

# Comprehensive Exam Notes

Patrick Dylan Barry

July 9, 2014

## **Abstract**

During my Ph.D. at the University of Alaska Fairbanks I needed to study for comprehensive exams. My committee made up of Anthony Gharrett, Megan McPhee, David Tallmon, and Eric Anderson were tasked with choosing topics that I needed to master in order to graduate and continue on the Ph.D. track. During my first committee meeting they made a list of topics: Relationship/Parentage analysis, population genetic theory, coalescence, inference models for molecular ecology, population conservation genetics as applied to fisheries management, evolutionary ecology of Pacific salmon, and molecular genetic methodologies and applications. This list scared the bejezus out of me, and so I decided that a good way of studying for these exams would be to create a short book that might aid in my preparation and be a good resource for the future. I think this should be worthwhile even if nothing more grows out of this than a study guide for my exams or possibly lecture notes for a course I instruct.

## Contents

# 1 A Brief History of the field of Genetics

Hooke (1665) coins the term 'cell' after observing cellular structure in cork.

Anton van Leeuwenhoek (1670s) invents compound-like microscope - observes animalcules.

Schwann (1847) - proposes animals made of tissues which were constructed of 'cells'  
Matthias Jakob Schleiden - plant cell structure - cells appear to be building blocks for complex organisms

Spontaneous generation Redi - flies prevented from laying eggs = no maggots Spallanzani - boiling hay infusion no animalcules Pasteur and Tyndall - put it to rest

Pangenesis - Aristotle eggs and sperm interact in mysterious way Hertwig used sea urchins to show developing embryo Strassburger did the same in plants - nucleus was important.

Fixity of Species - Linnaeus

Chromosomes - late 1800's Boveri Henkens Montgomery

Correns and deVries discover Mendel's work.

## 2 DNA - a review

In this section I will briefly review the discovery of DNA and review basic principles that will facilitate the understanding of preceding chapters.

### 2.1 DNA as the hereditary material

Major experiments that lead to the discovery of DNA as the hereditary material. Biological molecule mediates inheritance. Possibilities - Carbohydrate/polysaccharide, lipid, protein, nucleic acid Griffiths 1928 Dawson 1931 Alloway 1933 Avery, McCarty & MacLeod 1944 Hershey & Chase 1952 Fraenkel-Conrat and Singer 1957 Chargaff 1950 Linus Pauling Watson/Crick/Franklin

### 2.2 DNA structure

Nucleotides - Nucleosides etc.

### 2.3 Transcription

Brief description of DNA to mRNA

### 2.4 Translation

Brief description of mRNA to protein

## 3 Molecular Genetic Methods & Applications

### 3.1 Phenotypic Data

### 3.2 Codominant Phenotypic Data

### 3.3 Mitochondrial DNA (mtDNA)

#### 3.3.1 History and Description

Mitochondrial DNA (mtDNA) is, as its name suggests, DNA that is found within the mitochondria. First described by Carl Benda in 1898, mitochondria are organelles found in most eukaryotic cells (plants, animals, and fungi) that generate adenosine triphosphate (ATP) from the phosphorylation of ADP through cellular respiration. Details about the citric acid cycle (Krebs cycle) and the electron transport chain which take pyruvate (the product of glycolysis) to produce ATP are covered in most introductory biology text books and are omitted here.

Nuclear and mtDNA have different evolutionary histories. It is theorized that mtDNA was derived from the circular genome of an endosymbiotic bacteria 1.5 billion years ago. This theory was first described by Schimper who in 1883 noticed striking similarities between chloroplasts within green plants and freeliving cyanobacteria. The theory was later formalized by Mereschkowski, a Russian botanist, in 1910 and further advanced by Lynn Margulis in 1967.

Mitochondrial DNA is predominantly maternally inherited. There are a few explanations as to why this is the case. Mere dilution might be responsible. The egg has upwards of 1,000,000 mtDNA molecules while sperm have a high estimate of 1,000. The mtDNA on sperm is located on the tail to fuel the beating of the flagellum. When the acrosome reaction occurs, the tail may not enter the egg effectively eliminating the opportunity to pass on the male mtDNA. In 1999, it was reported by Sutovsky et al. that male mtDNA is tagged with ubiquitin to mark them for later destruction in the embryo.

Mitochondrial DNA in multicellular organisms is circular and double stranded. In unicellular organisms mtDNA is linearly organized DNA with telomeres and telomerase. In mammals, mtDNA encodes 37 genes (13 proteins, 22 tRNAs, and the large and small subunits of the rRNA). The guanine rich strand, referred to as the heavy strand (H-strand) encodes 28 genes, while the cytosine rich strand or light strand (L-strand) encodes the 9 other genes.

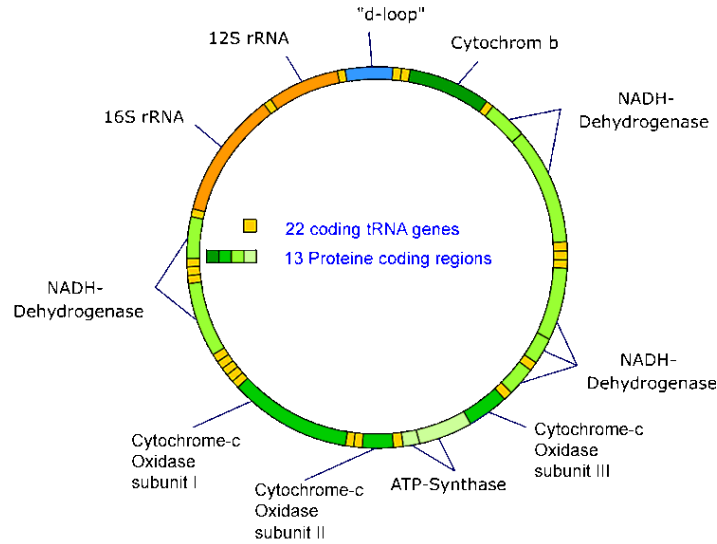


Figure 1: Structure of mammalian mtDNA molecule. The green segments represent the 13 protein genes, the yellow segments are the tRNA genes, and the orange are the rRNA subunits.

---

- 1970s to 1990 RFLP analysis of mtDNA dominated phylogenetics - switched to DNA sequencing of mtDNA - you can isolate mtDNA from other DNA by efficiently separation of cytoplasm from cell nuclei. Soft tissue is minced, homogenized and centrifuged at low speed (700 x g) to remove nuclei and cellular debris. then centrifugation at higher speed 20,000 x g pellets mitochondria. CsCl-EtBr gradient centrifugation at speeds in excess of 160,000 x g makes mtDNA appear as a discrete band that you can remove with a hypodermic needle and purify further with dialysis. - this was particularly fortuitous because organelle DNA can be transferred to the nucleus- sequences related to cytoplasmic genes now also exist as functionless derivatives in the nuclear genome Paralogous pseudogenes can wreak havoc on phylogenetic inference if they aren't accounted for.

'control region' responsible for initiating replication and transcription. variation exists in gene order and content useful in macrophylogeny.

for the most part homoplasmic -i single mtDNA sequence predominates all cells of tissues in the organism

---

### 3.3.2 Use in phylogenetics and population genetics

The analysis of mtDNA has been a mainstay of phylogenetics since XXXX and is still being used

### 3.4 Allozymes & Isozymes

Lacking the ability to look directly at the entire genome, early genetic studies relied on small bits of translated genomic regions. DNA is transcribed to mRNA which is then translated to protein (See Chapter 2). Consequently, proteins represent a way of looking at variation within DNA sequences. Proteins are polypeptides composed of strings of amino acids joined by covalent peptide bonds. Mutations within the protein coding regions can lead to different amino acids being incorporated into the polypeptide. Two general types of enzymes have been studied extensively: isozymes and allozymes. **isozymes** are all functionally similar forms of enzymes. **Allozymes** are a subgroup of isozymes that are coded by a single locus. Their name is derived from *allele* (alternative forms of a gene) and *enzyme*. Data collection for both types of marker rely on enzymatic reactions and staining so isozymes and allozymes can often be investigated simultaneously. Gel electrophoresis allows proteins to be separated based on their physical properties: charge, size, and shape. This method exploits the porous properties of a starch or cellulose acetate gel matrix and the differential charge of the amino acids that make up the allozyme. The rate of movement on a gel,  $u$ , is dependent on the net protein charge  $Q$ , shape  $r$ , strength of the electric field  $d$  and viscosity of the suspension medium  $n$ :

$$u = \frac{Qd}{4\pi^2n}$$

Charge differences among allozymes are resultant of the differential incorporation of positive (basic at neutral pH) amino acids lysine (Lys), arginine (Arg) and histadine (His) and negative (acidic at neutral pH) amino acids aspartic acid and glutamic acid.

The strength of allozymes

Cons: only observe non-synonous mutations, only look at water soluble proteins,

### 3.5 Restriction Length Polymorphisms (RFLPs)

#### 3.5.1 DNA barcoding

### 3.6 Microsatellites & Minisatellites

### 3.7 Sequencing

#### 3.7.1 Sanger Method

Who developed it, how it was done, pros&cons

#### 3.7.2 Pyrosequencing

Who developed it, how it was done, pros&cons

#### 3.7.3 RADsequencing

Who developed it, how it was done, pros&cons

### **3.8 Single Nucleotide Polymorphisms (SNPs)**

Who developed it, how it was done, pros&cons

#### **3.8.1 Genotyping by sequencing**