Project 1: The effects of mutation bias, codon bias, and amino acid bias on genome sequences.

Species differ in the nucleotides used across their genomes. Some of these differences are due to divergence in the error-correcting proteins and reflect simple mutational biases that act across the whole genome (also called “GC-bias”). Other differences are manifest as the non-random usage of synonymous codons (“codon bias”) and are due either to mutational biases or to selection for the usage of specific synonymous codons, usually to improve translational efficiency. Finally genomes also differ in the particular amino acids used, again either because of mutational biases or because of selection for the usage of less costly amino acids. But the relationship between these various biases and the contributions of mutation and selection to each of them is not well understood.

The sequencing of multiple genomes of closely related organisms offers the opportunity to begin to disentangle the causes and consequences of biased nucleotide usage. By looking at the correspondence between changes in nucleotide, codon, and amino acid biases across multiple yeast species, as well as the number of tRNA genes present in each genome, we can address the contributions of mutation and selection to each pattern.

Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis.
Hiroshi Akashi, and Takashi Gojobori

Nucleotide Bias Causes a Genomewide Bias in the Amino Acid Composition of Proteins.
Gregory A. C. Singer and Donal A. Hickey
Molecular Biology and Evolution 17:1581-1588 (2000)

Synonymous codon usage is subject to selection in thermophilic bacteria.
David J. Lynn, Gregory A. C. Singer and Donal A. Hickey

Project 2: Regulatory elements responsible for male-biased gene expression.

Approximately 20% of all genes show significant, preferential expression in either males or females. These sex-biased genes largely show this bias because they are expressed in germline cells: the testes of males and the ovaries of females. Specific sequences must drive this tissue-specific expression, yet no generally used sequences are yet known for either testes or ovaries. It is thought that these sequences must be short, as many duplicated genes that move from the X chromosome to an autosome evolve male-biased gene expression quite quickly.
In order to find the sequences involved, we must mesh data from whole genome sequencing with microarray experiments measuring expression in males and females, or in sex-specific tissues. By compiling sets of sex-biased genes in Drosophila and humans, we can begin to look for the sequence common to all of them.

Paucity of Genes on the Drosophila X Chromosome Showing Male-Biased Expression
Michael Parisi, Rachel Nuttall, Daniel Naiman, Gerard Bouffard, James Malley, Justen Andrews, Scott Eastman, Brian Oliver

A gene atlas of the mouse and human protein-encoding transcriptomes.
Andrew I. Su, et al.

Emergence of Young Human Genes after a Burst of Retroposition in Primates.
Marques AC, Dupanloup I, Vinckenbosch N, Reymond A, Kaessmann H