Comprehensive Exam Notes

Patrick Dylan Barry

July 21, 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. A major consequence of this is that many of the examples that I will use come from salmon. Similarly, many of the first few sections come from lecture notes that I took while in Tony Gharrett's Introduction to Genetics Course.

Contents

Chapter 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 Strssburger did the same in plants - nucleus was important.

Fixity of Species - Linnaeus

Chromossomes - late 1800's Boveri Henkens Montgomery

Correns and deVries discover Mendels work.

Chapter 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.

Possibilites - 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

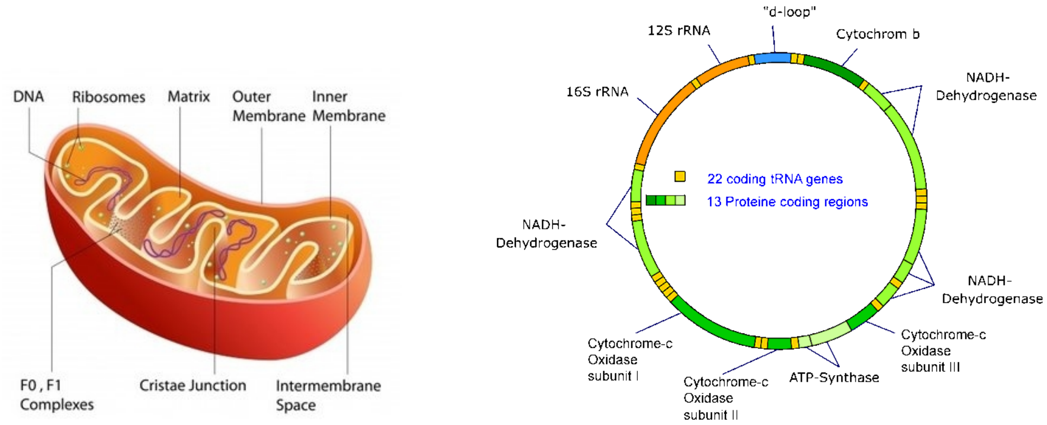
2.5 Mitochondrial DNA (mtDNA)

History and Description

Mitochondrial DNA (mtDNA) is, as its name suggests, DNA that is found within the mitochondria. The word mitochondrion, coined by Carl Benda in 1898, comes from the Greek word *mitos* meaning “thread” and *khondrion* meaning ‘ ‘little granule”. 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 extensivly covered in most introductory biology text books and are ommitted 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 similarties 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.

Mitochonrial DNA in multicellular organisms is cirular and double stranded. In unicellular organisms mtDNA is linearly organized DNA with telomeres and telemorase. 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) endoceds the 9 other genes. In addition to these coding regions is the control region (CR) that is responsible for initiating replication and transcription. The CR, sometimes refered to as the DLP, is composed of the displacement loop (D-loop) and associated transcription promoter regions.



(a)

(b)

Figure 2.1: (a) (b)

There are a few characteristics that make mitochondria a particularly appealing marker to molecular ecologists, however recent research has started to question the assumption that these traits hold true across a wide range of taxa (for detailed critiques see White et al. 2008 and Galtier et al. 2009).

Inheritance

First, mtDNA was thought to follow strict maternal inheritance in the majority of animals (Birky 1978 & Dawid and Blackler 1972). Paternal DNA is eliminated at different steps of fertilization depending on the taxa. In crayfish, which lack flagellum on their sperm, mtDNA is abscent from mature sperm Moses 1961 and thus cannot be inherited from the father. In the tunicate *Ascidia nigra*, paternal mtDNA is blocked from entering the egg preventing paternal inheritance(Ursprung and Schabtach 1965). In both cow and monkeys Sutovsky et al. (1999) suggest that ubiquination of mitochondria that occurs during spermatogenesis tags paternal mtDNA molecules for destruction by proteasomes and lysosomes in the embryo before the third embryonic cleavage. Although many studies have suggested otherwise and have been refuted (for a good review see Galtier et al. 2009).

Why is paternal DNA excluded from the cell? Bromaham 2003 proposes that (1)

Evolution of mtDNA molecule

The bacterial genome from which mtDNA arose has gradually shrunk over time. Genes from the original genome have been transfered either to the nucleus or to other organelles. As genes are transposed to the nucleus a copy can remain in the mitochondria making nuclear-mitocondrial dna (**numts**).

Selection for small genome size. Selective size advantage of faster replication

Small variations in mtDNA length is generally a result of homopolymer tracts generated by replication slippage. Larger differences may be due to duplication or deletions -¿ might be due to slippage as well?

Aerobic respiration takes place within the mitochondria and a major byproduct of ATP synthesis are reactive oxygen species, specifically superoxide and hydrogen peroxide. These free radicals produce a highly mutagenic envornment for mtDNA. This coupled with a limited repair mechanism leads to a much higher mutation fixation rate. Bogenhagen 1999 indicates that mtDNA can conduct base-excision repair, but cannot accomplish nucleotide excision repair, nor mismatch repair. Mutation rates vary among taxa. A universal molecular clock cannot be assumed. There also exists quite a bit of heterogeneity in mutaiton rate among mtDNA loci. The 12s and 16s rDNAs are more conserved than the CR. This is due to functional constraints on the proteins encoded. The CR for vertebrates is often divided into three parts based on base composition and mutational model (Review Saunders & Edwards Molecular Evolution 2000 51 97-109).

Transitions are more common than tranversions

Gene content and order has remained conserved among vertebrates, but some rearrangements have been observed. Comparisions of sea urchins and vertebrates show that tRNA genes are highly mobile with tanspositions and inversions occuring.

Zhang and Hewitt 2003 "mitochondrial DNA can be the cause of fortune, regret or headache" in describing the selection of mtDNA over other marker types. In plants the mutational rate is very slow Wolfe 1987, but it seems that mutational rates for plants is highly variable. Cho, Y 2004 show some of the highest mutational rates in plants! Very high recombination rate Palmer and Herbon 1988.

for the most part homoplasmic -¿ single mtDNA sequence predominates all cells of tissues in the organism new muation add to a hteroplasmic condition - 2 genotypes in 1 individual. Not really the case! mutations accumulate rapidly in nucleotide positions that have less selective constraints - protein coding then decrease Brown et al 1979. -¿ rapid evolving sequences are more useful for population level studies

mtDNA in molecular ecology

Population geneticists have used mtDNA to asses two salient questions: (1) what is the state of the exisitng genetic variation in the population and (2) identify reproductivly isolated populations. mtDNA displays substantial variation among individuals within and among populaitons. 1970s to 1990 restriction length polymorphism (RFLP) analysis of mtDNA dominated phylogenetics. 'Raw data' was restriction fragment digestion profiles - or restriction maps. The presence/absence of restriction sites could be used as qualitative data for parsimony trees.

corrleation between mtDNA diversity and nuclear diversity?

Maternally inherited means it is a promising marker for population genetics. Recent events - colonization, introductions and population bottlenecks sex biased dispersal Assumes neutrality and negligible back-mutation.

Mitocondrial psuedogens in the nuclear gemone complicate their use in population genetic studies. Lucily, back in the day geneticists used to isolate mtDNA from all nuclear DNA making this a moot point.

The analysis of mtDNA has been a mainstay of phylogenetics since XXXX and is still being used mtDNA is a single molecule and as such represents loci that are completely linked. Despite differences in the mutational rate among loci, it is just one look at the evolutionary history. Stochastic processes such as drift as well as selection will influence all mitochondrion loci. This look is also only the matrilineal history. Some studies have used both mtDNA and noncombining regions of the Y chromosome to look at population structure differences among sexes. The effective population size of mtDNA is at most (if we have 1:1 sex ratios) 1/4 that of nuclear autosomal DNA. A lower effective population size results in a faster lineage sorting rate and higher allele extinction rate. Selection not a huge deal if character states (nucleotides) are still synapomorphic. Does effect estimation of divergence times though!

Two issues: Incomplete lineage sorting can lead to discordanent gene and species trees constructued using mtDNA genotypes.

If a female has all male offspring that mtDNA lineage ends after that generation. Let's assume that females produce daughters according to a Poisson distribution with $\lambda = \mu$. The probability that a female will have no offspring in the next generation is simply $exp^{-\mu}$ (see below). The loss after G generations can be expressed as $P_G = e^{\mu(x-1)}$, where x is the probability of loss in the previous generation. Avise et al. showed that the probability of survival of two or more mtDNA lineages

$$P(\theta) = \frac{1}{\theta!}\lambda^{\theta} exp^{-\lambda},\ E(\theta) = \lambda,\ \&\ var(\theta) = \lambda \tag{2.1}$$

$$P(0) = \frac{1}{0!}\mu^{0} exp^{-\mu} \tag{2.2}$$

$$P(0) = exp^{-\mu} \tag{2.3}$$

```
# R code showing probability of losing an mtDNA lineage after 1 generation
off ← seq(from = 0, to = 10, by = 0.25)  # range of offpring you might have
d ← dpois(x = 0, lambda = off)  # calculate prob of 0 daughters for # offpring
plot(x = off, y = d, xlab = ”Number of female offspring produced”, ylab = ”
    Probability of having only sons”,
    type = ”l”)
```

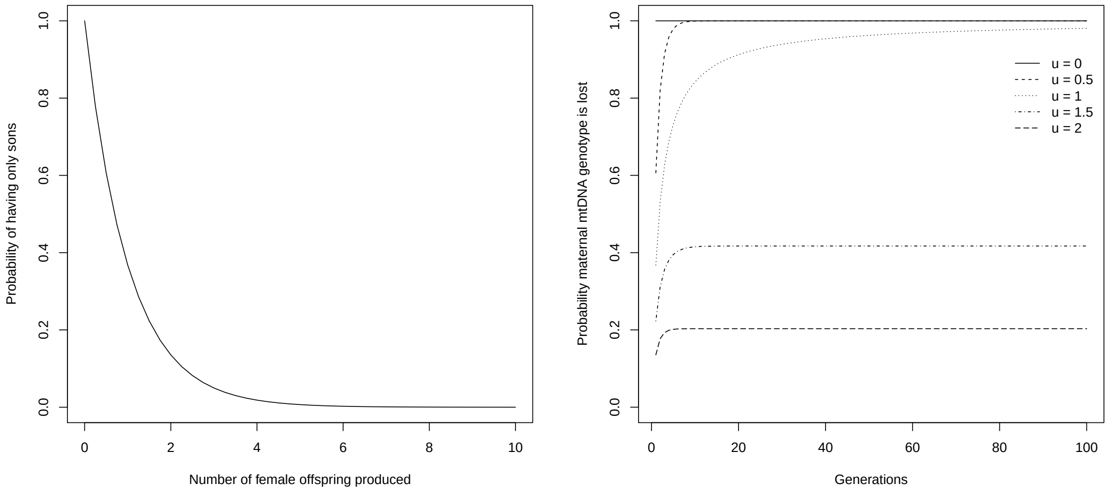
```
# R code showing how many maternal lineages are lost depending on how many
# female offering on average are left (u)
gen ← seq(from = 1, to = 100, by = 1)
u ← c(0, 0.5, 1, 1.5, 2)
PSA ← matrix(data = NA, ncol = length(gen), nrow = length(u))
for (fo in 1:length(u)) {
    for (g in 1:length(gen)) {
        if (g == 1) {
            PSA[fo, g] ← exp(-u[fo])
        } else {
            PSA[fo, g] ← exp(u[fo] * (PSA[fo, g - 1] - 1))
        }  #else
    }  #loop over fo
}  #loop over g
plot(x = gen, y = PSA[1, ], xlab = ”Generations”, ylab = ”Probability maternal mtDNA
     genotype is lost”,
    type = ”l”, ylim = c(0, 1))
for (p in 2:nrow(PSA)) {
    points(x = gen, y = PSA[p, ], type = ”l”, lty = p)
}
legend(x = ”topright”, inset = c(0.01, 0.1), bty = ”n”, legend = c(paste(”u”,
    ”=”, u, sep = ” ”)), lty = seq(from = 1, to = nrow(PSA), by = 1), )
```

Differential introgression If two species existed we can imagine there being nuclear loci that are responsible for their reproductive isolation. Patterns of introgression for nuetral nuclear loci will depend on how tightly linked they are to the markers responsible for isolation. mtDNA is unlinked to nuclear DNA and consequently the introgression of mtDNA variants will be much greater than for nuclear DNA. As a result inferences made from nuclear and mtDNA will be discordant.

DNA Barcoding

Recently, the use of short segments of mtDNA for taxonomic identification has been proposed. The basic theory is that we can use a short segment of DNA (a barcode) to identify an unidentified individual to species Hebert 2004. Later the amount of intraspecific divergence relative to interspecific divergence was proposed as a method to identify cryptic species DeSalle, Egan & Siddall, 2005. Many of these studies focused their efforts on the cytochrome oxidase I subunit (COI). This approach requires extensive examination of the intraspecies divergence one would expect to see and then categorizing the interspecies divergence.

Due to incomplete lineage sorting the COI gene may not reflect the true relationships among species. This is one of the various critiques of using a single locus for species delineation (Moritz & Cicero, 2004; Meyer & Paulay, 2005; Will, Mishler & Wheeler, 2005).



(a) (b)

Figure 2.2: (a) If we assume that females in a population produce female offspring according to the Poisson distribution with mean μ, we can calculate the probability that a female mtDNA lineage is lost after one generation according to equation 3. It is fairly intuitive that the more female offspring produced, the higher then chance that a mtDNA haplotype lineage persists to the next generation. If on average only 1 female offspring is left, then there is a 37% probability that the no daughters will be produced ending the mtDNA lineage for that family. We can also think of this in terms of how many maternal lineages will not persist. (b) We can expand this and look at the persistence of mtDNA lineages through time. We can vary the average number of daughters produced by females (μ). Again it is not surprising that if on average few daughters are produced, the time that it takes for an mtDNA lineage to go extinct is very short.

Chapter 3

Molecular Genetic Markers

3.1 Phenotypic Data

3.2 Codominant Phenotypic Data

3.3 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 *all*ele (alternative forms of a gene) and en*zyme*. 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:

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

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,

3.4 Restriction Length Polymorphisms (RFLPs)

3.5 AFLPs

3.6 Microsatellites & Minisatellites

3.7 RAPiD

3.8 Sequencing

Sequences are the ideal molecular data. If we had full genome sequences from all individuals in a study it would be awesome. All other molecular markers discussed so far are just small snippets of the genome that we assume, if randomly sampled, are descriptive of the rest of the genome. Many of the previous markers rely on sequencing for their development. Sequencing has been around since the mid 1970's, but new methods are currently being developed that improve upon these existing methods and make sequencing large pieces of the genome a much more economical endeavor.

When comparing each type of sequencing method the pros and cons that would determine its utility and eventually its use by researchers can be put into four broad categories:

- *Read length* - Ideally we would be able to generate one long sequence, but usually reads are 100bp -1000bp and short fragments need to be aligned to make larger sequences.
- *Speed* - The time it takes to generate samples involves sample preparation, sequencing, scoring, and alignment.
- *Accuracy* - One, if not the most important, attributes is generating good quality nucleotide scores.
- *Cost* - If a method is not economical it is unlikely to be adopted by many research labs.

There are often tradeoffs between each of these attributes and what one lab values in generating sequence data may not be what other labs value. As a result many labs differ in what technology they have adopted. The platforms on which the technology is developed and implemented are not cheap and the technology is rapidly changing.

3.8.1 Sanger Method

Sanger and Coulson developed the *plus and minus* method in 1975. This method took used *E. coli* DNA pol I and bacteriophage T4 DNA pol with different nucleoside triposphates. DNA pol I first extends the primer copying the template strand in the presence of four deoxyriobotriphoshates, one of which is labeled with ^{32}P. This should create many copes of different length. From previous research it was shown that DNA polymerase in the absence of one nucleotide, elongation would proceed until the position where the missing base should go. Using the partially elongated pieces from the first round of replication, Sanger & Coulson used a *minus* system, lacking one nucleotide, to deduce the sequences of each fragment up to the missing residue. Four different treatments need to be done, one for each nucleotide. The *plus* system is based on the research of Englund (1972) showing that in the presence of a single nucleotide, DNA pol from T4 will degrade dsDNA from its 3' end. The exonuclease activity is regulated by the presence of the nucleotides present - degradation only occurs in the absence of the nucleotides. This method was applied to the mixture obtained from the first elongation mixture produced, again each nucleotide was done separately, so that when fractionated by electrophoresis bands will indicate the positions of each residue in the sequence. The products in the *plus* system will be one residue larger than those in the *minus* system.

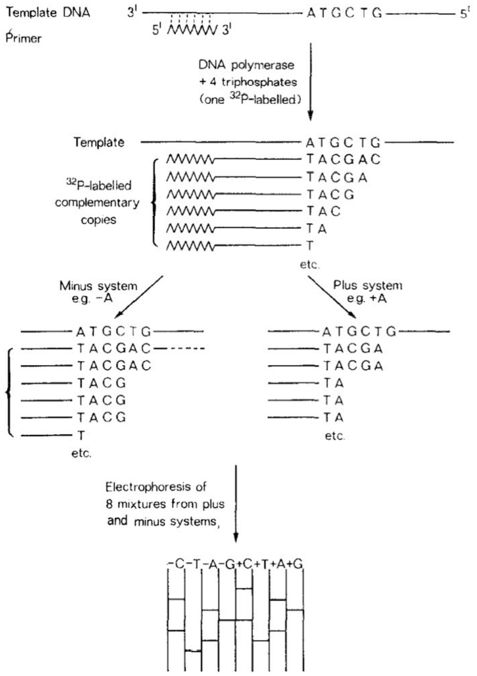


Figure 3.1: The *plus minus* method of sequencing DNA developed by Sanger and Coulson in 1975

A couple problems were encountered. First, all possible length fragments should be created in the initial elongation step. This was, often, not the case. Certain products were often found in high amounts while little or no yield of other fragment lengths were observed. Second, often artifact bands were observed making the inference of the target sequence difficult. Third, the accuracy of the sequence was somewhat suspect and relied on cooberation with amino acid sequence data. Fourth, it is only applicable to single stranded DNA. Lastly, 50 nucleotides in a couple days!

In 1977 Sanger et al. proposed the method of DNA sequencing with chain-terminating inhibitors (**the Sanger method**). Adkinson et al. 1969 observed that 2', 3'-dideoxythymidine triphosphate has inhibitory ability because ddT lacks a 3'-hydroxy group. If both ddTTP and dTTP are incubated in addition to the other three deoxyribonucleoside triphosphates with a primer and DNA polymerase fragments of different length will be produce each terminating where a dT should have been incorporated. The fragments can be fractionated by electrophoresis on an acrilamide gel. Mixtures with each terminator can be run side by side on an acrylamide gel to infer the complete sequence.

This method was better than the *plus minus* method because it did not require a preliminary

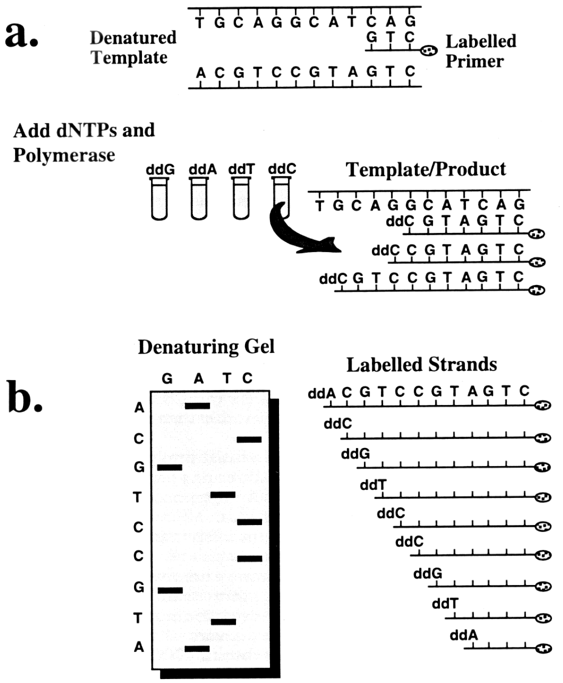


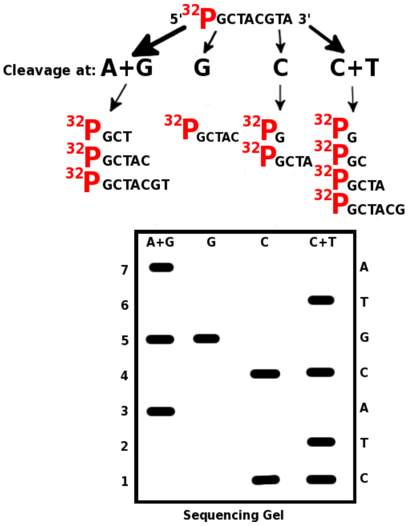
Figure 3.2: The Sanger method of DNA sequencing by chain termination.

extension, it only required one type of DNA polymerase, had fewer artifact bands, all individual nucleotides in long stretches came out as unique bands, and produced longer read lengths.

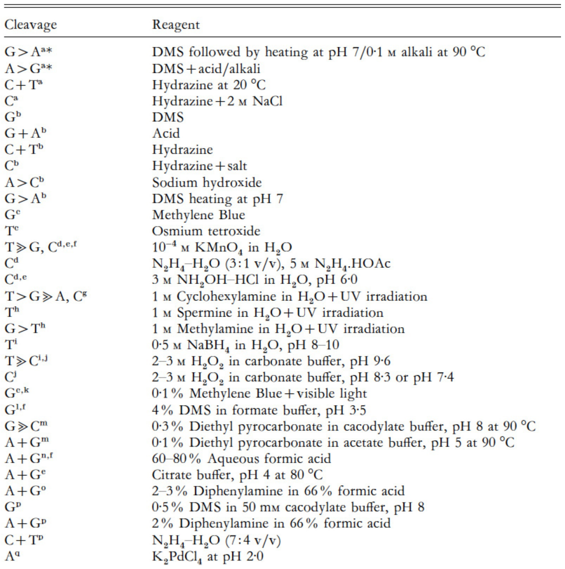
3.8.2 Maxam & Gilbert method

The Maxam & Gilbert method was described in 1977. The method takes advantage of the fact that certain chemicals can break the glycoside and phosphodiester bonds within DNA. The first step in the method consists of labeling the 5' end of a single strand of DNA, typically this was done with ^{32}P, but ^{35}S was also commonly used. Second, the bases were modified by breaking the glycoside bond between the ribose sugar and the base. Dimethyl sulfate attacks purines while hydrazine attacks prymadines (for a comprehensive list see Fig XX which is reproduced from Franca et al 2002). Third, piperdine is used to cleave the phosphodiester bond when the base is displaced. Generally for different treatments were done in the second step so that you could unambiguously call all four bases. After the fragments were cleaved the products could be fractionated by electrophoresis with an acrylamide gel.

When the method was developed Maxam & Gilbert were able to get 100bp reads within a few days. By 1980, they had refined the technique to get 250bp reads and by 1995 the process had been automated by Dolan et al (1995) with 500bp sequences. At the time when the method was developed this method was used because it did not rely on PCR amplification and it could be easily controlled in the lab. However, the necessity of handling some pretty nasty chemicals and the rather long wait times to produce sequence information made adopting new sequencing techniques



(a)



Cleavage	Reagent
$G>A^{a*}$	DMS followed by heating at pH 7/0·1 M alkali at 90 °C
$A>G^{a*}$	DMS + acid/alkali
$C+T^{a}$	Hydrazine at 20 °C
C^{a}	Hydrazine + 2 M NaCl
G^{b}	DMS
$G+A^{b}$	Acid
$C+T^{b}$	Hydrazine
C^{b}	Hydrazine + salt
$A>C^{b}$	Sodium hydroxide
$G>A^{b}$	DMS heating at pH 7
G^{c}	Methylene Blue
T^{c}	Osmium tetroxide
$T\gg G, C^{d,e,f}$	10^{-4} M $KMnO_4$ in H_2O
C^{d}	N_2H_4–H_2O (3:1 v/v), 5 M N_2H_4.HOAc
$C^{d,e}$	3 M NH_2OH–HCl in H_2O, pH 6·0
$T>G\gg A, C^{g}$	1 M Cyclohexylamine in H_2O + UV irradiation
T^{h}	1 M Spermine in H_2O + UV irradiation
$G>T^{h}$	1 M Methylamine in H_2O + UV irradiation
T^{i}	0·5 M $NaBH_4$ in H_2O, pH 8–10
$T\gg C^{i,j}$	2–3 M H_2O_2 in carbonate buffer, pH 9·6
C^{j}	2–3 M H_2O_2 in carbonate buffer, pH 8·3 or pH 7·4
$G^{c,k}$	0·1 % Methylene Blue + visible light
$G^{l,f}$	4 % DMS in formate buffer, pH 3·5
$G\gg C^{m}$	0·3 % Diethyl pyrocarbonate in cacodylate buffer, pH 8 at 90 °C
$A+G^{m}$	0·1 % Diethyl pyrocarbonate in acetate buffer, pH 5 at 90 °C
$A+G^{n,f}$	60–80 % Aqueous formic acid
$A+G^{e}$	Citrate buffer, pH 4 at 80 °C
$A+G^{o}$	2–3 % Diphenylamine in 66 % formic acid
G^{p}	0·5 % DMS in 50 mM cacodylate buffer, pH 8
$A+G^{p}$	2 % Diphenylamine in 66 % formic acid
$C+T^{p}$	N_2H_4–H_2O (7:4 v/v)
A^{q}	K_2PdCl_4 at pH 2·0

(b)

Figure 3.3: (a) (b)

attractive.

3.8.3 Pyrosequencing

Pyrosequencing is real time DNA sequencing (sequencing by synthesis) that detects the release of PPi as DNA polymerase forms phosphodiester bonds between nucleotides. The technique was developed by Mostafa Ronaghi and Pål Nyrén in XXXX. The really cleaver mechanism behind pyrosequencing is the use of the enzyme luciferase. Luciferase is a general class of enzymes used in bioluminescence; the specific one used in pyrosequening is that of fireflies *P. pyralis*. The entire process is accomplished in three reactions:

$$(DNA)_n + dNTP \xrightarrow{DNApolymerase} (DNA)_{n+1} + PPi$$

$$PPi + APS \xrightarrow{ATPsulphurylase} ATP + SO_4^{-2}$$

$$PPi + luciferin + O_2 \xrightarrow{luciferase} AMP + PPi + oxyluciferin + CO_2 + hv$$

First DNA is exposed to one dioxyribonucleic acid at a time. If the nucleotide introduced is the complementary base for the next position in the sequence DNA polymerase will encorporate it creating a phosphodiester bond and releasing PPi. Another enzyme ATP sulphyrylase converts the PPi and APS to ATP and SO_4^{-2}. This second reaction is needed to generate ATP that will be used by the enzyme luciferase. In the final step the luciferin, a light-emitting compound, and PPi react in the presence of oxygen to form AMP, PPi, oxyluciferin, carbon dioxide and light (hv). After introducing a single nucleotide base, light emission indicates if the base was used by DNA polymerase. The amount of light produced indicates if more than one base is added in sequence.

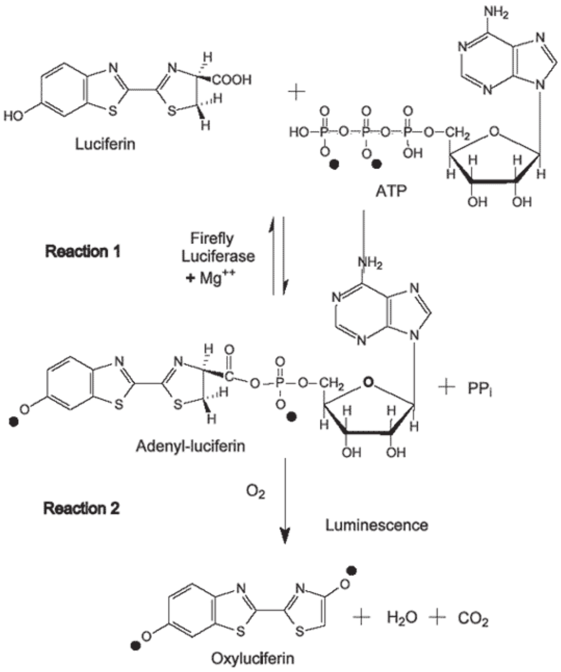


Figure 3.4

Pyrosequencing can be divided into two general types: solid and liquid phase. Solid phase was described by Ronaghi et al. in 1996. This method requires introducing a single nucleotide, washing the templete thouroughly before subsequent nucleotide additions to remove non-incorporated dNTPs and ATP. The liquid phase method was describe by Ronaghi et al. in 1998. Instead of washing the templete between dNTP introductions the nucleotides are degraded by the enzyme apyrase.

Pyrosequening has some pretty clear advantages over previous methods. The is no need for labelled primers or nucleotides. Sequencing solely relies on the emission of light! There is also no need for gel electrophoresis, so there is less dealing with acrylamide and other nasty chemicals. Sequencing is done rapidly. Chain extension is accomplished in the 2min cycle time. All sequencing is accomplished at room temperature and at physiological pH. Over the course of adopting this technology some difficulties have been encountered. The solid phase method has issues with signal loss as the template is washed repeatedly. Similarly, in the liquid phase method apyrase activity decreases over time such that nucleotides and ATP are not completely digested before the next introduction. Despite its advantages pyrosequening is not a technique that many people will be adopting in the future. The GS FLX platform could read 400Mb in a 10 hour run, but each run cost between $5,000 and $7,000. Roache purchased the technology from Qiagen and promptly discontinued the entire line in 2013 and will only be servicing these platforms through mid 2016.

3.8.4 RADsequencing

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

3.9 Single Nucleotide Polymorphisms (SNPs)

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

3.9.1 Genotyping by sequencing

Chapter 4

Population Genetics

4.1 Hardy-Weinberg Principle

4.2 Mutation

4.3 Drift

4.4 migration

4.5 Selection