Jargon Dictionary

Zane Billings

EPID 8200 assignment

2024-05-06

* Metadata: used in this paper to refer to non-genomic data associated with a sequence, different from how I would normally use this word.
* Ontology: a formalization for concepts that establishes names and relationships and allows for standard language about a topic.
* Pipeline: a sequence of code/analysis steps which can be executed in sequence (data flows between modules on the pipeline).
* Auditability: ensuring data can be validated to some standard in addition to being accessible.
* Interoperability: the ability to exchange information and operate across systems.
* Deleterious: reduces fitness
* Fidelity: ability of a polymerase to exactly replicate a nucleic acid sequence.
* Pleiotropic: has multiple different effects / affects many things
* Endogenous: from inside of an organism (e.g. endogenous to the host rather than made by the virus). Endo inside, genous origin
* Nucleobase: nitrogenous base compound, A(T/U)CG
* Nucleoside: sugar + nucleobase
* Nucleotide: phosphorylated nucleoside
* Mutation: localized alternation of a nucleic acid residue.
* Residue: a single unit of a polymer.
* Silent or synonymous mutation: those that do not give rise to an amino acid substitution despite being located in a protein-coding region of the genome (change one codon to another codon that encodes the same amino acid).
* Neutral theory: evolution of organisms at the molecular level occurs mainly due to random drift.
* Episodic positive selection: transient events of positive (directional) selection due to one or a few AA substitutions in a string of synonymous, tolerated mutations.
* Internal oligoadenylate tract: a bunch of adenoside nucleotides that are not at the end of a polynucleotide (as opposed to a poly-A tail).
* Homopolymeric tract: part of a polymer with different types of units which is all composed of the same unit.
* Polymerase slippage: when there are too many repeats of the same base, the polymerase can slide off the template and stop copying or mess something up.
* Exonuclease activity: enzymatic activity that allows a protein to cleave individual residues off of a nucleic acid polymer.
* RNA replicases: aka RDRP, binds to a SSRNA strand and synthesizes its complementary strand (as opposed to DDRP for transcription).
* Moiety: a part of a molecule.
* APOBEC: apolipoprotein B mRNA editing complex; a protein complex which mutates RNA transcripts.
* ADAR: adenosine deaminase acting on ds-RNA; same thing but different.
* Reverse genetics: method for studying genetics where a gene is altered to examine the consequences on the phenotype.
* Compensatory mutation: beneficial mutations which allow other deleterious mutations to remain in the genome.
* Epistasis: the joint effect of different mutations in the same genome on the same trait.
* Recombination: the formation of a new genome by covalent linkage of genetic material from multiple parent genomes.
* Defective interfering (DI) genomes: virus particles that contain only part of the genome and thus interfere in the normal replicative cycle by generally being in the way.
* Multipartite virus: a virus with a segmented genome where each viral particle only contains part of the total genome; a cell must be coinfected by all parts of the genome simultaneously to produce a new virus.
* Clonal interference: different lineages with different beneficial mutations compete with each other, which can lead to exclusion of one beneficial lineage in viruses without recombination.
* ADE: severe symptoms of a disease are mediated by interaction with antibodies in some way
* EIP: interaction between strength of immune response, kinetics of viral adaptation, and timing of transmission
* Selective sweep: the process by which a beneficial mutation becomes fixed in the population through selective reduction in genetic variation at the locus of that mutation.
* Immunological distance: the extent to which the immune system is able to generate cross-reactive immune responses between two antigens.
* Variant: a mutant pathogen lineage which is qualitatively different from other observed lineages.
* Mechanistic model: a model that uses real-world theory and aims at predicting behavior based on theory, instead of just quantifying statistical correlations. Typically, a differential equation model.
* Homogeneous sampling: all samples are drawn from the same underlying population.
* Representative sampling: the source population of sampling is the same as the target population for inference.
* Internal/external tree branch: an internal branch is not connected to a leaf/end node, but an external branch is.
* THE coalescent model: a model of the process by which two gene sequences could have evolved from a common ancestor.
* Emergence: when a mutant variant of a pathogen spreads and takes hold in the population instead of dying out.
* Coalesce: to join.
* Antigenic drift: the gradual process of mutation that occurs as a virus persists in the population under some amount of selection.