



# From Linkage to Gene Detection

Ben Hayes and Mike Goddard



## Why detect QTL?

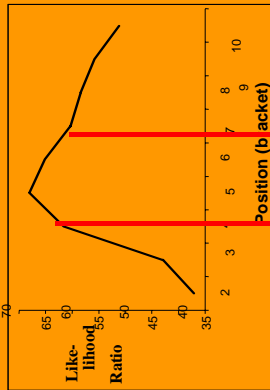
- Use markers linked to QTL in MAS
  - > genetic gain (esp. hard to select for traits)
- Use markers/ marker haplotypes in LD with QTL in MAS
  - >> genetic gain
- Find genetic mutation underlying QTL effect
  - patent = \$\$\$\$ (maybe)

## Pathway for gene mapping

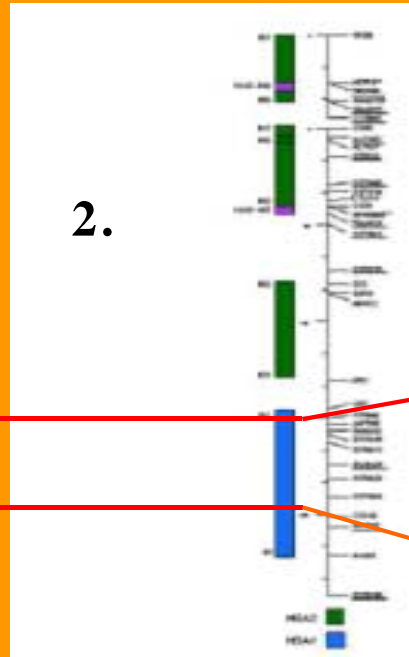
- If human genome ~ 3000cM long
- And there are between 30 000 and 100 000 genes
- 3cM segment contains between 30 and 100 genes
  - but don't know what these genes are in livestock, no livestock genome project completed (yet)
  - even if we did know what genes were, too many genes to evaluate/sequence for causative mutation

# Pathway for gene discovery

1.



2.



3.

*25DAP 4EBP1  
6P2ase ABF1  
ACT ACTH actin  
ADCC adrenaline  
Af1 AF2 Ag  
aIbb3 AKT AP1  
Apaf1 Apaf3*

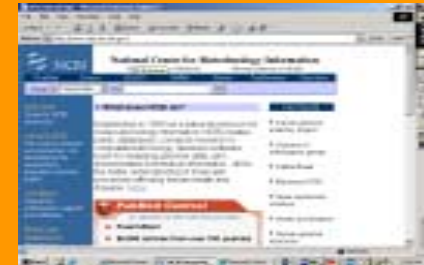
6.

**BLAST**

5.

ACTGGGTCCG  
ACTGGGACCG

4.

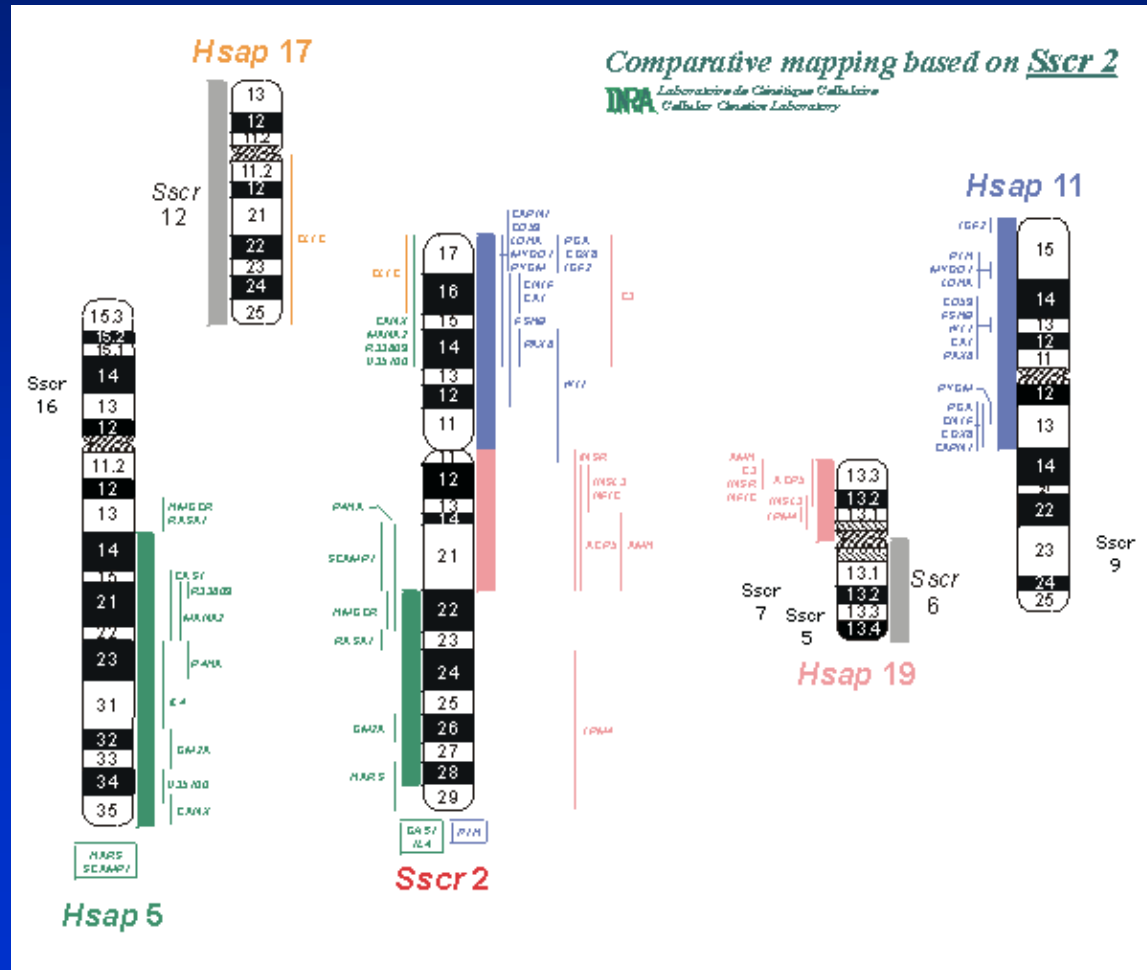


# Comparative mapping with humans

- Following human genome project, great deal of information about human genes
  - map livestock chromosome segment (C.I.) to corresponding human chromosomes
  - radiation hybrids, bi-directional painting

# Comparative mapping with humans

## Pig Chromosome 2



# Selecting candidate genes in an interval

- For the segment of human chromosome, can return a list of genes on that segment

The screenshot displays the NCBI Entrez Genome Map Viewer interface. At the top, there are navigation links for PubMed, Entrez, BLAST, OMIM, Taxonomy, and Structure. A search bar is present with the text "Find in This View" and "Find". Below the search bar, the page title is "Homo sapiens Map View build 31" with a link to "BLAST the Human Genome".

The main content area shows the chromosome selection: "Chromosome: [ 1 ] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y". Below this, there are tabs for "Master Map: Genes On Sequence" and "Maps & Options". The text indicates "Total Genes On Chromosome: 2430 [5 not localized]" and "Region Displayed: 0-244M bp" with a link to "Download/View Sequence/Evidence". It also states "Genes Labeled: 20 Total Genes in Region: 2425".

The gene list is presented in a table with columns for "Contig Uni...", "Genes\_seq", "symbol", "orient.", "links", and "evidence cyto.". The genes listed are:

Contig Uni...	Genes_seq	symbol	orient.	links	evidence cyto.
		<a href="#">MASP2</a>	+	<a href="#">sv</a> <a href="#">ev</a> <a href="#">hm</a> <a href="#">seq</a> <a href="#">mm</a>	C 1p36.3-
		<a href="#">CAPZB</a>	+	<a href="#">sv</a> <a href="#">ev</a> <a href="#">hm</a> <a href="#">seq</a> <a href="#">mm</a>	C 1p36.1
		<a href="#">PAFAH2</a>	+	<a href="#">sv</a> <a href="#">ev</a> - <a href="#">seq</a> <a href="#">mm</a>	C 1p34.3
		<a href="#">KHDRBS1</a>	+	<a href="#">sv</a> <a href="#">ev</a> <a href="#">hm</a> <a href="#">seq</a> <a href="#">mm</a>	C 1p32
		<a href="#">FLJ12666</a>	+	<a href="#">sv</a> <a href="#">ev</a> <a href="#">hm</a> <a href="#">seq</a> <a href="#">mm</a>	C 1p34.2
		<a href="#">KIAA1511</a>	+	<a href="#">sv</a> <a href="#">ev</a> <a href="#">hm</a> <a href="#">seq</a> <a href="#">mm</a>	C 1p34.1
		<a href="#">FLJ40201</a>	+	<a href="#">sv</a> <a href="#">ev</a> - <a href="#">seq</a> <a href="#">mm</a>	C 1p32.2
		<a href="#">MGC27382</a>	+	<a href="#">sv</a> <a href="#">ev</a> - <a href="#">seq</a> <a href="#">mm</a>	C 1p31.1
		<a href="#">FLJ10287</a>	+	<a href="#">sv</a> <a href="#">ev</a> - <a href="#">seq</a> <a href="#">mm</a>	C 1pter-q
		<a href="#">MOV10</a>	+	<a href="#">sv</a> <a href="#">ev</a> <a href="#">hm</a> <a href="#">seq</a> <a href="#">mm</a>	C 1p13.1
		<a href="#">ERp60</a>	+	<a href="#">sv</a> <a href="#">ev</a> - <a href="#">seq</a> <a href="#">mm</a>	C 1q21

## Selecting candidate genes in an interval

- Literature search (PubMed, etc) on key words (eg. “Lactation”)
  - record number of ‘hits’ to narrow list of candidates
- Next step: sequence heterozygous sires for candidate gene (or part thereof)
  - sequence of gene from NCBI
  - BLAST search to determine if any SNPS are function, eg change amino acids or insert stop codon

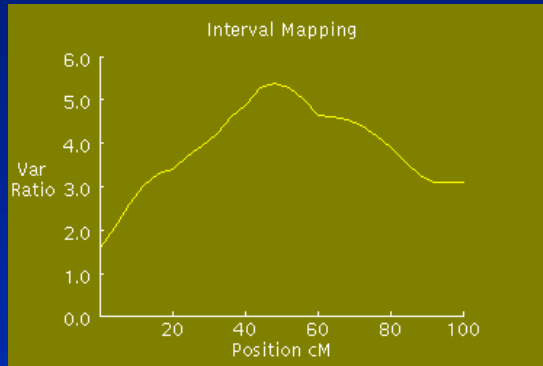


## Selecting candidate genes in an interval

- Example of Inverdale gene (Galloway et al 2001)
  - July 2001, AgResearch (NZ) reported discovery of a mutation affecting litter size in sheep
  - Single copy of mutation in heterozygous ewes increased litter size by 0.6 lambs
  - Double copy small non-functional ovaries, infertile

# Selecting candidate genes in an interval

- Example of Inverdale gene (Galloway et al 2001)



X chromosome

Candidate gene  
GDF9, selected,  
based on Inverdale  
like effects in  
mutant mice

But GDF9 not  
on X  
chromosome in  
humans, mice  
or sheep

Point mutation  
discovered, use in  
sheep breeding  
schemes

Alternative  
candidate,  
GDF9B, does  
map to X  
chromosome

## Selecting candidate genes in an interval

- Example of Inverdale gene (Galloway et al 2001)
  - aided by unambiguous phenotype
- Nevertheless, this example and particularly DGAT1, prove it is possible to find the mutations which underlie variation in livestock production traits