http://fire.biol.wwu.edu/trent/trent/direct_detection_of_genotype.html

Direct Detection of Genotype using the model organism *Arabidopsis thaliana*

Part 1 of Lab Exercise

Part 2 of Lab Exercise

Links to student g3l photos

Lab Report & Data Workup

Associated Lecture Material:

Intro to Arabidopsis & Reverse genetics

Review PCR basics with this Biol 321 lecture

Technical info about reagents used in this lab:

Hazards of Ultraviolet light

About Ethidium Bromide

Ethidium Bromide Fact Sheet

Like most other model organism Arabidopsis thaliana has a sequenced genome?

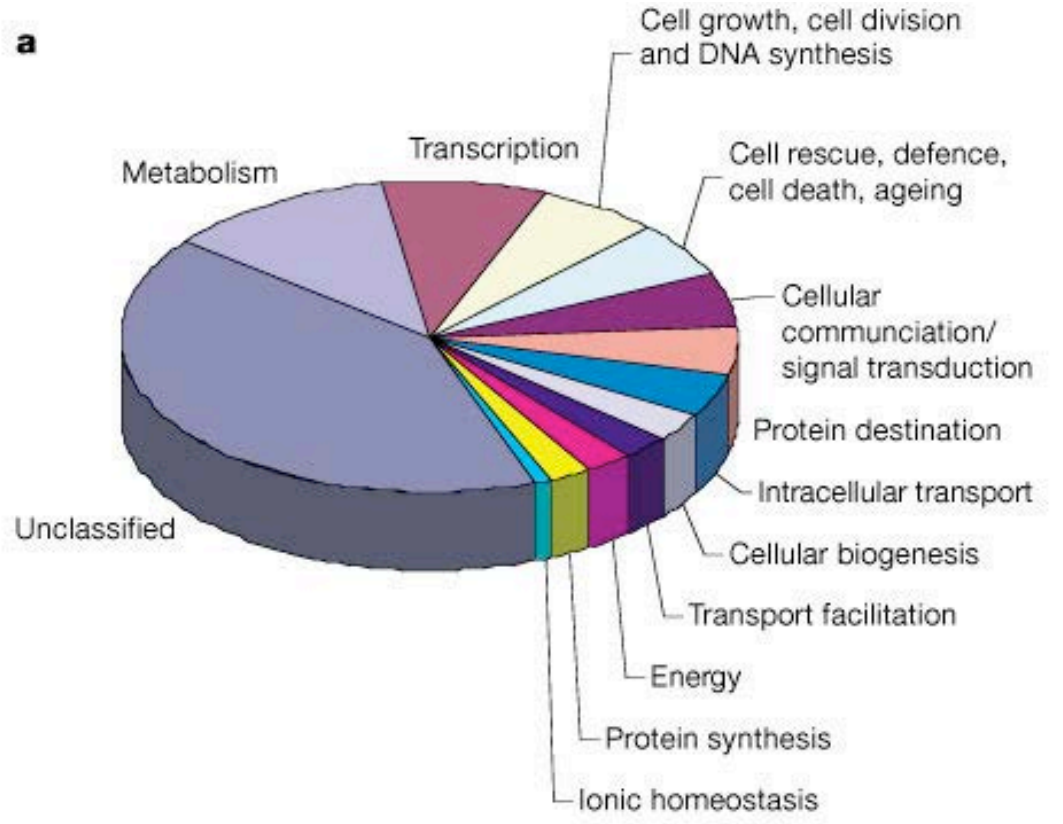
What do we mean by "sequenced genome"?

What sort of info does a complete sequence provide?

Genome project defined 25,500 genes

What did they know about these genes?

The Arabidopsis thaliana (weedy plant) genome project was completed in 2000. Here is a breakdown of the functional analysis of the 25,500 genes discovered in the genome of this organism.



Nature 408: 796 Dec. 14, 2000

From NATURE REVIEWS GENETICS July 2006

After completion of the sequencing of the Arabidopsis thaliana genome in 2000, the plant biology community faced the new challenge of assigning biological functions to all of the genes in this 120 Mb genome.

- ***Computational annotation*** of the genome was initiated to predict the locations of the genes and their basic structural elements (introns, exons and putative regulatory sequences) and led to rapid annotation of more than 25,000 Arabidopsis thaliana genes.
- Although extremely useful, such *ab initio* (from scratch) annotation generates numerous inaccuracies — at least 40% of gene predictions were subsequently found to be erroneous.
- Further refinement and validation of the computational gene models and identification of additional genes not predicted by gene-finding algorithms has been achieved using various experimental approaches
- At the time of completion of the genome sequence, only ***~10% of the 25,500 genes that were initially predicted had an experimentally assigned function.***
- Determination of the functions of the remaining ***90% of genes presented a tremendous challenge,*** not only because of the large number of genes to be examined, but also because defining what constitutes a 'gene' is itself a complex problem.

How to figure out what these genes do?
The paradigms of Forward and Reverse Genetics

Reverse genetics

Gene sequence (may reveal specific molecular function) → Knock-out (targeted mutagenesis) → reveal phenotype → infer function

Forward genetics

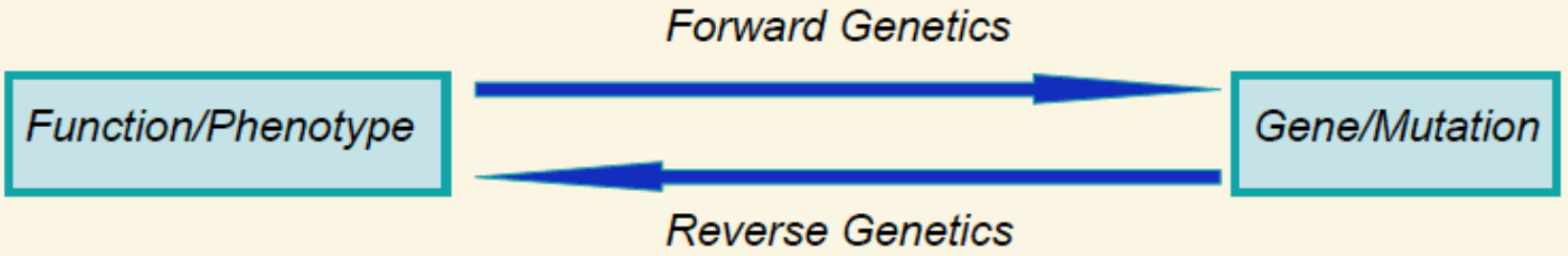
Random mutagenesis → look for phenotypes of interest → identify mutated gene(s) → examine DNA and protein sequences → infer molecular function and specific role in process

Assigning Gene Function Using a Reverse Genetics Approach

What is the genetic basis of a particular phenotype?
(How does one determine the function of a gene, or the identity of genes responsible for a trait?)

Forward Genetics:
Starts with a phenotype and moves towards the gene

Reverse Genetics:
Starts with a particular gene and assays the effect of its disruption



25,500 in 11,000 gene families

What is a gene family?

Arabidopsis H^+-*ATPase* Gene Family

Gene	Location	Function
AHA1	whole plant	?
AHA2	root cortex	?
AHA3	phloem	?
AHA4	root endodermis	nutrient uptake
AHA5	whole plant	?
AHA6	-	?
AHA7	-	?
AHA8	-	?
AHA9	anthers	?
AHA10	seeds	?
AHA11	hypocotyl	?
AHA12	-	psuedogene

Gene Family: a set of genes in one genome, all descended from the same ancestral gene

http://ghr.nlm.nih.gov/geneFamily

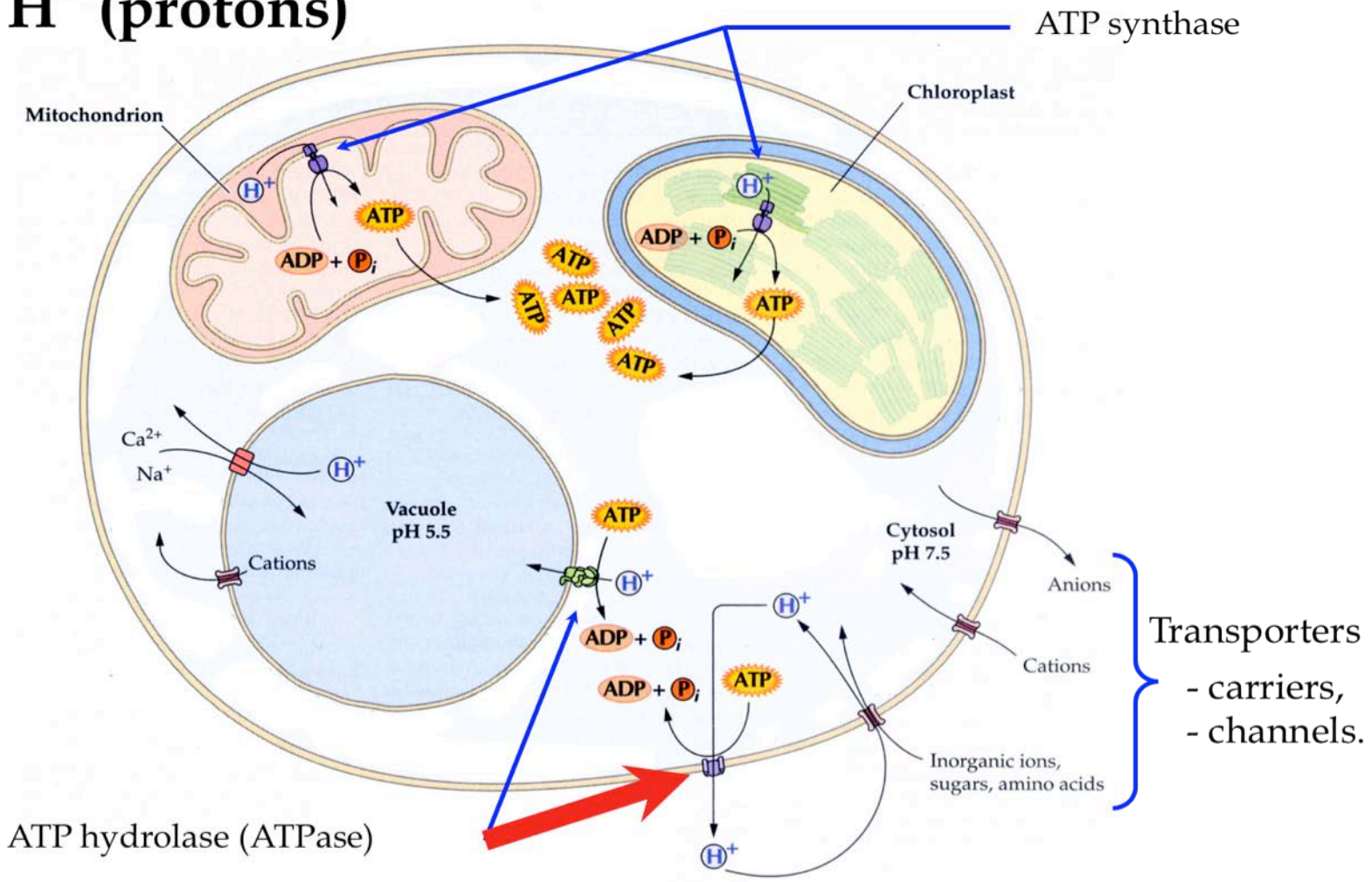
Gene Families

A gene family is a group of genes that share important characteristics. Classifying individual genes into families helps researchers describe how genes are related to each other. For more information, see What are gene families? in the Handbook.

The following families, defined by the HUGO Gene Nomenclature Committee, are included in Genetics Home Reference.

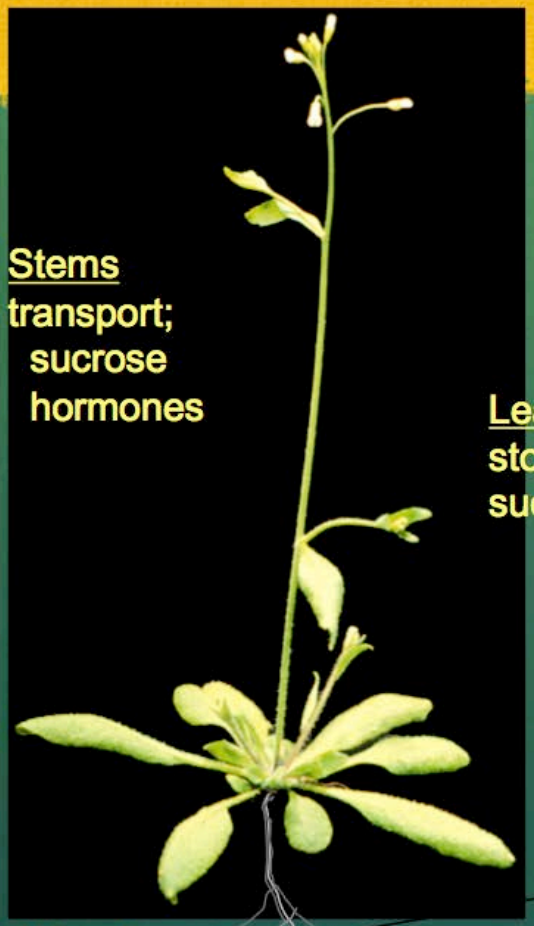
- aaRS (aminoacyl tRNA synthetases)
- ABC (ATP-binding cassette transporters)
- ABHD (abhydrolase domain containing genes)
- ACS (acyl-CoA synthetase family)
- ADAMTS (ADAMTS metallopeptidase with thrombospondin type 1 motif family)
- ALDH (aldehyde dehydrogenases)
- ALOX (arachidonate lipoxygenases)
- ANKRD (ankyrin repeat domain containing)
- ARHGEF (Rho guanine nucleotide exchange factors)
- ATP (ATPase superfamily)
- bHLH (basic helix-loop-helix)
- BIRC (baculoviral IAP repeat-containing genes)
- blood group (blood group determining genes)
- CACN (calcium channels)
- CATSPER (cation channels, sperm associated)
- CD (CD molecules)
- CDH (cadherin superfamily)
- CDK (cyclin-dependent kinases)
- CHMP (charged multivesicular body proteins)
- chromatin-modifying enzymes (chromatin-modifying enzymes)
- CLCN (chloride channels, voltage-sensitive)
- CNG (cyclic nucleotide-regulated channels)
- COLEC (collectins)

H^+ (protons)



Adapted from Biochemistry and Molecular Biology of Plants, pp. 115

Proton Pumps *in planta*



Arabidopsis

Stems
transport;
sucrose
hormones

Leaves
stomata (gas exchange)
sucrose transport

Roots
root hair
growth
mineral uptake

Pollen
tip growth

Anthers
cell elongation

Embryo/Seeds
loading

Arabidopsis H^+-*ATPase* Gene Family

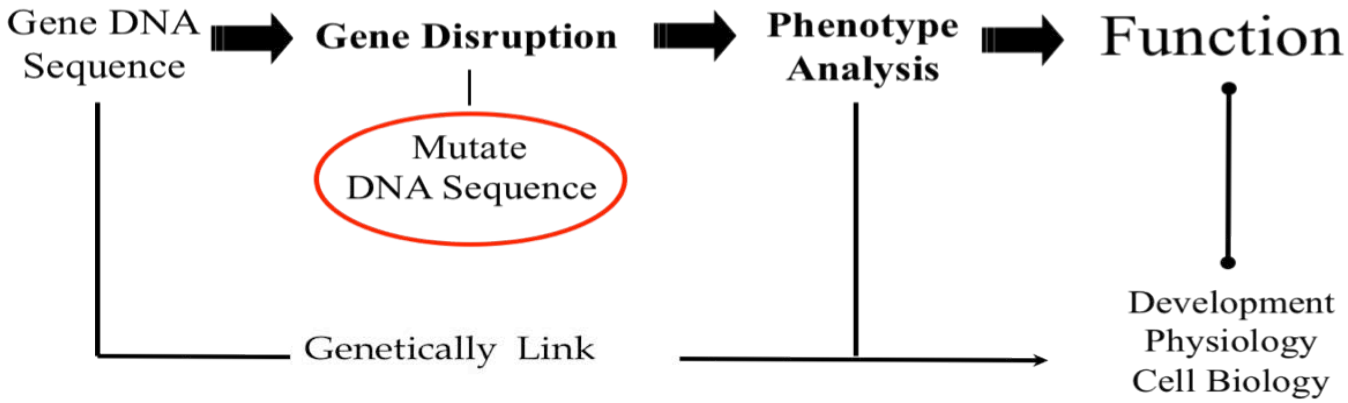
Gene	Location	Function
AHA1	whole plant	?
AHA2	root cortex	?
AHA3	phloem	?
AHA4	root endodermis	nutrient uptake
AHA5	whole plant	?
AHA6	-	?
AHA7	-	?
AHA8	-	?
AHA9	anthers	?
AHA10	seeds	?
AHA11	hypocotyl	?
AHA12	-	psuedogene

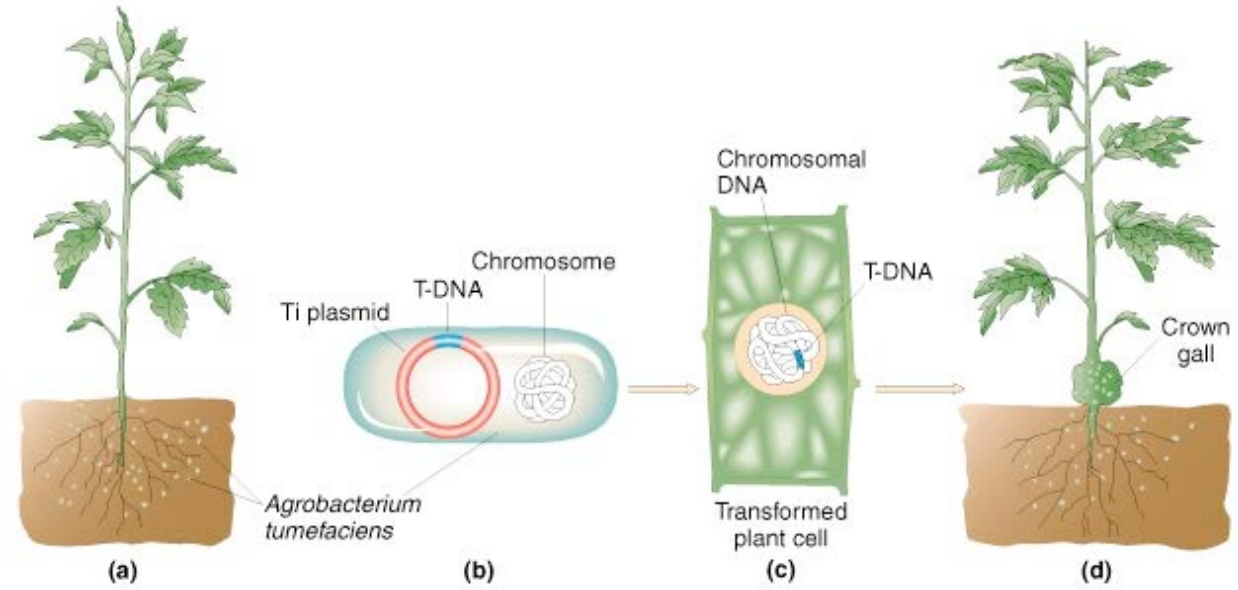
What does aha3 do for the organism?

What biological processes is this gene involved in?

Reverse Genetics

Functional Genomics

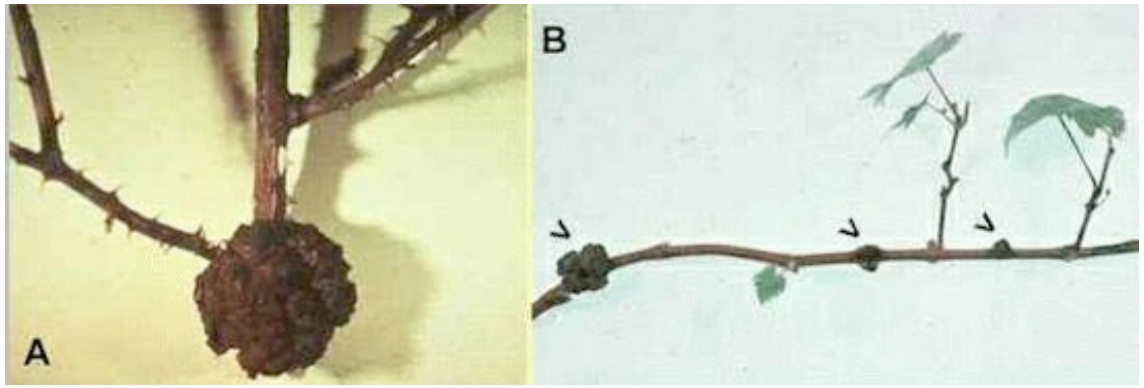




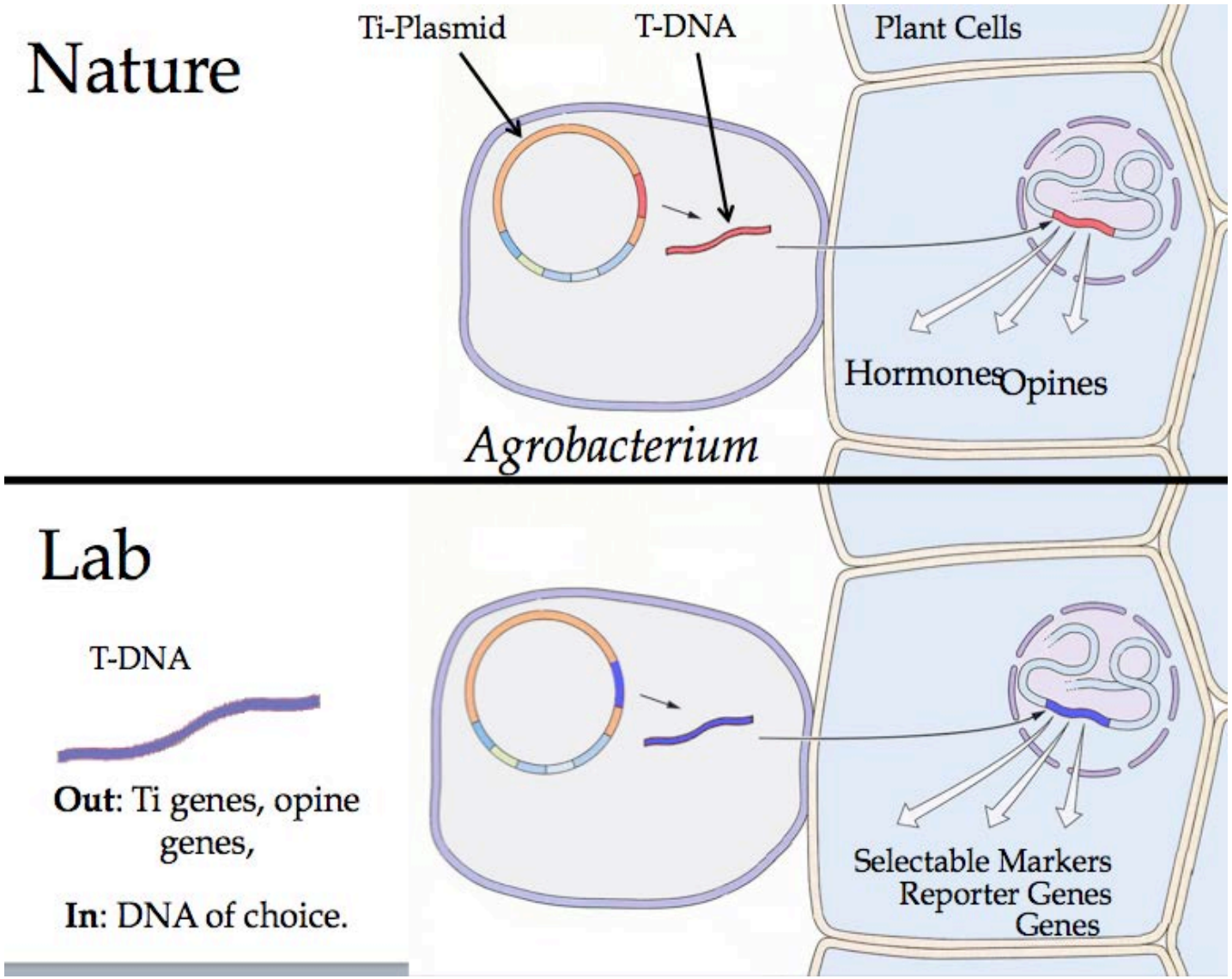
In the process of causing crown gall disease, the bacterium *Agrobacteria tumefaciens* inserts a part of its Ti plasmid (a region called the T DNA) into a chromosome of the host plant.

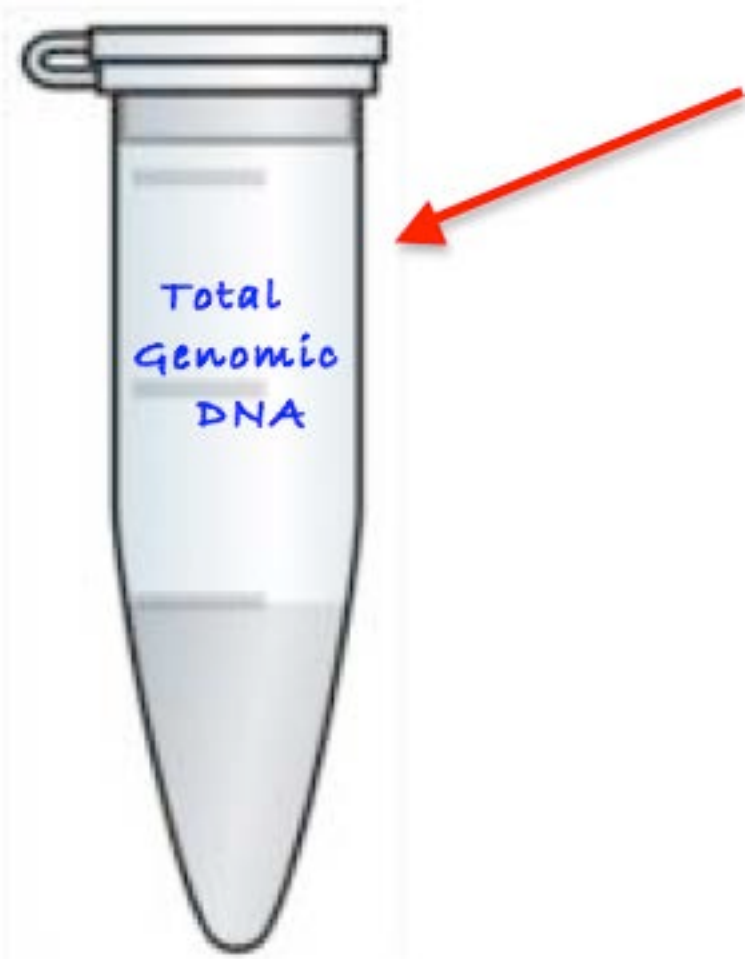
Figure 8-31

The genes on the T DNA direct the synthesis of cytokinins which stimulate plant cell division.



crown gall on a rose plant


Nature
Ti-Plasmid
T-DNA
Plant Cells
Hormones
Opines
Agrobacterium
Lab
T-DNA
Out: Ti genes, opine genes,
In: DNA of choice.
Selectable Markers
Reporter Genes
Genes

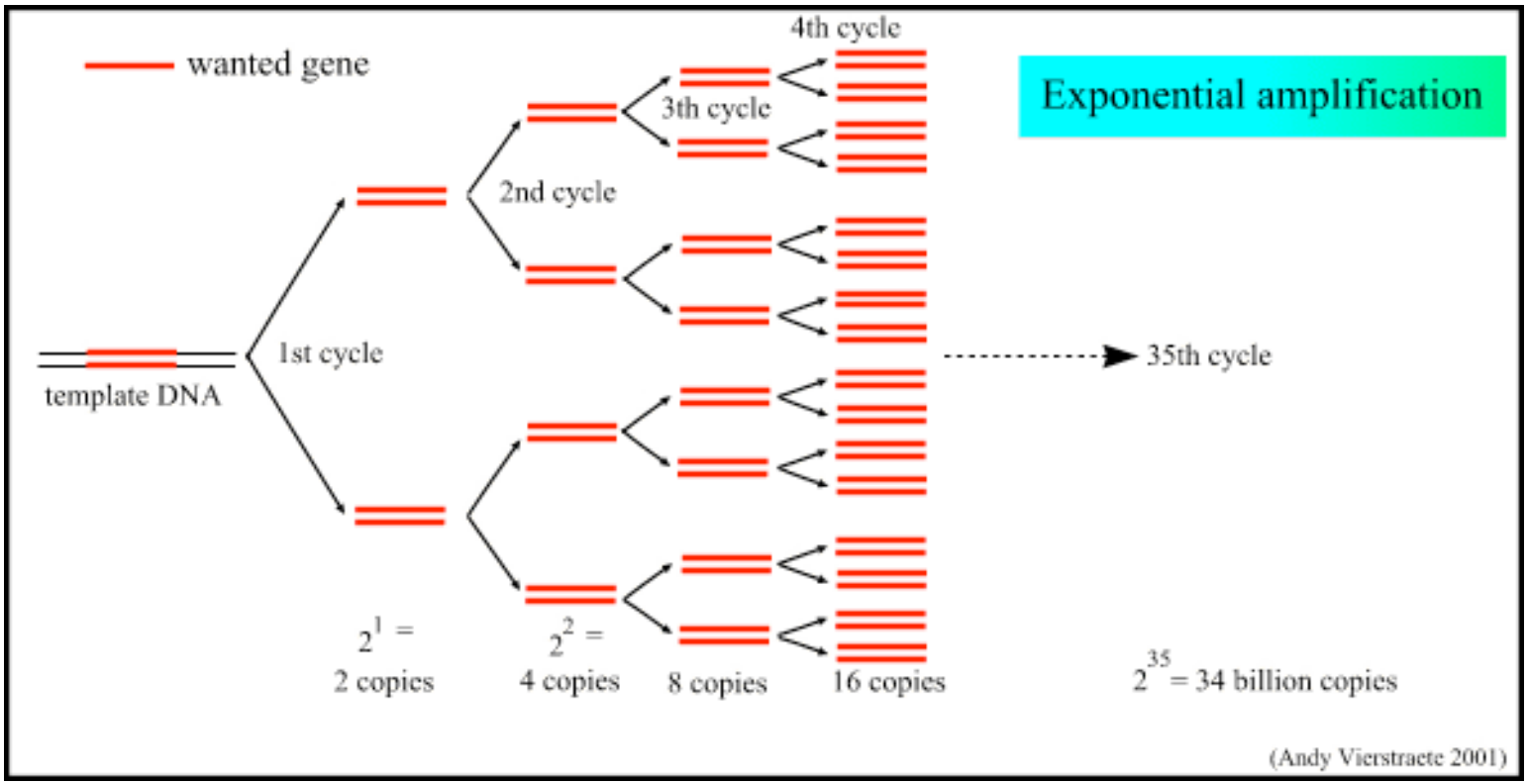


Direct Detection of Genotype using PCR

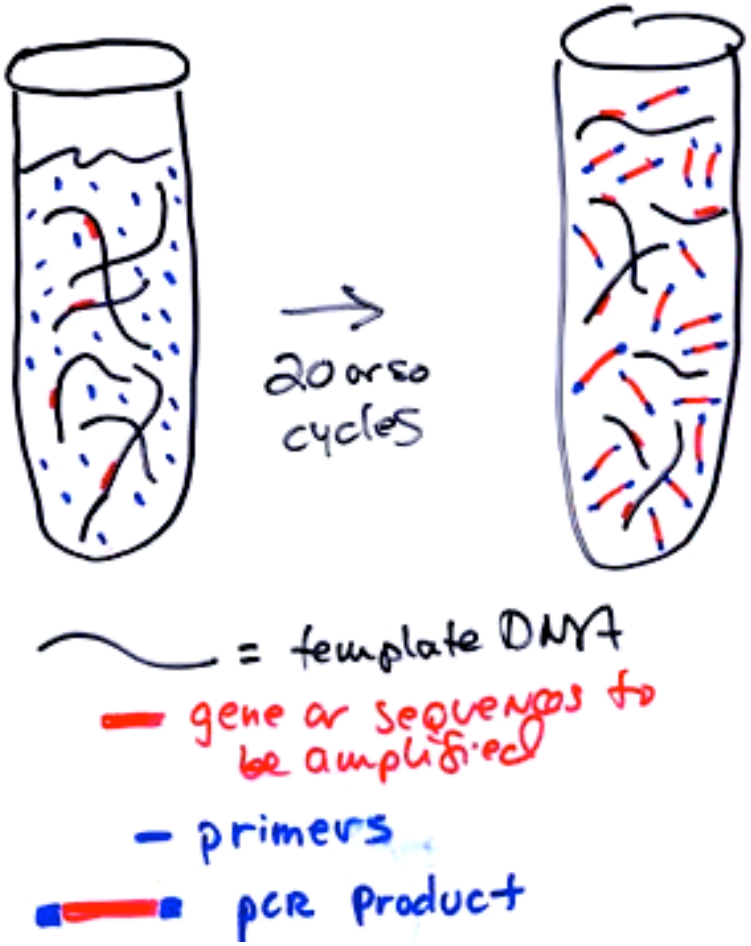
http://fire.biol.wwu.edu/trent/trent/10.05.19lecture.pdf

How do we target amplification to our specific sequences of interest?

How come only the red sequence is amplified from the starting template: 



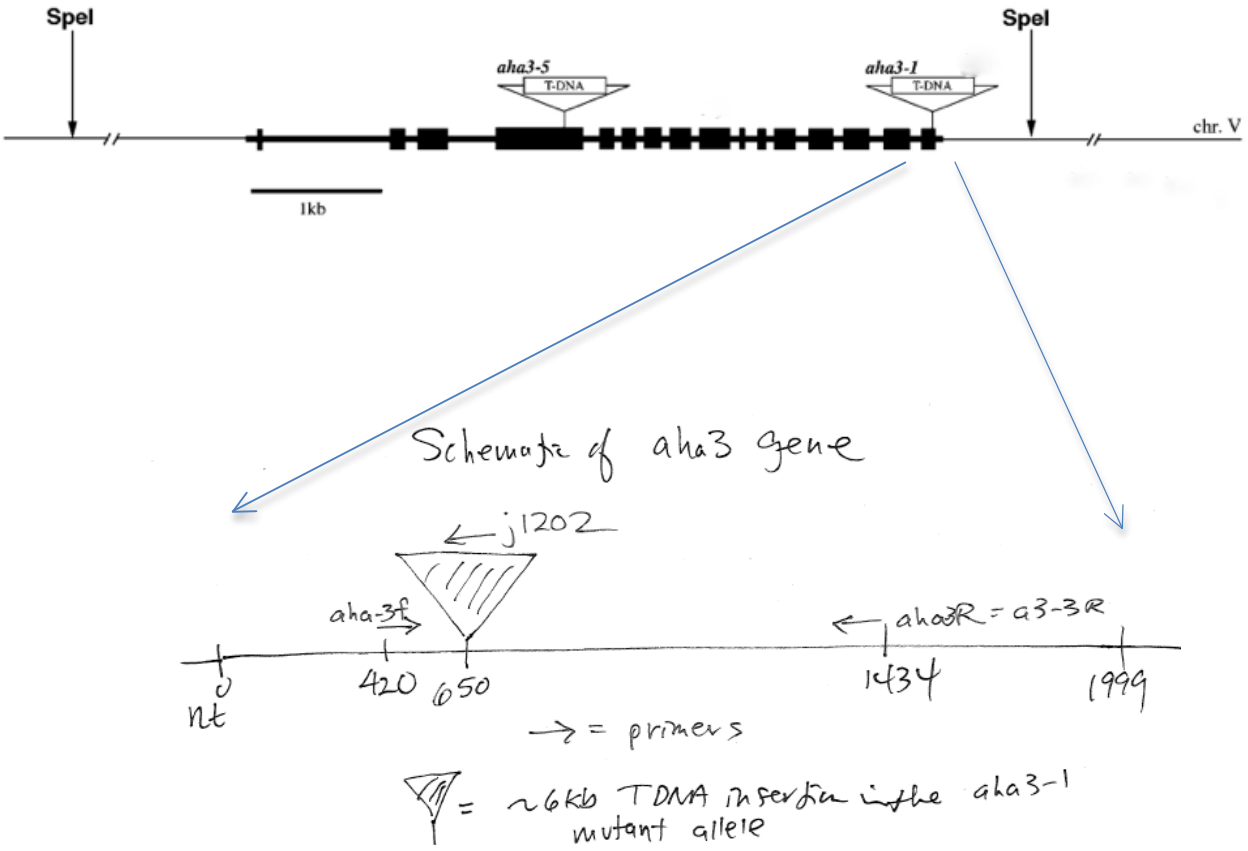
Specificity of amplification is controlled by the primers added to the reaction –WHY?



What are the components of a PCR Reaction?

T-DNA insert in the aha-3 gene

Structures of *aha3 gene and mutant* alleles. Boxes represent exons; lines represent introns. T-DNA is not drawn to scale.



T-DNA inserts in aha3-1 span about 6 kb. The nucleotide scale of the map directly above corresponds to the wildtype gene copy (3 'end only as indicated by the arrows). The T-DNA insert site is at nucleotide 650 in the wildtype sequence and the 5' end of the T-DNA-specific jl202 primer is at nucleotide 814 in the aha3-1 mutant sequence. The 5' end of the aha3f primer corresponds to nucleotide 420 and the 5' end of aha3r to nucleotide 1434.

PCR master mix Recipe

Fill in blanks in this table and tape it into your lab notebook

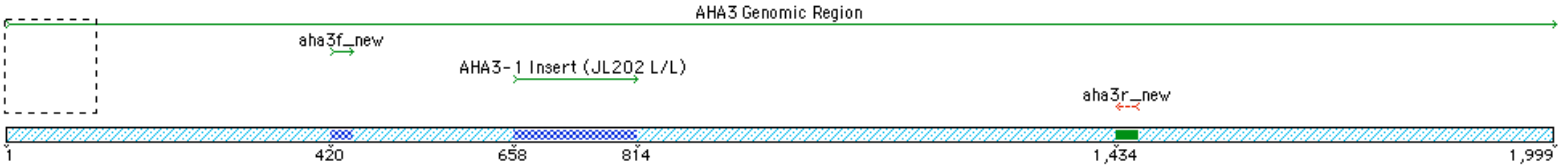
Master Mix 1: primers aha3f and a3-3R
Master Mix 2: primers aha3f and jl202

Reagent (Stock conc.)	vol per 100 Rxns	vol per one 25 µl reaction: 23µl of master mix + 2µl DNA	final conc
H_20 (to make 23 µl per reaction)		14.3 µl	
10X Taq Buffer		2.5	
dNTP's (10 mM)		2.0	
$MgCl_2$ (25mM)		1.5	
primer 1 (10 µM) aha3f		1.25	
primer 2 (10 µM) either a3-3R or jl202		1.25	
Taq polymerase 5U/µl		0.2 µl	

Setting up PCR Reactions:

Primer Pairs:
Aha3f and a3-3R (aka Aha3r): amplify wildtype allele only (why?)
Aha3f and jl202: amplify mutant allele only (why?)

Primer	# bases	Sequence
aha3f	29	CAC AAA GGA CTT TAC ACG GTC TTC AGA AC
jl202	29	CAT TTT ATA ATA ACG CTG CGG ACA TCT AC
a3-3R	29	GTC GTG GTG TGA AGA TTT ACA ACA GAT TG



Aha3-1: a single TDNA insert is about 6kb. The aha3 mutant allele has at least two tandem inserts oriented in opposite orientations.

Genotype	**Primer Pairs**	
	aha3f + jl202	aha3f + a3-3R
+/+		
+/ *aha3-1*		
aha3-1/aha3-1		

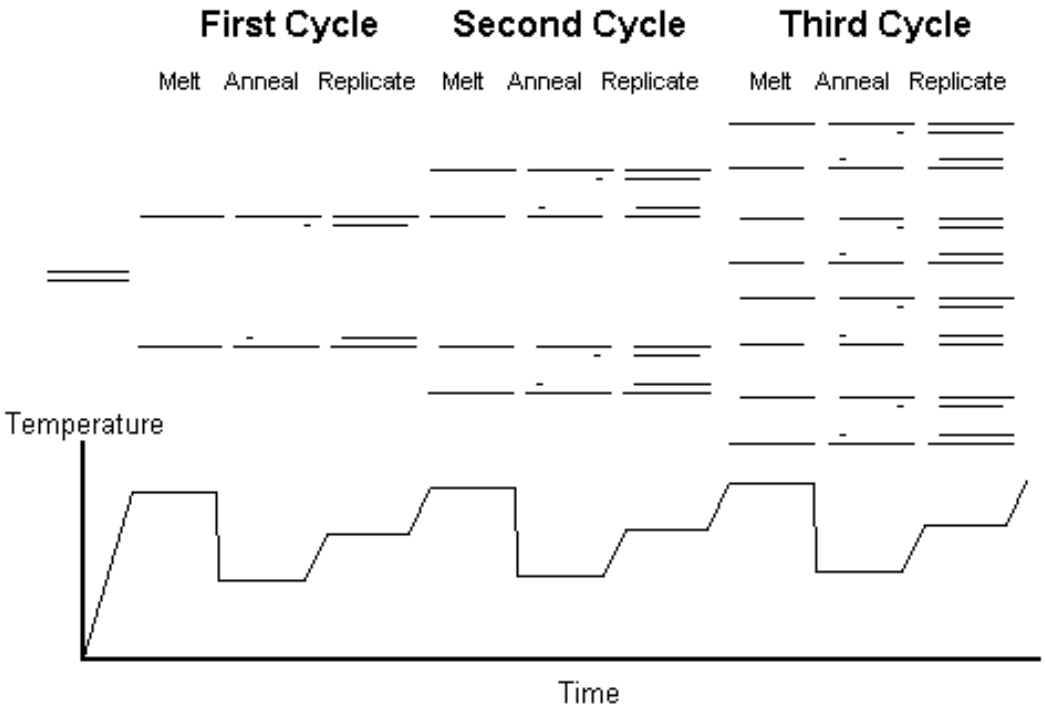
PCR animations
http://www.dnalc.org/ddnalc/resources/animations.html

PLEASE note the position of your tubes after you place them in the thermocycler

The thermocycler will be programmed as follows:

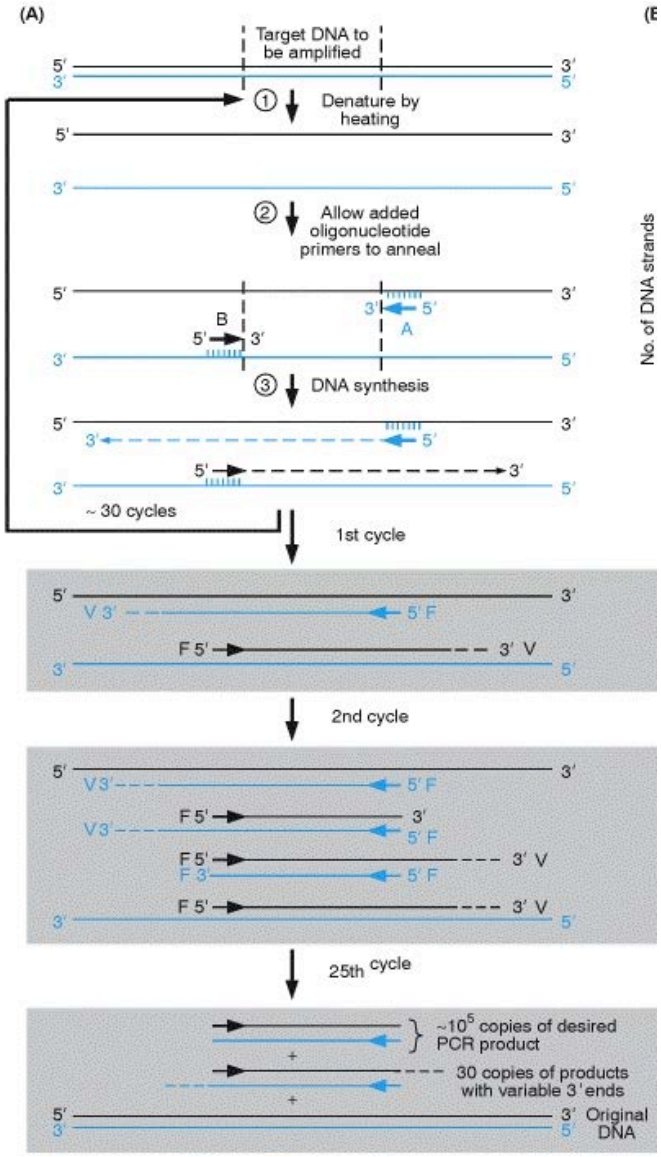
Step	Time/temp	Purpose? What is happening at this step?
1	5 min 94ºC	
2	1 min 90ºC	
3	1 min 60ºC	
4	1 min 72ºC	
5	Repeat steps 2-4 39 more times	
6	10 min 72ºC	
7	Hold 4ºC	

See next pg for more detailed schematic



What is temperature scale in ^{o}C?

Melt = denature DNA with heat
Anneal = allow primer to hydrogen bond with complementary sequences on the template DNA
Replicate = allow DNA polymerase to extend primer and synthesize complementary copy of template


(A)
Target DNA to be amplified
5′
3′
3′
5′
① Denature by heating
② Allow added oligonucleotide primers to anneal
B
A
③ DNA synthesis
~ 30 cycles
1st cycle
V
F
2nd cycle
25th cycle
~10^5 copies of desired PCR product
30 copies of products with variable 3′ ends
Original DNA
No. of DNA strands