Background-dependent mutation rate and strong deleterious effects of *de novo* transposition of endogenous retrovirus in mouse inbred strains

Fernando Pardo-Manuel de Villena[1], Anwica Kashfeen[2], Timothy A Bell[1], Darla Miller[1], Gary A Churchill[3], Martin T Ferris[1], Leonard McMillan[2]

1, Department of Genetics, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill
2, Department of Computer Science; University of North Carolina at Chapel Hill
3, The Jackson Laboratory

The explosive growth in the number of individuals for which whole genome sequence (WGS) has reinvigorated the field of transposable element (TE) biology. TEs represent a large fraction of the murine genome. Some TE insertions, such as the one at the *agouti* locus, have well known and dramatic phenotypic effects and recent studies demonstrate that the complement of TEs is variable among laboratory strains. However, many fundamental questions remain unanswered, including the mutation rate, the spectrum of functional effects of *de novo* TE and their impact on rigor and reproducibility of mouse studies. We have taken advantage of our unique resources/tools to address these important questions. These resources/tools include a newly developed computational pipeline to identify, map and characterize TEs, WGS of many related individuals within well-defined pedigrees and the ability to test the effect of *de novo* TE insertions in both gene expression, and ultimately on fitness. We conclude that the rate of mutation for a given TE family varies dramatically among inbred strains. Mutation rates are highly consistent among substrains of a given strain-group, such as BALB/c and C57BL/6, independently of their commercial and historical origin. This observation, coupled with the substantial variation in the number of *de novo* TEs observed among CC strains, strongly supports the conclusion that genetic background represents a key factor in mutation rate. In addition, almost half of *de novo* TEs are deleterious as demonstrated by the statistically significant differences in the pattern of insertions (exon, intron and intergenic) relative to ancestral neutral TEs, their rapid purge from populations, the departure from HW equilibrium, the causal effect on expression of nearby genes and the detection of TE/gene chimeric transcripts. Our results indicate that in some inbred strains *de novo* TE can be a substantial source of phenotypic variation and may impact reproducibility of results.