Glossary for *Ancient DNA and Human Evolution*

DNA STRUCTURE & ORGANIZATION

Allele: Alternative variant forms at the same locus.
- **Heterozygotes:** Have two different alleles at a locus.
- **Homozygotes:** Have two identical alleles at a locus.

Chromatin: DNA wrapped around histone proteins.
- **Euchromatin:** Open, allowing information to be read.
- **Heterochromatin:** Tightly wrapped and inactive.

Chromosomes: Discrete strands of packaged DNA.
- **Allosomes:** Chromosomes that determine sex (XY, with Y-Chromosome inherited paternally).
- **Autosomes:** All other non-allosomal chromosomes. Do not differ between the sexes.
- **Mitochondrial DNA (mtDNA):** Maternally inherited DNA found only in the mitochondria.

DNA: The molecule of inheritance, consisting of sequences of the four nucleotide building blocks (ATGC).
- **Sequence:** The linear order of the building blocks, which encodes individual form and function.

Genome: All DNA in a cell. Also refers to the DNA sequence that typifies an individual or species.
- **Genetics:** The study of genes and their inheritance.
- **Genomics:** The study of genome structure/function.

Haplotype: A set of alleles at distinct positions in the genome which are inherited together.
- Individuals in **Haplogroups** share a given haplotype.

Histones: Chief protein components of chromatin and can be chemically modified as part of epigenetics.

Karyotype: Chromosome number in the cell nucleus.
- **Diploid:** Two sets of paired chromosomes.
- **Haploid:** One set of unpaired chromosomes.

Locus (pl. Loci): A unique physical position on a chromosome.
- **Exons:** Sequences at a locus that encode proteins.
- **Introns:** Sequences between exons, don't encode proteins.

Variant: DNA that differs among groups studied.

Recombination: Exchanges between chromosomes that causes independent inheritance of alleles.
- **Linkage Disequilibrium:** Non-random inheritance of alleles at different loci (due to low recombination).

INFORMATION ENCODED IN DNA

Codon: A sequence of three nucleotides along a DNA or RNA chain encoding a single amino acid.

CpG site: Locus where a cytosine nucleotide is followed by a guanine nucleotide in the linear sequence of bases. Cytosines in CpG dinucleotides can be methylated to form 5-methyl cytosine, a common epigentic mark.

Enhancer: Short region of DNA that can be bound by proteins to alter transcription of a gene.

Epigenetic: Information not encoded directly in DNA.

Epigenome: Molecular modifications of the DNA and its associated histone proteins, affecting its function.

Functional DNA: Encodes biological information.
- ~2% of all DNA: Codes for proteins.
- ~80% of all DNA: Regulates gene activity.

Gene: DNA whose information encodes a function.
- **Post-translation Modifications:** Alter mature protein.
- **Transcription:** DNA sequence converted into RNA.
- **Translation:** mRNA converted into a protein sequence.

Gene Regulation: Alterations of gene expression/activity.
- **lncRNA:** Long non-coding RNA.
- **miRNA:** Short non-coding regulatory microRNA.
- **RNA Binding Proteins (RBP):** Proteins that bind RNA.
- **Transcription Factor Proteins:** Alter gene expression by binding directly or indirectly to DNA.

Genotype: The two alleles at one or more diploid loci.

Mutation: Change of a DNA sequence.
- **Indels:** Insertions or deletions of DNA sequence.
- **Single Nucleotide Polymorphisms (SNPs):** Single nucleotide differences (e.g. A vs. T).
- **Silent Mutations:** No change to the phenotype.
- **Synonymous/Non-synonymous Mutations:** No change to the protein; changes to protein, respectively.

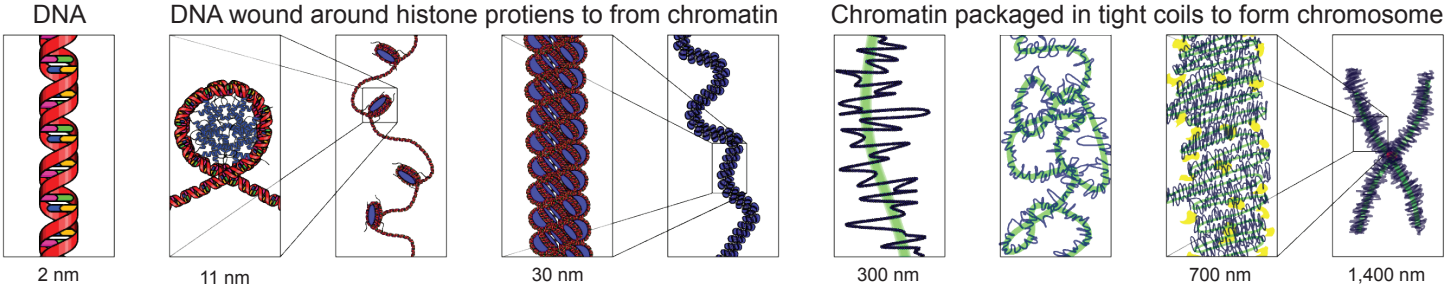
Phenotype: Observable traits of an organism (result from interactions between genes and environment).

Promotor: Region of DNA that initiates transcription of a particular gene.

Transposable Elements (TE): Sequences that replicate in a genome by inserting copies of themselves at other loci (a type of "molecular parasite").

DNA ORGANIZATION AND CHROMOSOME STRUCTURE

Modified from www.wikipedia.org



METHODS OF GENOME ANALYSIS

Alignment: Arranging related sequences by position.

Cloning: Making a copy of an organism or sequence.
Organisms are cloned by moving an entire genome from a cell into an egg. DNA sequences are cloned by moving copies into a bacteria using a vector.

Genome Wide Association Study (GWAS): An approach for "gene mapping" in which hundreds of thousands of SNPs are tested statistically for genetic associations with a phenotype.

Genotyping: Characterizing genetic variants at one or more loci.

CRISPR: A method that can mutate a specified locus.

PCR: A method of copying a specified locus.

Sequencing: Reading the order of nucleotides in DNA.
Coverage: The number of reads for a given locus.
Shotgun: Sequencing cuts the genome into short chunks that are read and reassembled by a computer.

Vector: DNA molecule used to direct the replication of a cloned DNA fragment ("insert") in a host cell.

EVOLUTION

Adaptation: Evolution of a phenotype by selection because it improved reproduction and/or survival.
Coalescence: Time since common ancestor.
Coalescent Theory: Models evolution backward in time to infer historical population size, mutation rate, allele age, and allele frequency change by selection and drift.

Divergence: Change in genetic content or phenotype between isolated populations or species.

Effective Population Size (N_e): The size of an idealized population (random mating, no selection, mutation or migration) with the same rate of genetic drift as the study population.

Genetic Drift: Loss of alleles by chance.

Homology: Similarity in DNA or phenotype because of shared evolutionary history from a common ancestor.

Homoplasy: Similarity in DNA sequence or phenotype that has evolved independently.

Phylogeny: Historical relationships of species or loci.

Polymorphism: An allelic difference observed in more than 1% of the population studied.
Allele Frequency: The proportion of all alleles within a population that are a particular type.
Derived Alleles: Variants arising since last common ancestor.
Fixed Alleles: Replaced all other alleles in a population.

Population: A defined group of similar individuals.
Demography: Study of population size over time.
Gene Flow: Movement of alleles between populations.

Selection: Allele frequency change over time caused by the different replication rate of specific alleles.

Species: A population whose individuals can mate with one another to produce viable and fertile offspring. (debated definition)

OTHER TERMS

Admixture: Breeding between isolated populations.

Archaic Admixture: DNA from ancient, divergent, and now extinct populations found in current people.

Atapuerca: An archaeological site in Spain with fossils and stone tools of the earliest known hominins in Western Europe.

Denisovans: A population of extinct hominins contemporary with Neandertals.
Our knowledge of Denisovan morphology is based on two small fossils (a finger bone and a molar) found in the Eurasian Steppe.

Dental Calculus: Calcified dental plaque, provides information on diet, disease, health, microbiome and protects the genetic information within the tooth from degradation.

Homo: The genus that comprises the species *Homo sapiens*, which includes modern humans, as well as several extinct species classified as ancestral to or closely related to modern humans.

***Homo erectus*:** An extinct species of hominin with fossil evidence dating from 1.9 million (possibly earlier) to 70 thousand years ago and found from Africa to Indonesia.
May have been the first hominin to leave Africa.
H. erectus DNA may be retrievable from other species due to archaic admixture.

Introgression: Transfer of alleles between species.

Middle Pleistocene: A period of geological time (781-126,000 years ago). An important time for the diversification of hominins, including the emergence of Neandertals and *Homo sapiens*.

Morphology: Shape or form (outward appearance) of an organism.

Neandertals: An extinct Eurasian hominin species.
Neandertals existed from over 500,000 to 30,000 years ago, and hybridized with ancient humans.

This ancient DNA glossary is the product of the Anthropogeny Graduate Specialization students, CARTA faculty and staff, and Stevan Springer at UC San Diego.