

## OF MICE AND MEN REVISITED (AGAIN)

DAVID PILBEAM\*

Department of Human Evolutionary Biology, Harvard University, Cambridge MA 02138, U.S.A. pilbeam@fas.harvard.edu

I've known Louis Jacobs for almost fifty years and have learned to greatly value his wit (by which I mean both his wonderfully unique sense of humor and his intelligent astuteness). It has been both pleasure and privilege. This brief contribution is an *Hommage*.

I first met Louis in Pakistan in 1975 at the Khaur Rest-House on the Potwar Plateau. Lou (I still think of him as such) was beginning his doctoral research as part of the Dartmouth-Peshawar research group, which soon began collaborating closely with a Yale-Geological Survey of Pakistan group which I was coordinating. In the following couple of years we began discussing the important role fossils play in calibrating “molecular clocks” (the now widely accepted proposal that genetic differences between species can be used to make reasonable predictions about speciation times). These discussions led to “Of mice and men” (Jacobs and Pilbeam, 1980) in which we urged closer collaboration between paleontologists and molecular systematists. We focused in particular on the important role of the increasingly abundant murid fossil record, particularly as it was then believed to document the divergence of *Mus* and *Rattus*.

When we were discussing and then sometime in 1979 writing the paper, the most common molecular approaches involved immunological comparisons, protein sequencing, and DNA-DNA hybridization. Each served as proxies for the genomic difference (“distance”) between a pair of living species. One contribution was particularly interesting to us because it summarized the genetic distances, based on albumin immunological and DNA-DNA hybridization differences, both between *Mus* and *Rattus* and between *Homo* and *Pan* (Sarich, 1972). For the two proxies, *Mus-Rattus* differences were an order of magnitude greater than for *Homo-Pan*. At that time, most paleoanthropologists accepted a date of 14 Ma or more for the divergence of *Homo* and *Pan*, while Sarich and his colleague Allan Wilson (1967) proposed 5 Ma

for the human chimpanzee divergence. For the *Mus-Rattus* divergence, Sarich (1972) inferred a date of 35 to 40 Ma, rather than the fossil-based 8 to 14 Ma estimated in Jacobs and Pilbeam (1980).

The genetic difference between genomes of two species will be an average of the range of coalescent times (ages of separation) of their different alleles (Edwards and Beerli, 2000), chromosomes (Patterson et al 2006), or arbitrarily-selected DNA sequence “windows (Foley et al, 2023). These times vary across the genome, so this average must be older than the age of the splitting or speciation of the descendant lineages (Patterson et al, 2006, well-explained in their Fig. 1). Over the past six or more decades, there has been considerable progress in understanding molecular evolutionary processes, and gene sequences are now available for many mammal species, along with a range of model approaches to using such data and generating phylogenetic frameworks with age estimates for splitting of ancestral lineages (for example, Foley et al, 2023).

An important paper from the 1970's (King and Wilson, 1975) emphasized that the genetic variation between sister species has two components: variation within their common ancestor before the lineages speciated; and a second portion, differences accumulated along the diverging lineages. This critical insight was developed more formally over the following quarter century (for example, Edwards and Beerli, 2000), emphasizing the difference between genomic divergence time and population divergence in speciation or splitting time. In a useful paper for paleontologists, Steiper and Young (2008) drew attention to the significance of these differences. Determining that a split has occurred requires recovering fossils preserving hard-tissue features with clear apomorphies showing a relationship to at least one of the lineages. Clearly, even the oldest plausible member of a lineage will be younger than split time, how much younger remaining

---

\*Corresponding author

unclear and effectively indeterminate. This splitting time will of course also be younger than genomic divergence time.

The “earliest” acceptable fossil representative of a lineage is used in a “phylogenetic” approach to estimating mutation rates and speciation times. Formally, determining a mutation rate per time interval would require using the averaged genomic difference (distance) as numerator and the genomic divergence time as denominator. But because divergence time is unknown, the (unavoidably) younger population splitting times are used in calculations. The denominator must therefore be reduced and mutation rates over-estimated; they will be least over-estimated if there is reason to believe that the fossil record is a sufficiently reliable predictor of the splitting times used in calculations. Taking a “phylogenetic” approach, Nachman and Crowell (2000) used a range of fossil-based estimates for the splitting time of humans and chimps, along with estimates of ancestral effective population size (the fraction of population reproducing), and estimated a human mutation rate of  $\sim 2.5 \times 10^{-8}$  mutations per nucleotide per generation.

An important recent contribution to estimating mutation rates has been the expanding use of “pedigree” studies in living species. These involve the “trio” method in which genomes of parents and offspring are sequenced to recognize new mutations in offspring, hence determining mutation rates per generation. Accumulating data, especially for primates, show that mutation rates vary across species, with variation based on different factors. The current consensus is that for primates estimated mutation rates are low, on the order of  $\sim 10^{-8}$  per base pair per generation (Chintalapati and Moorjani, 2020). The best available data are for humans:  $\sim 1.2 \times 10^{-8}$  base pairs per generation. Note that the phylogenetic approach of Nachman and Crowell (2000) estimated a human mutation rate twice as great. Using a human generation interval of  $\sim 30$  years, Chintalapati and Moorjani (2020) estimated a human mutation rate of  $0.4$  to  $0.5 \times 10^{-9}$  per base pair per year.

While there is not yet a completely stable estimate for hominoid genomic divergence and populations splitting times, Moorjani et al (2016) and Chintalapati and Moorjani (2020) provide what are the current best estimates for the *Pan-Homo* clade: a *Pan/Homo* divergence date of  $\sim 12.1$  Ma and a splitting time of  $7.9$  Ma (midpoint of estimate range from  $9.3$  to  $6.5$  Ma). Around  $8$  Ma for the split between *Pan* and *Homo* lineages is the best that can currently be offered, not least because the fossil record of early hominins is poor

and contested. Murids present a clear contrast, with an outstanding fossil record.

Using the murid fossil record as documented and interpreted at the time, Jacobs and Pilbeam (1980) had proposed a *Mus-Rattus* divergence time of between  $8$  and  $14$  Ma, based on the assignment of species of *Progonomys* and *Karnimata* respectively to *Mus* and *Rattus* lineages. A recent phylogenetic analysis (Kimura et al, 2015) confirmed the *Progonomys-Mus* relationship, but revised that of *Karnimata* to support a link with *Arvicanthis* with *Rattus* now seen as more distantly related to *Mus* and *Arvicanthis*.

The several decade-long campaign of intense screen-washing by Lou and his colleagues (especially the late, greatly missed, Will Downs) in the fossiliferous Miocene sequence of the Potwar Plateau greatly expanded collections of small mammal fossils (Jacobs and Flynn, 2005, Flynn et al, 2023, Flynn et al, in prep.). Intense stratigraphic surveys and use of paleomagnetic analyses made it possible to place most small-mammal samples within  $100,000$  year-long bins (Flynn et al, 2023, Flynn et al, in prep.).

Sample sizes of murines in particular have steadily expanded, as have analyses by Lou and his students and colleagues: systematic, phylogenetic and functional analyses, along with those of morphometrics and carbon and oxygen stable isotopes of tooth enamel (Kimura et al, 2013a, 2013b, 2015, 2016, Aghova et al, 2018, Flynn et al, 2020, Kimura et al, 2021). For my purpose here the most relevant of these recent analyses for calibrating clocks involves the detailed documentation over several million years of the speciation of the murines *Progonomys* and *Karnimata*, especially the recent work of Dr. Yuri Kimura (Kimura et al, 2013a, 2016, Flynn et al 2020, Kimura et al 2021). The record is abundant with well-dated specimens under excellent stratigraphic control, detailed morphometric analyses making possible plausible systematic and phylogenetic inferences (This material is likely to be discussed in greater detail in other contributions to this Festschrift.).

Good samples of *Potwarmus* and *Antemus* specimens document the evolving stem lineage ancestral to *Progonomys* and *Karnimata*, with good records between  $15$  and  $12.5$  Ma and especially between  $14.2$  and  $12.8$  Ma; this is followed by a sampling gap until the next good murine localities ranging in age between  $11.6$  and  $11.2$  Ma, which specimens present a very interesting pattern (Kimura et al. 2021). The most likely interpretation is that a single species is represented, but one

that is intriguingly variable. First molar patterns show continuous variation but with end-members resembling respectively *Progonomys* and *Karnimata* morphs. Other molars do not show this pattern. Kimura et al, 2021, label this sample “Pre-*Progonomys*, Indeterminate *Progonomys-Karnimata* grade”, reflecting its morphological complexity and intermediacy, and systematic ambiguity. Following the 11.2 Ma level there is another fossil-free interval until the next good samples starting at 10.5 Ma; by this time clear morphological differences of the *Progonomys* and *Karnimata* lineages are present which become even more marked by 9.2 Ma.

How is this interval between 12.5 and 10.5 Ma best interpreted, during which *Progonomys* and *Karnimata* lineages separated and then diverged? What might we be observing? The sequence appears to document an ancestral species and its subsequent split into descendant lineages. The separation of these had definitely happened by 10.5 Ma, but until the preceding 700 ka has been well sampled the best that can be said with confidence is that splitting happened between 11.2 and 10.5 Ma. What might explain the intriguing variation patterns seen in the “Pre-*Progonomys*, Indeterminate *Progonomys-Karnimata* grade” taxon? One plausible scenario is that speciation is being documented.

The first decades of this century saw continuing and expanding interest in the speciation process (Harrison, 2012). A very recent whole-genome analysis of 241 placental mammal genomes (Foley et al, 2023) showed that for many splitting events there was evidence of introgression (gene flow from hybridization) in the early stages of lineage divergence. (The phenomenon is also well-documented in plants as well as animals: an example being a recent study (Zhou et al, 2017) of introgression among Pine subspecies.) Conveniently for us, introgression can also be observed in the house mouse, *Mus musculus*, a well-studied model organism for phylogenetic studies and not just for biomedical research -- and the descendant of the *Progonomys* lineage.

The oldest specimen of the genus *Mus* recorded anywhere is from an 8.0 Ma Potwar locality (Flynn et al, in prep.). *Mus musculus* evolved over the last 3 million years (Lawal et al, 2022), giving rise during the most recent million years (White et al, 2009) to three primary subspecies: (*M. m. domesticus* native to Western Europe, *M. m. musculus* present across Eastern Europe and Siberia, and *M. m. castaneus* across South and Southeast Asia. (Lawal et al 2022). The subspecies are well-supported genetically, but show considerable

introgression among them (Lawal et al 2022, White et al 2009). Of interest, given our necessary hard-tissue paleontological approach, morphological differences in dentitions across the subspecies are present (Darvish, 2008), offering an additional perspective on variation seen in “Pre-*Progonomys*, Indeterminate *Progonomys-Karnimata* grade”.

Hence, one plausible hypothesis, as noted by Kimura et al. (2021) and Flynn et al (in prep.), is that the Potwar murine record between 12 and 11 mya records a species that is in the process of differentiating, subspecific lineages showing some differences in tooth morphology while continuing to introgress before the lineages “emerge” and differentiate as full species (no doubt continuing exchanging genes). It is worth noting that the degree of morphological differences between molars of *Mus musculus* subspecies (Darvish, 2008) are equivalent to those between the *Progonomys* and *Karnimata* morphs observed between 12 and 11 Ma (Flynn, pers comm.) It is unfortunate that there is still too little understanding of the genomic determinants of hard-tissue phenotypic features, or how these might be or might not be involved in the introgressed fraction of genomes.

The history of the *Progonomys* and *Karnimata* lineages provides one of highest quality mammalian fossil records currently available for recognizing and addressing divergence and split times; best estimates are a population separation/splitting time around 11 Ma with a genomic divergence time (averaged across the genome) around 12 Ma.

More such high-quality records are needed.

## ACKNOWLEDGMENT

As a final note, in recognition and thanks to Lou for his contribution to my Festschrift (Jacobs and Flynn, 2005) I echo and analogize the great American philosopher Lawrence Peter Berra: “Always go to other people’s funerals; that way they’ll come to yours.”

## LITERATURE CITED

- Aghová, T., Kimura Y., Bryja, J., Dobigny, G., Granjon, L., & Kergoat, G. J. (2018). Fossils know it best: using a new set of fossil calibrations to improve the temporal phylogenetic framework of murid rodents (Rodentia: muridae). *Molecular Phylogenetics & Evolution*, 128, 98–111. doi: 10.1016/j.ympev.2018.07.017.
- Chintalapati, M. & Moorjani, P. (2020). Evolution of the mutation rate across primates. *Current Opinion in Genetics and Development*, 62, 58–64. <https://doi.org/10.1016/j.gde.2020.05.028>.

- Darvish, J. (2008). Biosystematic approach to geographic variations of house mouse group, *Mus musculus* L.1766. *Iranian Journal of Animal Biosystematics*, 4, 31–54.
- Edwards, S. V. & Beerli, P. (2000). Perspective: gene divergence, population divergence, and the variance in coalescent time in phylogeographic studies. *Evolution*, 54, 1839–1854.
- Foley, N. M., Mason, V. C., Harris, A. J., Bredemeyer, K. R., Damas, J., Lewin, H. A., Eizirik, E., Gatesy, J., Karlsson, E. K., Lindblad-Toh, K., Zoonomia Consortium, Springer, M. S., & Murphy, W. J. (2023). A genomic timescale for placental mammal evolution. *Science*, 380, eabl8189. DOI: 10.1126.
- Flynn, L. J., Kimura, Y., & Jacobs, L. L. (2020). The murine cradle. In G. V. R. Prasad & R. Patnaik (Eds.), *Biological Consequences of Plate Tectonics: New Perspectives on Post-Gondwanaland Break-Up* (pp. 347–362). Springer, Ostende.
- Flynn, L. J., Morgan, M. E., Barry, J. C., Mahmood Raza, S., Cheema, I. U., & Pilbeam, D. (2023). Siwalik rodent assemblages for NOW: Biostratigraphic resolution in the Neogene of South Asia. In I. Casanovas i Vilar, C. Janis, J. Saarinen, & L. van den Hoek (Eds.), *The NOW database of fossil mammals*. Springer, Ostende.
- Flynn, L. J., Jacobs, L. L., Kimura, Y., Lindsay, E. H., Cheema, I. U. (In prep.) Siwalik Glires. In C. Badgley, M. E. Morgan, & D. Pilbeam (Eds.), *At the Foot of the Himalayas: Paleontology and Ecosystem Dynamics of the Siwalik Record of Pakistan*. Johns Hopkins University Press, Baltimore.
- Harrison, R. G. (2012). The language of speciation. *Evolution*, 66, 3643–3657.
- Jacobs, L. L. & Flynn, L. J. (2005). Of mice....again: the Siwalik rodent record, murine distribution, and molecular clocks. In D. E. Lieberman, R. J. Smith, & J. Kelley (Eds.), *Interpreting the Past: Essays on Human, Primate, and Mammal Evolution in Honor of David Pilbeam* (pp. 63–80). American School of Prehistoric Research Monograph Series 5. Boston, MA: Brill Academic Publishers, Inc.
- Jacobs, L. L. & Pilbeam, D. (1980). Of mice and men: fossil-based divergence dates and molecular “clocks”. *Journal of Human Evolution*, 9, 551–555. doi: 10.1016/0047-2484(80)90062-7.
- Kimura, Y., Flynn, L. J., & Jacobs, L. L. (2016). A palaeontological case study for species delimitation in diverging fossil lineages. *Historical Biology*, 28, 189–198. doi: 10.1080/08912963.2015.1022175.
- Kimura, Y., Flynn, L. J., & Jacobs, L. L. (2021). Tempo and Mode: Evidence on a protracted split from a dense fossil record. *Frontiers in Ecology and Evolution*, 9, 642814. doi:10.3389/fevo.2021.642814.
- Kimura, Y., Hawkins, M. T. R., McDonough, M. M., Jacobs, L. L., & Flynn, L. J. (2015). Corrected placement of *Mus-Rattus* fossil calibration forces precision in the molecular tree of rodents. *Scientific Reports*, 5, 14444. doi: 10.1038/srep14444.
- Kimura, Y., Jacobs, L. L., Cerling, T. E., Uno, K. T., Ferguson, K. M., Flynn, L. J., & Patnaik, R. (2013a). Fossil mice and rats show isotopic evidence of niche partitioning and change in dental ecomorphology related to dietary shift in late Miocene of Pakistan. *PLOS One*, 8, e69308. doi: 10.1371/journal.pone.0069308.
- Kimura, Y., Jacobs, L. L., & Flynn, L. J. (2013b). Lineage-specific responses of tooth shape in murine rodents (Murinae, Rodentia) to late Miocene dietary change in the Siwaliks of Pakistan. *PLoS One*, 8, e76070. doi: 10.1371/journal.pone.0076070.
- King, M. C. & Wilson, A. C. (1975). Evolution at two levels in humans and chimpanzees. *Science*, 188, 107–116.
- Lawal, R. A., Mathis, V. L., Barter, M. E., Charette, J. R., Garretson, A., & Dumont, B. L. (2022). Taxonomic assessment of two wild house mouse subspecies using whole-genome sequencing. *Scientific Reports*, 12, 20866 | <https://doi.org/10.1038/s41598-022-25420-x>.
- Moorjani, P., Amorin, C. O. G., Arndt, P. F., & Przeworski, M. (2016). Variation in the molecular clock of primates. *PNAS*, 113, 10607–10612.
- Nachman, M. W. & Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*, 156, 297–304.
- Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S., & Reich, D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature*, 441, 1103–1108. doi:10.1038/nature04789.
- Sarich, V. M. (1972). Generation time and albumin evolution. *Biochemical genetics*, 7, 205–212.
- Sarich, V. M. & Wilson, A. C. (1967). Immunological time scale for hominid evolution. *Science*, 158, 1200–1203.
- Steiper, M. & Young, N. (2008). Timing primate evolution: lessons from the discordance between molecular and paleontological estimates. *Evolutionary Anthropology*, 17, 179–188. <https://doi.org/10.1016/j.gde.2020.05.028>.
- White, M. A., Ané, C., Dewey, C. N., Larget, B. R., & Payseur, B. A. (2009). Fine-Scale phylogenetic discordance across the House Mouse genome. *PLOS Genetics*, 5(11), e1000729.
- Zhou, Y., Duvaux, L., Ren, G., Zhang, L., Savolainen, O., & Liu, J. (2017). Importance of incomplete lineage sorting and introgression in the origin of shared genetic variation between two closely related pines with overlapping distributions. *Heredity*, 118, 211–220. doi:10.1038/hdy.2016.72.