there is a phase transition as predicted by Chang99 when individuals are either the ancestor of everybody or nobody. How does the size of these two groups fluctuate as we go back in time? What is the distribution of the Chang time? [we know its mean.]

k locus model: the addition of a second locus does not change anything that only relates to genealogy or genealogy and a single locus. The interesting questions clearly relate to 2 loci [with or without genealogy].

k -locus model (with r recombination rate between the loci): as k and r becomes large and individuals descendants will also have inherited genetic material and the concepts of genetic and genealogical ancestry should converge. But how quickly dependent on k , r , size of population (n) and number of generations (t)? And the investigation could also be refined by assigning amount to genetic material inherited and also by adding a multiplicity of the genealogical descendency. After Chang-time, what is the distributed of the ancestral material in the present population? How many are genetically empty?

References

- Matsen and Evans (2008) To what extent does genealogical ancestry imply genetic ancestry? *Theoretical Population Biology* 74:182–190
Rohde, D.L.T., Olson, S., Chang, J.T., 2004. Modelling the recent common ancestry of all living humans. *Nature* 431, 562–566.
Manrubia, S., Derrida, B., Zanette, D., 2003. Genealogy in the era of genomics. *Am.Sci.* 91, 158–165.
Chang, J.T., 1999. Recent common ancestors of all present-day individuals. *Adv. Appl. Probab.* 31, 1002–1026. 1027–1038.
Derrida, B., Manrubia, S., Zanette, D., 1999. Statistical properties of genealogical trees. *Phys. Rev. Lett.* 82, 1987–1990.
Derrida, B., Manrubia, S., Zanette, D., 2000a. Distribution of repetitions of ancestors in genealogical trees. *Physica A* 281, 1–16.
Derrida, B., Manrubia, S., Zanette, D., 2000b. On the genealogy of a population of biparental individuals. *J. Theoret. Biol.* 203, 303–315.
Donnelly, P., Wiuf, C., Hein, J., Slatkin, M., Ewens, W.J., Kingman, J.F.C., Chang, J.T., 1999. Discussion: Recent common ancestors of all present-day individuals. *Adv. Appl. Probab.* 31, 1002–1026. 1027–1038.