Dating past events using genetic evidence remains one of the holy grails of molecular anthropology (Ho et al. 2014b). Accurate estimation of the relative and absolute timing of past events and processes is crucial if they are to be placed in their appropriate anthropological context, as framed by archeology and linguistics. However, dating is more than just pinning times on events (Ho and Duchêne 2014). Uncertainty in genetic dating can lead to real contention about how we interpret human history. As just one key example, it is still surprisingly unclear when our species first left Africa. The out-of-Africa expansion is generally considered to have occurred around 50 kya (kya = thousand years ago), but mitochondrial DNA (mtDNA) evidence has variously been used to support this date (Underhill and Kivisild 2007) or the much earlier time of 75 kya (Petraglia et al. 2007). This has important implications for how lithic assemblages outside Africa are interpreted. For instance, the Jwalapuram lithics of southern India span the ash layer laid down by the Toba eruption and are therefore securely dated to 74kya. But who made them? Depending on the dating of the out-of-Africa expansion, this lithic assemblage might reflect (i) the first movements of modern humans out of Africa, (ii) an early expansion of modern humans that turned out to be an evolutionary dead end, or (iii) one of the many archaic hominin groups that are now known to have populated Eurasia over much of the last million years. Dates matter, and uncertainty in dating is preventing us from answering fundamental questions about our history.
The role that molecular evidence can play in assigning dates to past events was first identified in the early 1960s with the discovery of the molecular clock. Zuckerkandl and Pauling (1962) proposed that lineages acquire mutations in a regular fashion, allowing sequence divergence to act as a proxy for time. Just five years later, Sarich and Wilson (1967) used this logic to conclude that humans and chimpanzees diverged ~five million years ago, overturning the widespread belief that humans and our sibling species are much more distantly related. This claim was soon backed up by new fossil evidence from East Africa, most notably the *Australopithecus afarensis* specimen “Lucy.” Since then, the development of molecular-dating techniques has flourished, and the approaches used today are vastly more complex than those initial forays. However, many of the simplest questions in the field of molecular dating are still unanswered, and our ability to date molecular events with conviction remains surprisingly constrained.

Mathematics tells us that molecular dating will always be an uncertain science. Ideally, we would like to know how the ages of lineages within populations and species vary across the genome. A branch of mathematics known as coalescent theory has made this possible. One of its most important outcomes was the derivation of equations that describe the mean and variance of the time back to the most recent common ancestor ($T_{\text{MRCA}}$) of a population sample (Hudson 1990; Donnelly and Tavare 1995; Wakeley 2008a).

\[
E(T_{\text{MRCA}}) = 2n \left( 1 - \frac{1}{n} \right)
\]

\[
\text{var}(T_{\text{MRCA}}) = n \left( 8 \sum_{i=2}^{n} \frac{1}{i^2} - 4 \left( 1 - \frac{1}{n} \right)^2 \right)
\]

As the sample size increases, this expectation rapidly converges to 2 and the variance to 1.16 (in units of $N$ generations). The take-home message is that the variance is large relative to the mean, which implies that even if we knew the ages of a set of genes perfectly, they would still vary considerably from each other. In practice, this raises concerns about our past reliance on individual loci, such as mitochondrial DNA or the Y chromosome. Assigning any one molecular date to a historical event has considerable potential to be misleading, which explains why molecular anthropology and, indeed, most subfields of genetics are now turning rapidly to genome-scale data. The entire human complement of DNA comprises thousands of small genetic regions, each representing an independent historical record due to recombination breaking the physical links between loci. With key exceptions, such as studying a particular gene or sex-biased process, looking at thousands of genes simultaneously across the autosomes allows us to “average across” histories at different loci, and thus generate more robust estimates of the dates of influential historical events.

Other evolutionary features also impose strong confounding effects when reconstructing dates from molecular evidence. For instance, both genes and
lineages evolve at different rates (reviewed in Ho 2014). It is perhaps unsurprising
that loci change at variable speeds, as some genetic regions are strongly constrained
(such as the ubiquitin gene, which differs little between mammals and nematodes; Castillo-Davis et al. 2004), while other regions are free to vary. Indeed, the fast
rate of the mtDNA control region, which offers improved resolution over the
time scale of recent human history, is widely perceived as one of the key benefits
of this locus (Endicott et al. 2009). That genes evolve at different rates has been
known since at least the early 1960s, and was suggested by Dickerson (1971) to
reflect variation in the proportion of sites that are free to vary. Such gene-specific
effects demand a careful choice of locus when studying different aspects of human
history, and a variety of loci are employed today precisely to study evolutionary
processes at different time scales.

However, even for the same gene, rates may vary among different human line-
ages. Genes were originally considered to evolve according to a “strict” molecular
clock, whereby all lineages change at some fixed regular rate, much like the ticking
of a metronome. Today, there is greater awareness that rates can vary with gener-
atition time, metabolic rate, and the accuracy of DNA repair mechanisms (among
other drivers). These influences can be pivotal at deeper time depths (such as
comparisons between species; Bromham 2011), but are also important within
species (Endicott and Ho 2008). Most contemporary methods now employ a
“relaxed clock,” in which mutation rates are free to vary among lineages (Heath
and Moore 2014; Yang 2014). Within this framework, different branches in a
gene tree can evolve at different rates. Although there is still considerable
discussion around which relaxed-clock methods perform best (Lepage et al. 2007;
Heath et al. 2012; Ho et al. 2014a), there is a general consensus that relaxed
clocks are preferable to strict-clock alternatives (Drummond et al. 2006), even
when dating recent divergences (Brown and Yang 2010).

Even within a lineage, different parts of a gene may evolve at substantially differ-
ent rates. This has been studied in most detail for human mtDNA: the control
region evolves more quickly than regions that encode genes, and even within genes,
rates vary considerably between first, second, and third codon sites due to the
redundancy of the genetic code (Endicott et al. 2009). Even more nuanced muta-
tion patterns are still being identified in humans. Although the mechanism remains
unknown, mutation events are clustered along the genome and therefore lack
independence, perhaps due to localized failures of DNA repair (reviewed in Ségurel
et al. 2014). Furthermore, mutation rates strongly covary with paternal (but not
maternal) age, primarily due to the marked difference in the number of cell divi-
sions required to produce eggs (an average of 31 cell divisions) versus sperm (~400
cell divisions at age 30 and ~650 at age 40) (Ségurel et al. 2014). Many existing
molecular-dating tools do not readily accommodate these and other similarly com-
plex determinants of mutation rates. As genome sequencing continues apace, it is
becoming increasing clear that our understanding of mutation processes is still in
its infancy, with all that implies for the accuracy required of molecular dating.
A crucial, but often underappreciated, feature of molecular dating is that all methods return only the order and relative ages of events. Real chronological dates only emerge from the process of calibration. At least within molecular anthropology, calibrated dates are most commonly obtained by rescaling a tree with a given mutation rate (Endicott et al. 2009). (This is not necessarily true for other disciplines; see Hipsley and Muller 2014.) Unfortunately, mutation rates are never known with perfect accuracy, which in turn increases the uncertainty of calibrated dates. Three primary methods have been developed to obtain mutation rates: (i) phylogenetic comparisons, (ii) archeologically dated events, and (iii) pedigrees. Each has its benefits and disadvantages, and importantly, the three approaches often do not return the same mutation rate (as discussed in greater detail in this chapter).

One of the most common calibration points used in molecular anthropology is the divergence time between humans and chimpanzees. The mutation rate ($\mu$) is typically calculated as:

$$\mu = \frac{d}{2t}$$

where $d$ is the sequence divergence (i.e. the number of differences per base pair [bp] between humans and chimpanzees for a given genetic region), and $t$ is the divergence time (in years). While sequence divergence can be calculated with high accuracy (Prüfer et al. 2012), the date at which humans and chimpanzees diverged remains contentious. This separation event is often assumed to have occurred six million years ago (Endicott et al. 2009), but molecular estimates support values anywhere from four to eight million years ago (Bradley 2008), while the fossil record places the split at least six million years ago, in part due to a lack of clarity around which fossils are direct human ancestors versus sister taxa (Benton and Donoghue 2007). Phylogenetic rates also typically assume that humans, chimpanzees, and intermediate forms have broadly similar generation intervals (Langergraber et al. 2012) and that the human–chimpanzee split was a relatively discrete event (Patterson et al. 2006; Wakeley 2008b; Presgraves and Yi 2009). Both of these assumptions have been questioned. Nevertheless, the phylogenetic comparison is still among the most widely used methods for calibrating molecular rates for humans and their hominin cousins today.

An alternative approach favors calibration points closer to the present typically employs archeologically dated events, such as the arrival of modern humans in Australia, the settlement of the Americas, or other well-defined migration events (Endicott and Ho 2008; Henn et al. 2009). Because these events are younger, they are often known with more temporal precision than older events, such as human–chimpanzee divergence. Unfortunately, many potential calibration points are also highly contentious in their own right, and therefore often the preferred target rather than basis of molecular dating. For example, early and late dates have been proposed for the settlement of both Australia (O’Connell and
CALIBRATING THE CLOCK

Allen 2004; Hudjashov et al. 2007; Rasmussen et al. 2011) and the Americas (Fagundes et al. 2008; Ho and Endicott 2008; O’Rourke and Raff 2010; Reich et al. 2012). Selecting the wrong age for a calibration point will obviously bias subsequent dating, and at times, archeologically defined mutation rates can introduce hidden circular arguments. However, even if a date is well accepted, determining whether a particular genetic lineage is associated with that historical event can still be problematic. An alternative, and increasingly feasible, approach employs ancient genomes as known calibration points. Sequence divergence between modern and ancient genomes is readily calculated, and archeological remains, particularly those within the radiocarbon limit, are often robustly dated. We even have increasingly extensive genome sequences from a number of our sister species, including Neanderthals (Briggs et al. 2009; Green et al. 2010) and Denisovans (Reich et al. 2010, 2011a; Meyer et al. 2012). This general approach is now used extensively with nonhuman systems (Drummond et al. 2003; Shapiro et al. 2004), and calibrations against ancient human genomes show considerable promise for molecular anthropology as well (Brotherton et al. 2013; Fu et al. 2013b; Rieux et al. 2014).

The third calibration method uses direct estimates of mutation rates from human pedigrees (1000 Genomes Project Consortium 2010; Roach et al. 2010).

\[
\mu = \frac{d}{t_g}
\]

Sequence divergence is determined between two or more generations (such as parent–child pairs), but it must be corrected for the mean generation interval (\(t_g\)), which can vary widely between different human communities (Fenner 2005; Matsumura and Forster 2008). As sequencing costs decrease, this approach is growing in popularity. Although it has the major limitation that sufficient new mutation events must be identified to estimate the mutation rate accurately, this concern has largely dissipated with the advent of genome-scale sequencing. Identifying false-positive and false-negative mutations remains an important consideration, as these can respectively inflate and decrease the observed mutation rate (Conrad et al. 2011). Nevertheless, pedigree methods are rapidly supplanting many earlier approaches of estimating mutation rates.

It is important to note that application of a single mutation rate, regardless of how accurately it is known, cannot account for rate variation among lineages. An alternative approach, which is implemented in some of the more popular software solutions (see Box 4.1), is to instead infer a gene tree and fit that tree to multiple calibration points (Duchêne et al. 2014). As described in this chapter, these calibration points might be archeologically dated events, ancient genomes, or some other well-dated feature. Because many modern dating methods have adopted Bayesian statistics, calibration points are now often provided as tight priors on particular nodes in the gene tree, thereby facilitating uncertainty in the inferred date rather than employing a single point estimate. This approach has recently
Box 4.1

A wide range of tools is now available for molecular dating, many of which target specific research goals. Recent years have seen a noticeable switch from relatively simple dating methods to far more computationally intensive likelihood and Bayesian approaches. Some early methods, such as the $\rho$ statistic (http://www.fluxus-engineering.com/sharenet.htm; Forster et al. 1996), remain appealing due to their simplicity, but unfortunately they produce extremely biased dates under many real-world demographic scenarios (Cox 2008). In contrast, modern likelihood and Bayesian methods explicitly measure uncertainty using standard statistical frameworks and are frequently robust to minor violations of their assumptions. However, the range of demographic models offered by these methods is still often limited, especially models of population structure, and the software is often not trivial to run, typically requiring a high level of familiarity with advanced software and access to substantial computational resources.

Some older methods, such as Genetree (Griffiths and Tavare 1996), still exhibit excellent statistical properties but are limited in application (e.g. the genetic dataset can contain no recombination events or recurrent mutations; Woerner et al. 2007). Other software, notably DAMBE (http://dambe.bio.uottawa.ca/DAMBE/dambe.aspx; Xia 2013) and the “chronos” function of the “ape” package in R (https://cran.r-project.org/web/packages/ape/index.html; Paradis 2013), assign dates to a tree via calibration points at internal nodes, but do not consider uncertainty in the tree topology or branch lengths. Yet other programs, such as RevBayes, provide a general environment for implementing bespoke Bayesian phylogenetics (https://revbayes.github.io). However, the software perhaps most commonly used for molecular dating today is BEAST (http://beast.bio.ed.ac.uk; Drummond and Rambaut 2007), and increasingly its successor BEAST2 (http://beast2.org; Bouckaert et al. 2014). Using Markov chain Monte Carlo (MCMC) and Bayesian statistics, the primary aim of this program is to infer dates for rooted trees under a range of strict and relaxed molecular clock models. The variety of demographic models implemented, including a “skyline” model that allows population sizes to vary freely over time (Drummond et al. 2005), makes BEAST2 one of the most general solutions available for molecular dating today.

Finally, more specialized tools also exist. The IM suite of packages (https://bio.cst.temple.edu/~hey/software; Pinho and Hey 2010) can reconstruct divergence times between pairs of populations, as opposed to dating the divergence of genetic lineages. In most instances, determining the sizes, divergence times, and migration rates of populations is closer to real research goals than obtaining dates for particular genetic lineages on a gene tree.
been used to date (and place geographic paths on) the expansion of the Indo-European language family (Bouckaert et al. 2012), thus illustrating its utility beyond genetics. Multiple calibration points are required for this method to yield any accuracy, but as a subsequent benefit, inferred molecular dates are typically robust to one or a few calibration points that happen to be inaccurate or even simply wrong.

It might be thought that the three calibration methods will return the same mutation rate for a given gene, at least in the limit of large sample sizes, where substantial amounts of information can be brought to bear. Surprisingly, this is not true; systematic differences in estimated mutation rates have been observed using the three different approaches. This discrepancy has been explored in most detail for mtDNA. Considering just the noncoding control region, pedigree studies estimate a mutation rate around $7.95 \times 10^{-7}$ mutations/bp/year ($4.30 \times 10^{-7} - 10.2 \times 10^{-7}$) (Sigurðardóttir et al. 2000; Heyer et al. 2001; Howell et al. 2003; Santos et al. 2005). Conversely, assuming a human–chimpanzee divergence time of six million years (and recalibrating published rates for this value, as in Henn et al. 2009), phylogenetic rates for this same genetic region fall around $2.14 \times 10^{-7}$ mutations/bp/year ($1.15 \times 10^{-7} - 3.60 \times 10^{-7}$) (Hasegawa and Horai 1991; Vigilant et al. 1991; Ward et al. 1991; Stoneking et al. 1992; Tamura and Nei 1993; Horai et al. 1995; Forster et al. 1996). These two estimates are radically different; the pedigree rate is 3.7 times larger than the phylogenetic rate, and the two distributions do not overlap.

This difference has been attributed to the “time dependency” of molecular rates (Ho and Larson 2006). Rates do not simply differ between the pedigree and phylogenetic methods, but have been shown to change regularly through time, exhibiting an exponential decline from fast rates in the present to slow rates in the past (Henn et al. 2009; Ho et al. 2011). This pattern likely reflects the fact that not all changes we observe are neutral. Most genetic regions studied in molecular anthropology are assumed to evolve neutrally, and under neutral theory, the mutation rate is determined only by the rate of spontaneous mutation, independent of population size (Kimura 1968). In practice, however, most regions experience some level of selection: either purifying selection, which removes disadvantageous mutations that arise in the population (Endicott and Ho 2008), or positive selection, which sweeps advantageous mutations (and any mutations physically linked to them; Smith and Haigh 1974) to higher frequency. Most of the mutations that arise within a population are therefore transient and quickly lost through drift or selection. Pedigree studies observe and count these spontaneous mutations, but only a small proportion of them survive long enough to contribute to sequence comparisons at older time depths (such as between humans and chimpanzees). Put differently, pedigree studies measure the spontaneous mutation rate, while phylogenetic studies instead measure the substitution rate (i.e. fixed differences between species). This distinction likely explains the apparent discrepancy between pedigree and phylogenetic rates, at least for mtDNA. Ad hoc methods for
correcting mitochondrial mutation rates for time dependence have been proposed (Soares et al. 2009; Gignoux et al. 2011), but because these methods do not account for other sources of uncertainty (such as rate variation among lineages; Endicott et al. 2009), alternative analytical solutions are actively being explored.

It is worth emphasizing again that mutations are independent of population size only when they are neutral. If selection pressures are acting, changes in population size can also lead to changes in the observed mutation rate. Even for neutral markers, various aspects of population demography can substantially alter the shape of the gene tree, and hence the regularity with which mutations arise and persist in the population (Cox 2008; Henn et al. 2009). Unfortunately, these demographic features – population structure, founder events, and bottlenecks – have all been extremely common throughout human prehistory, thus leveling additional uncertainty on dating estimates.

In the modern genomics era, molecular anthropologists are rapidly turning to genome-wide datasets, and in concert, questions about mutation rates have been moving away from mtDNA and focusing instead on the nuclear chromosomes (Scally and Durbin 2012). Genome-wide data have been available for some time – initially with datasets of short loci distributed across the genome (Wall et al. 2008), and increasingly with genome-wide sequencing of the entire DNA complement of individuals and populations (1000 Genomes Project Consortium 2012). Phylogenetic rates, calibrated by fossil evidence of the divergence between humans and orangutans, or humans and macaques, typically fall around $1 \times 10^{-9}$ mutations/bp/year (Takahata and Satta 1997; Green et al. 2010). Conversely, pedigree studies of dominant disease loci, genome-wide coding regions (“exomes”) and the whole genome suggest a rate around $13 \times 10^{-9}$ mutations/bp/generation ($9.7 \times 10^{-9} – 18.5 \times 10^{-9}$) (Kondrashov 2003; 1000 Genomes Project Consortium 2010; Awadalla et al. 2010; Lynch 2010; Roach et al. 2010; Kong et al. 2012; O’Roak et al. 2012; Sanders et al. 2012). When corrected for a human generation interval of 30 years (Fenner 2005; Matsumura and Forster 2008), this equates to $0.43 \times 10^{-9}$ mutations/bp/year ($0.32 \times 10^{-9} – 0.62 \times 10^{-9}$). Again, a 2.3-fold discrepancy is observed between the two rates, but in striking contrast to mtDNA results, the pedigree rate is lower than the phylogenetic rate. When calibrations are made against a range of primate divergences (in descending temporal order: macaques, orangutans, gorillas, and chimpanzees), a strong trend is observed of lower mutation rates toward the present (Scally and Durbin 2012), a process that has been termed the “hominoid slowdown” (Goodman 1961). Superficially, it appears that mutation rates have decreased through time during the evolution of humans and the great apes. There is still considerable discussion around the veracity of this explanation, and alternatives such as errors in fossil divergence estimates, higher effective population sizes, longer generation intervals, false negatives leading to missing mutations, and the sheer complexity of the mutation process have all been proposed (Ségurel et al. 2014; Veeramah and
Hammer 2014). While the driver of this pattern remains a matter of debate, its discovery highlights how much we have still to learn about mutation processes, even in extremely well-studied organisms like humans.

Molecular dating has also been performed with length-variable microsatellites instead of point mutations (Sun et al. 2009), and recently, new forms of molecular dating have been developed that do not require mutation rates at all. A dominant characteristic of the human past has been rampant admixture between long-separated groups (Cox et al. 2010; Pugach et al. 2013; Hellenthal et al. 2014; Lipson et al. 2014). As these previously separated populations interact and intermarry, they produce offspring that carry chromosomes from each of the parent groups. Recombination events create mosaic chromosomes with ancestry blocks from each parent population, and as recombination events accumulate over time, these ancestry blocks become progressively smaller (Winkler et al. 2010). If the recombination rate is known, the admixture time can be inferred by measuring the distribution of block sizes (Pugach et al. 2011; Xu et al. 2012; Sanderson et al. 2015). At present, this technique is restricted to dating admixture events, but it emphasizes that calculating dates with molecular data requires only that a given process changes in some regular way, and this is not the sole purview of mutation. Still, recombination rates provide no simple panacea to the issues raised here: like mutation rates, recombination rates are increasingly recognized as being variable between species (Jeffreys and Neumann 2009; Laayouni et al. 2011), populations (Laayouni et al. 2011), and individuals (Broman et al. 1998; Kong et al. 2002; Neumann and Jeffreys 2006; Calabrese 2007; Khil and Camerini-Otero 2009). There is also a growing body of evidence showing that recombination rates can change over time (Cox et al. 2013). While new dating methods are always welcome, attempts to date past events using patterns of recombination simply emphasize that most sources of uncertainty in molecular dating are generic and will presumably require broad universal solutions.

Molecular dating has come a long way from the strict molecular clocks and relatively simple tests of the 1960s. Sophisticated analyses are now routine, key sources of uncertainty in molecular dating have been identified, and in many cases, reasonable practical solutions have been developed to address them. Nonetheless, dating past events using molecular evidence is by no means a solved problem. Ongoing research shows that we know less about mutation rates than we once thought; new sources of variation are being discovered, and old sources are being revisited in the light of new genomic information. Uncertainty in the basic foundations of molecular dating, particularly with mutation rates or even whether a mutation rate for a given locus exists in any singular sense, is causing considerable upheaval for our interpretation of major anthropological questions. Many of these fundamentals of molecular dating are areas of active research and are likely to remain so for the foreseeable future.
ACKNOWLEDGMENTS

The author thanks Simon Ho (University of Sydney, Australia) and Phillip Endicott (Musée de l’Homme, France) for their constructive and detailed comments. The author was supported by a Rutherford Fellowship from the Royal Society of New Zealand (RDF-10-MAU-001), the New Zealand Marsden Fund (11-MAU-007), and the Allan Wilson Centre for Molecular Ecology and Evolution.
References


REFERENCES


Bromham, L. 2011. The genome as a life-history character: why rate of molecular evolution varies between mammal species. Philosophical Transactions of the


Burgner, D., S.E. Jamieson, and J.M. Blackwell. 2006. Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better? The Lancet Infectious Diseases 6(10):653–663. doi:10.1016/S1473-3099(06)70601-6


Buss, D. 1985. Human mate selection: opposites are sometimes said to attract, but in fact we are likely to marry someone who is similar to us in almost every variable. American Scientist 73(1):47–51.


Cadavid, L.F., B.E. Mejia, and D.I. Watkins. 1999. MHC class I genes in a new world primate, the cotton-top Tamarin (Saguinus oedipus), have evolved by an active process of loci turnover. Immunogenetics 49:196–205. doi:10.1007/s002510050480


REFERENCES


Ersts, P.J. 2015. [Internet]. Geographic Distance Matrix Generator v. 1.2.3. American Museum of Natural History, Center for Biodiversity and Conservation.


REFERENCES


Fenech, M. 2012. Folate (vitamin B9) and vitamin B12 and their function in the maintenance of nuclear and mitochondrial genome integrity. *Mutation Research* 733(1–2):21–33. doi:10.1016/j.mrfmmm.2011.11.003


REFERENCES


REFERENCES


Haak, W., I. Lazaridis, N. Patterson, et al. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. Nature 522(7555):207–211


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Ishida, Y. 2009. Sewall Wright and Gustave Malécot on isolation by distance. Philosophy of Science 76:784–796. doi:10.1086/605802


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Ortiz, A.M., S.R. Dinapoli, and J.M. Brenchley. 2015. Macrophages are phenotypically and functionally diverse across tissues in simian immunodeficiency virus-infected and
REFERENCES


Overholtzer, L. 2012. Empire and everyday material practices: a household archaeology of Aztec and Spanish imperialism at Xaltocan, Mexico. PhD dissertation, Northwestern University, Chicago, IL.


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


