A Genealogical Look at Shared Ancestry on the X Chromosome

Vince Buffalo[1], Stephen M. Mount[2], Graham Coop[1]

[1]Department of Evolution and Ecology,
Center for Population Biology
UC Davis

[2]Department of Cell Biology and Molecular Genetics,
Center for Bioinformatics and Computational Biology
University of Maryland

April 6, 2016

bioRxiv preprint and blog post

Details of this work are in our bioRxiv preprint: http://bit.ly/x-preprint

We've also written a blog post for a general audience about X chromosome ancestry and genetic genealogy: http://bit.ly/x-ancestry

the blind play of genes

research blog notes book publications about

A Genealogical Look at Shared Ancestry on the X Chromosome

An example of a present-day female's X material being broken up across her X ancestors in her X genealogy back through the generations.

Graham Coop, Steve Mount, and my preprint *A Genealogical Look at Shared Ancestry on the X Chromosome* has been recently been posted on bioRxiv. In the spirit of both outreach and continuing Graham's terrific series of blog posts[1] on genetic genealogy, I'm writing about our paper on X chromosome genealogy and recent ancestry. Before diving into the details of X chromosome ancestry work, I'll review the concepts of genealogies and ancestry. Then, in the next section we'll look at how one's genetic ancestors —the subset of ancestors that you share genetic material with— vary back through the generations. With these concepts reviewed, we'll look at the genealogy that includes all of our *X ancestors*, which due to the special inheritance pattern of the X chromosome is only a subset of one's genealogy. The embedded X genealogy has some properties that impact how segments of DNA are shared between individuals with recent common ancestry (e.g. 6th degree cousins), which we look at through a simple probability model. Finally, we'll look at what we can learn about the relationships of individuals that share sections of their X chromosome due to sharing a recent common ancestor.

[1] For example, see Graham's posts on how many genomic blocks you share with a cousin, how your number of genetic ancestors grows back in time, and how much of your genome is inherited from a particular ancestor

Genealogies

Each human, as a sexually reproducing species with two sexes, has two parents. You have two parents, four grandparents, eight great-grandparents, 16 great-great grandparents, and k generations back have 2^k great$^{(k-2)}$ grandparents, and in general 2^k ancestors k generations back. An example genealogy back five generations is shown below:

A genealogy back five generations. k generations back, one has 2^k ancestors. Circles indicate females, and squares males. The shaded individual is a present-day female.

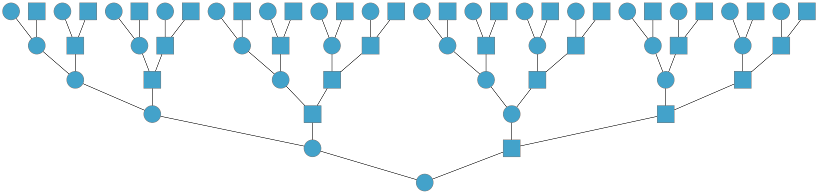
Of course, these genealogical ancestors are not necessarily all *distinct* individuals; as you go further back through the generations, some of these 2^k individuals aren't unique—they're the same person. Intuitively, this occurs when one's two parents are actually related some number of generations back. For example, one's two parents could be 9th degree cousins—e.g. if we assume a generation time of about 30 years, this means these parents shared an ancestor around 270 years ago. This phenomenon is known as *pedigree collapse*, and it's the same thing as inbreeding. The further back through the generations you go back, pedigree collapse *must* happen—it's exceedingly unlikely that 20 generations ago, your 1,048,576 ancestors are all distinct.[2] While pedigree collapse definitely occurs, throughout the rest of this blog post (and in our paper) we ignore it, as we model ancestry that's recent enough where pedigree collapse isn't a large problem.

[2] Some beautiful probability theory by Chang (1999) has shown that the **most recent common ancestor** (the ancestor from which all current individuals in the population descend from) of a population of size N lived $T_N = \log_2(N)$ generations ago. Furthermore, rather amazingly, any individual $1.77 \log_2(N)$ generations ago that has *any* present-day descendents is actually (with very high probability) ancestors of *all* modern day individuals.

Genetic Ancestry

Since each of us have two parents, we receive ½ of our autosomal (i.e. not including the sex chromosomes) genetic material from each parent. We share ½ of our genome with our mother, and ½ with our father. Since your mother shares ½ her genetic material with her two parents, you share ¼ of your genetic material with each grandparent. In general, on average you'll share $½^k$ of your genome with an ancestor k generations in the past. Since the number of crossovers per chromosome is limited, close relatives are likely to share large contiguous segments of their genetic material; a beautiful visualization of this is Morgan's 1916 illustration of crossing over:

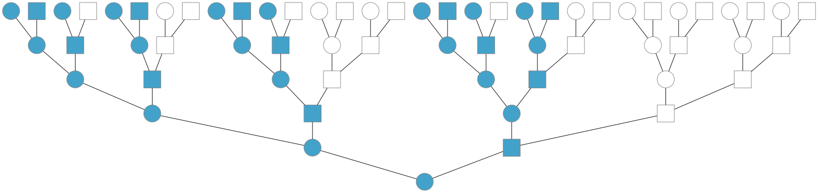
Genealogies



One's **genealogy** contains all biparental relationships back in time of a present-day individual.

These include your two parents, four grandparents, eight great-grandparents, . . . , your 2^k great^{k-2} grandparents, and in general the 2^k ancestors k generations back. These are one's **genealogical** ancestors.

X Genealogies

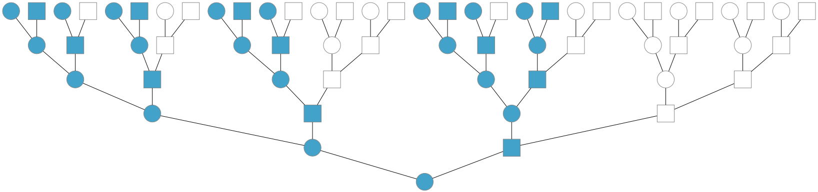


The X chromosome inheritance pattern:

- Every individual receives an X chromosome from his/her mother.
- Every female receives an X from her father (and sons do not receive an X from their fathers[1].).

[1] We call this the *no two adjacent male conditon*, as it means no two males are adjacent in an X genealogy

X Genealogies

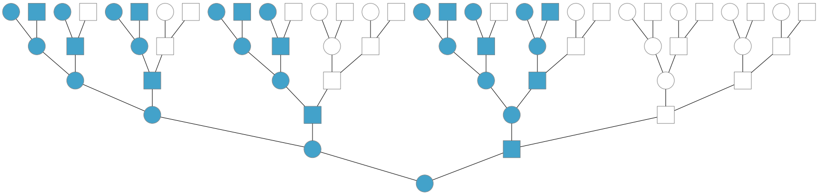


Thus, a present-day female has 2, 3, 5, 8, 18, . . . X ancestors.

Encoding X inheritance rules as a set of recursion equations for the number of females (f_k), males (m_k), and total X ancestors $n_k = f_k + m_k$ for generation k in the past:

$$
\begin{aligned}
f_k &= n_{k-1} && \text{every individual receives an X chromosome from his/her mother} \\
m_k &= f_{k-1} && \text{every female receives an X chromosome from her father}
\end{aligned}
$$

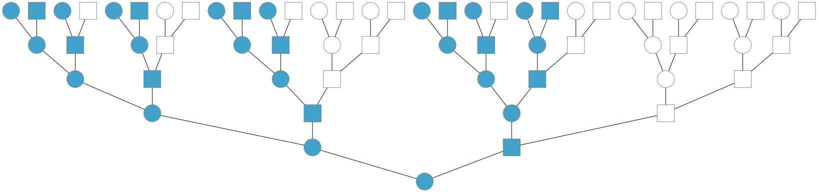
X Genealogies



$$n_k = n_{k-1} + n_{k-2} \tag{1}$$

Which is the famous Fibonacci recurrence. Thus, k generations back a present-day female has $\mathcal{F}_{k+2}$ X ancestors in her X chromosome genealogy.

The X Chromosome



$$\mathbb{E}[N] = \frac{1}{2^d}(\nu d + c) \qquad (2)$$

Acknowledgments

Graham Coop
Steve Mount

Coop Lab