Computational Systems Biology: Biology X

Bud Mishra

Room 1002, 715 Broadway, Courant Institute, NYU, New York, USA

L#4:(October-0-4-2010)
Cancer and Signals

Outline

Outline

Evidence in Favor

- Somatic mutations, Aneuploidy, Copy-number changes and LOH (loss of heterozygosity)
- Two kinds of genes play important roles. **Oncogenes**: Act dominantly and act through gain of function; **TSG (Tumor Suppressor Genes)**: Act recessively and act through loss of function.
- Effect of radiation and chemotherapy. They are assumed to induce additional further genetic instability to trigger apoptosis.
- **Synthetic Lethality**: Target a “paired gene” to dysfunctionalize a tumor cell.
- **Knudsen’s Two Hit Hypothesis**:
  - *Familial vs Sporadic*: Acquired (or non-inherited) mutations are often same as inherited mutations

- Cancer risk in families... *Li-Fraumeni syndrome*: Mutant variety of one copy of the pair of p53 genes is inherited. Increases risk of several forms of cancer... [Breast carcinoma, sarcoma and leukemia]
- Hereditary Colon Cancer:
  - Multiple independent primary tumors
  - Same mutant gene (APC) in colon stem cells, on the other hand, is characterized by monoclonality and slower progression
- Hereditary Mutations: Leukemia by age 3... Retinoblastoma by age 5: Caused by a mutation in RB1 gene in chromosome 13.

Evidence Against

- Not all carcinogens are mutagens
- Incidence of various cancer varies from population to population (even after correction for gender and age)
- Only about 5% of childhood cancer have a clear hereditary basis.
- *Field Effects*: Multiple primary tumors in the same organ, but no familial disposition. Locally recurrent tumors.
- **Debulking and Remission**:

Outline

Population Genetics

- Study of Genetic Basis of Evolution.
- Frequencies and fitness of genotypes in natural population.
- Evolution is the change in the frequencies of genotypes through time, perhaps due to their differences in fitness.
- Fitness of genotypes are determined by the phenotypes.
- Associating genotypes to phenotypes (either in an individual or a population).

Forces in Genotypic Evolution

- **Mutations/Polymorphisms**
- **Genetic Drift**
- **Selection**
- **Migration**

Mendelian Genetics

- Mendelian genetics apply to virtually all sexual organisms: **metazoa** (multicellular animals) as well as **metaphyta** (multicellular plants).
- Mendel's insights:
  1. Genetic information is passed in particular forms from an organism to its offspring. (Genes, genome, epigenome)
  2. Constitution of an organism could be divided into a series of discrete, separable entities. (Traits: Particulate theory)
  3. Each observable trait of an individual might be traceable to a separate gene.
  4. **Genotypes** and **Phenotypes**

Genotypes

- In **diploid** organisms, a gene has two fold redundancy – with the exception of the sex chromosomes.
- Two copies of a gene could convey different, possibly conflicting information.
- Different versions of genes are called **alleles**. A gene is called **biallelic**, if it has primarily two different versions in the population.
- If an organism carries two identical alleles of a gene, it is said to be **homozygous**, If it has two distinct alleles, it is **heterozygous**.

Genotypes

- If a single allele determines the phenotype (irrespective of what the other allele is), then that phenotype is called **dominant**. If, on the other hand, both copies of the gene must have the same allele to determine the phenotype, then it is **recessive**.
- There is a spectrum of phenotypes determined by the control of the two alleles: **incomplete dominance** (two phenotypes blend: flower color); **co-dominance** (both phenotypes are present, blood types)

Mutations

- Genetic information is corruptible. **Mutations** modify the information content of a gene (creating a new allele), regulation of a gene (changing the dosage) or copy-number of a gene (also, changing the dosage). Thus, mutations *may* be able to affect the phenotypes. It also modifies the gene-frequencies within a population.
- An allele that is present in the great majority of the population is termed **wild type** — compatible with normal structure and function.
- The collection of alleles present in the genomes of all members of a species: the **gene pool**.

Selection

- **Selection**: Most mutations are neutral. Otherwise, they are very likely lethal (killing the organism with the mutation) and do not enter the gene pool. If a mutation is advantageous, it spreads in the gene pool through a selective sweep.
- Heterozygous Advantage

Neutral Mutations

- The great bulk of the genetic information in human genomes (generally, in all mammals) is **non-functional** (non-coding, non-regulatory, etc.): **junk DNA**.
- Certain modification in certain base-positions (**synonymous modification**) in a gene will not affect the encoded protein. **Silent mutations**
- Such mutations are **neutral mutations**.

Mutations

- **Single-base substitutions**: *Missense mutations* (altered amino-acid), **Nonsense mutations** (changes the STOP codon), **Silent mutations**, **Splice-site mutations** and **Regulatory-site mutations**. Loss of Heterozygousity (LOH)...
- **Insertions and Deletions**: Frame-shift, Splice-variants...
- **Translocation**
- **Copy-number changes** (Duplication, Hemizygous Deletion, Homozygous Deletions, Loss of Heterozygousity)
- **Somatic vs. Germline Mutations**

Cancer in Population

- Cancer is a disease of the genome and has its origin in mutations occurring at mitosis. There are 10^{13} cells in the body. (The total number of cells during an average human lifetime $\approx 10^{16}$. Amount of cell turn over – involving cell death and replacement – about 10^7 events per second.
- Conclusion: The cancer risk (of each kind) should have similar distributions in all human sub-populations... (and both genders).
- This is not true (except for certain pediatric tumors).
- Role of **heredity** and **environment**

Cancer Mutagens

- Katsusaburo Yamagiwa (1915): Repeated painting of localized areas of the skin of rabbits' ears resulted in carcinoma.
- Bruce Ames (1975): Showed that carcinogens can act as mutagens. Experiments on laboratory mice and rats.
- Theory: Cancer is a disease of mutant genes and that carcinogenic agents induced cancer through their ability to mutate genes.
- Some carcinogens are not mutagens (they promote tumorigenesis through non-genetic mechanisms): **tumor promoters**.

Standard Approach: “Bean Bag Genetics”

- Strategy:
  - Ignore the complexities of real populations and focus on the evolution of just one (or a few loci)
  - Treat the population as mating at random or, if subdivided, cross-migrating in a simple pattern
- Though successful, this strategy was mocked as “Bean Bag Genetics:” Ernst Mayr.
- In the absence of large number of genomic sequence data and powerful statistical algorithms, this strategy appeared to be the only viable one...

Example: DNA Variation in Drosophila

- Marty Kreitman: “Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*.”
- Sequence variation in a sample of natural (wild-type) alleles: 11 alleles from Florida (F1), Washington (Wa), Africa (Af), Japan (Ja) and France (Fr)
- In the region of 11 ADH alleles: No two alleles matched in their DNA sequences

Polymorphisms

- In the coding regions, some alleles did have the same sequence; 14 sites have two alternative nucleotides (**biallelic**); A site with different nucleotides in independently samples: **a sgregating site**, or **a polymorphic site**.

Polymorphisms

*

allele	39	226	387	393	441	513	519	531	540	578	606	615	645	684
Ref	T	C	C	C	C	C	T	C	C	A	C	T	A	G
Wa-S	.	T	T	.	A	A	C	.	.	.	.	.	.	.
F1-1S	.	T	T	.	A	A	C	.	.	.	.	.	.	.
Af-S	.	.	.	.	.	.	.	.	.	.	.	.	.	A
Fr-S	.	.	.	.	.	.	.	.	.	.	.	.	.	A
F1-2S	G	.	.	.	.	.	.	.	.	.	.	.	.	.
Ja-S	G	.	.	.	.	.	.	.	T	.	T	.	C	A
F1-F	G	.	.	.	.	.	.	G	T	C	T	C	C	.
Fr-F	G	.	.	.	.	.	.	G	T	C	T	C	C	.
Wa-F	G	.	.	.	.	.	.	G	T	C	T	C	C	.
Af-F	G	.	.	.	.	.	.	G	T	C	T	C	C	.
Ja-F	G	.	.	A	.	.	.	G	T	C	T	C	C	.

Polymorphisms

- About 1.8 of every 100 sites are segregating in the ADH sample (typical for *D. melanogaster*
- The variation at 13 of the 14 segregating sites is **silent** (i.e., they represent synonymous mutations, that changes the codon but not the encoded amino acid)
- The variation at the 578th nucleotide position results in a change of the amino acid at position 192 in protein, where a lysine (**AAG**) or a threonine (**ACG**) is found. This is a *replacement or non-synonymous polymorphism* as this nucleotide polymorphism causes an amino acid polymorphism.

Questions

- What causes these diversities within the same species?
- Why are there so many synonymous polymorphisms? Random mutations are mostly lethal...
- Note: Alcohol Dehydrogenase is an important enzyme as flies and their larvae are often found in fermenting fruits with high alcohol concentration
- Alcohol Dehydrogenase is used in the detoxification of ingested alcohol... A small change in the protein could have a serious consequence.

Variation Across Species

- Comparison of the coding region of the ADH loci in *D. melanogaster* vs. *D. erecta*:
- 36 out of 768 nucleotides differ between the two species
- Of the 36 differences, only 10 (26%) are non-synonymous

Loci and Alleles

- **Locus**: A chromosomal location referring to a segment of DNA (which may or may not have a phenotypic effect). A locus is a template for an allele.
- **Allele**: A segment of DNA sequence at a locus. An allele is an instantiation of a locus.
- The genome consists of a sequence of loci: one for each haploid chromosome. A diploid human has two alleles at a particular autosomal locus (one from father and the other from the mother). If they differ in nucleotide sequences, the human is heterozygote; otherwise, homozygote. They can also vary in copy-numbers... (things get a bit complicated)...

Different Alleles

- **By Origin**: They come from the same locus on different chromosomes (perhaps belonging to different individuals)
- **By State**: They have different nucleotide sequences
- **By Descent**: They do not share a common ancestor allele (i.e., during a relatively short time period in the recent past)
- **Identity by origin**, **Identity by state** or **Identity by descent**...

Why Cancer?

- If cancer is a disease of the genome...
- And since it should be under a negative selection pressure...
- *Why hasn't cancer been eliminated from multi-cellular metazoans by evolution*?

Few plausible answers

- *Shadow of evolution...*
- *Cancer has a selective advantage...* (surveillance vs. growth or fertility: P53, P63 & P73)... What is bad in somatic cells, could be an advantage for germline cells.
- *Cancer is a natural evolutionary process...* and intimately associated with multicellularity.

[End of Lecture #4]