JEWELS IN JUNK DNA:

Using cross-species sequence comparisons to sift through the genome.

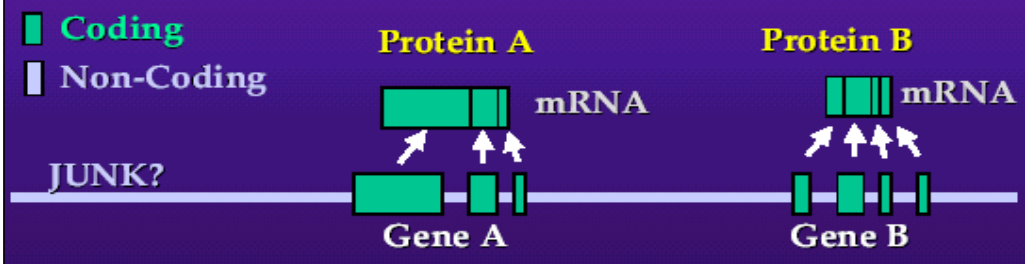
Marcelo A. Nobrega

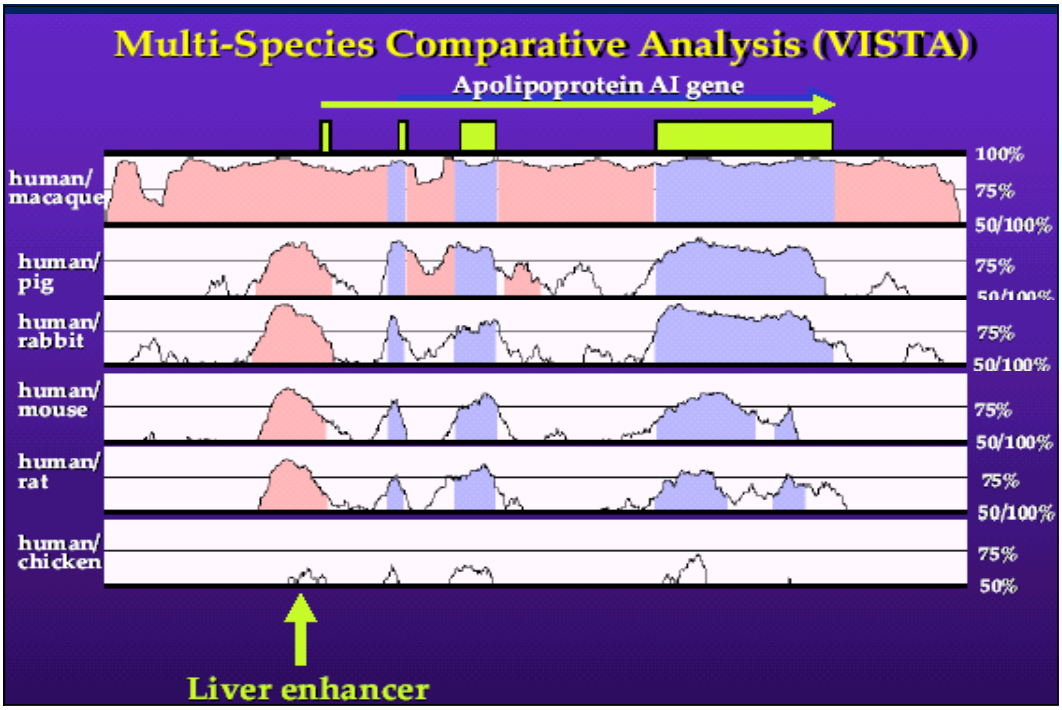
Genome Sciences Department
Lawrence Berkeley National Laboratory

Categories of DNA

Coding ~5%

Non-coding ~95%





Pufferfish (*Fugu rubripes*)

- genome 8 times smaller than human's
- Very little "junk DNA"
- _ of proteins have strong homology to human counterparts
- Several genomic segments conserved synteny with human orthologs
- Most recent common ancestry with humans at ~540 MY.

Aligning vertebrate genomes:

Step 1. Break Human chromosomes into 225kb pieces overlapping by 25kb

Step 2. Repetitive elements are masked in each human sequence

Step 3. Non-repetitive sequence is mapped to mouse and fugu genomes using external web tools

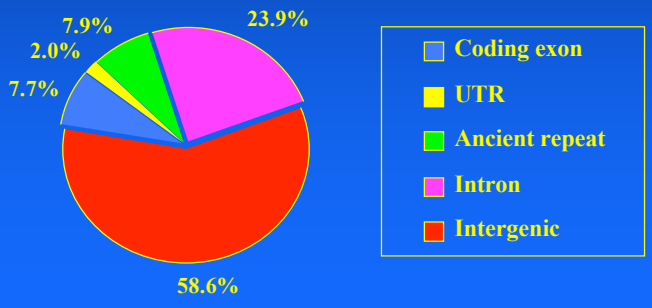
BLAT @ UCSC
http://genome.ucsc.edu/cgi-bin/hgBlat

BLAST @ JGI
http://aluminum.jgi-psf.org/prod/bin/blast.fugu6.cgi

Aligning Human and mouse genomes

10% of the sequence is highly conserved

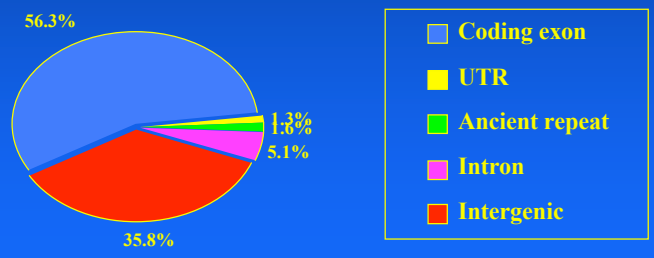
1,257,793 ECRs (Evolutionary Conserved Regions)



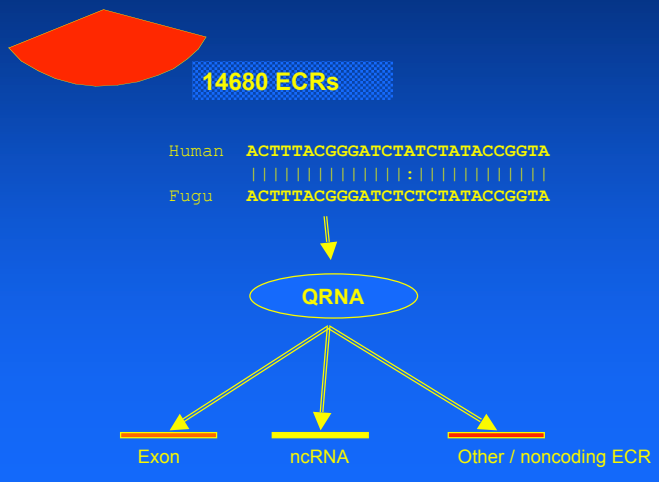
Aligning Human and fugu genomes

0.2% of the human, 1.7% of the fugu genomes are conserved

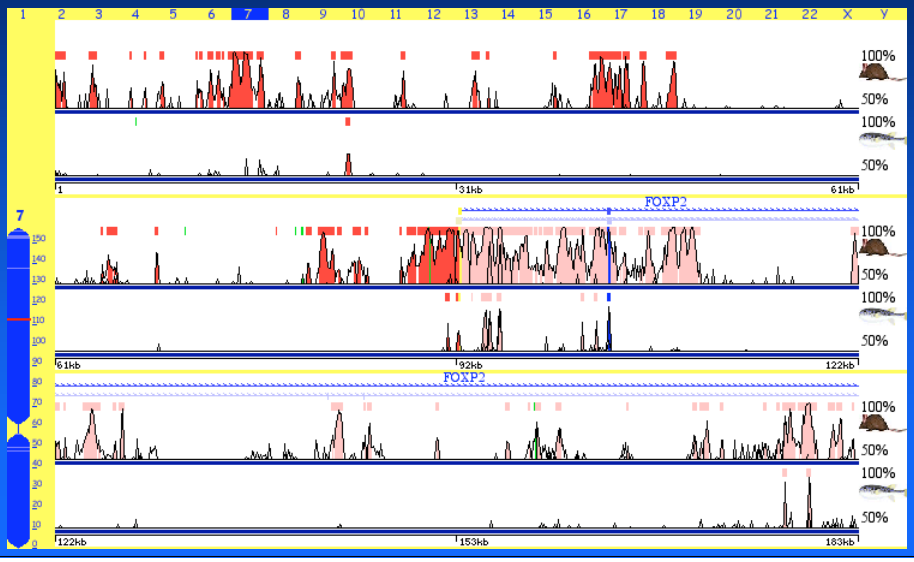
41,067 ECRs



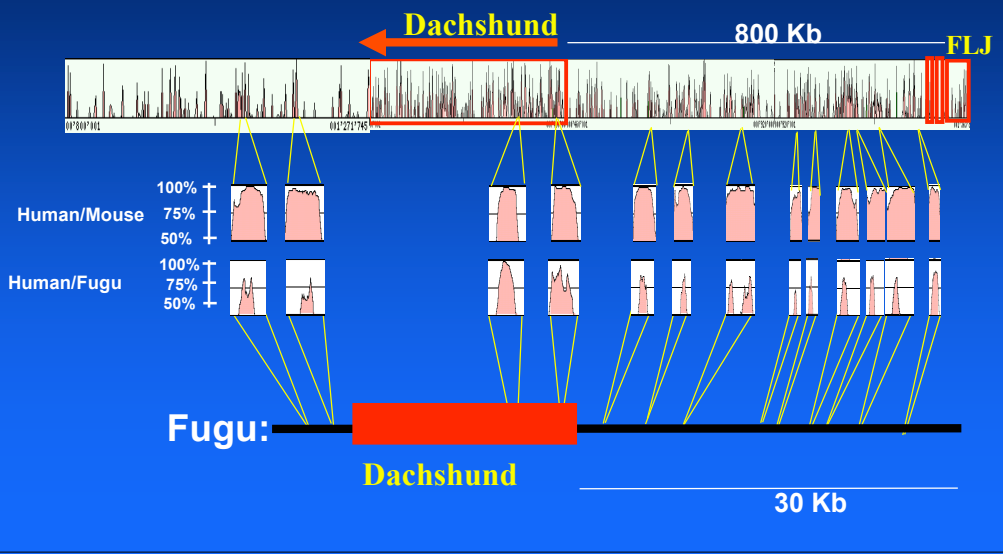
Human/Fugu 'noncoding' ECRs - qRNA analysis

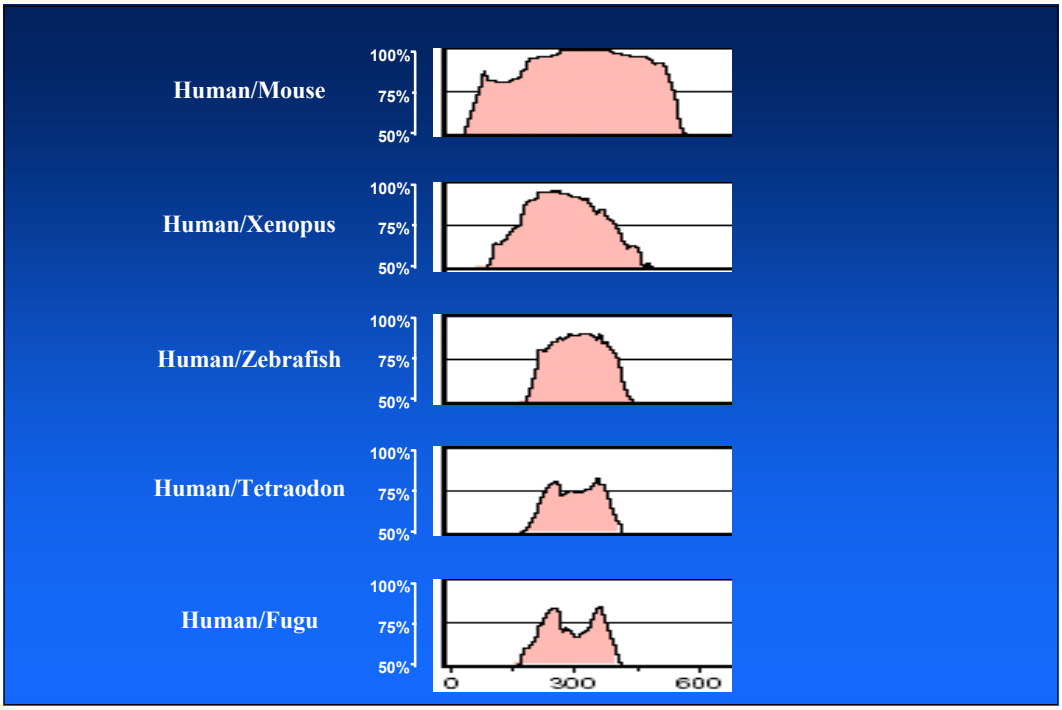


http://nemo.lbl.gov/~ovcharen/

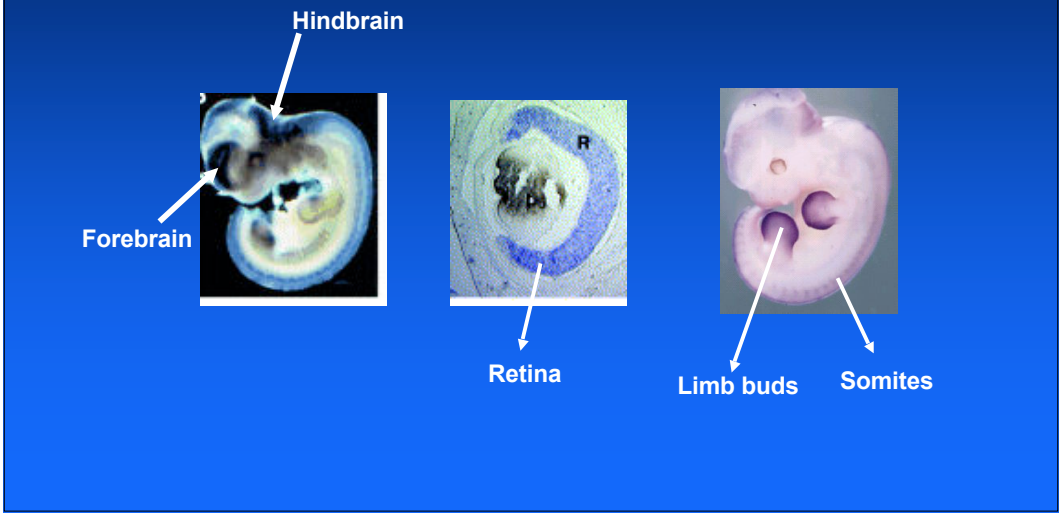


Identification of 13 distal CNSs around Dachshund

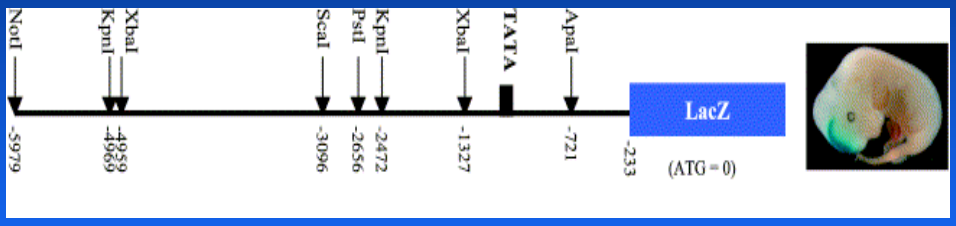




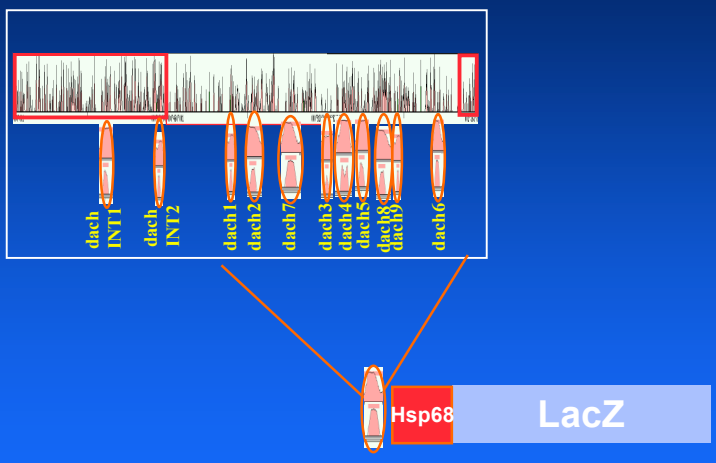
Dachshund expression



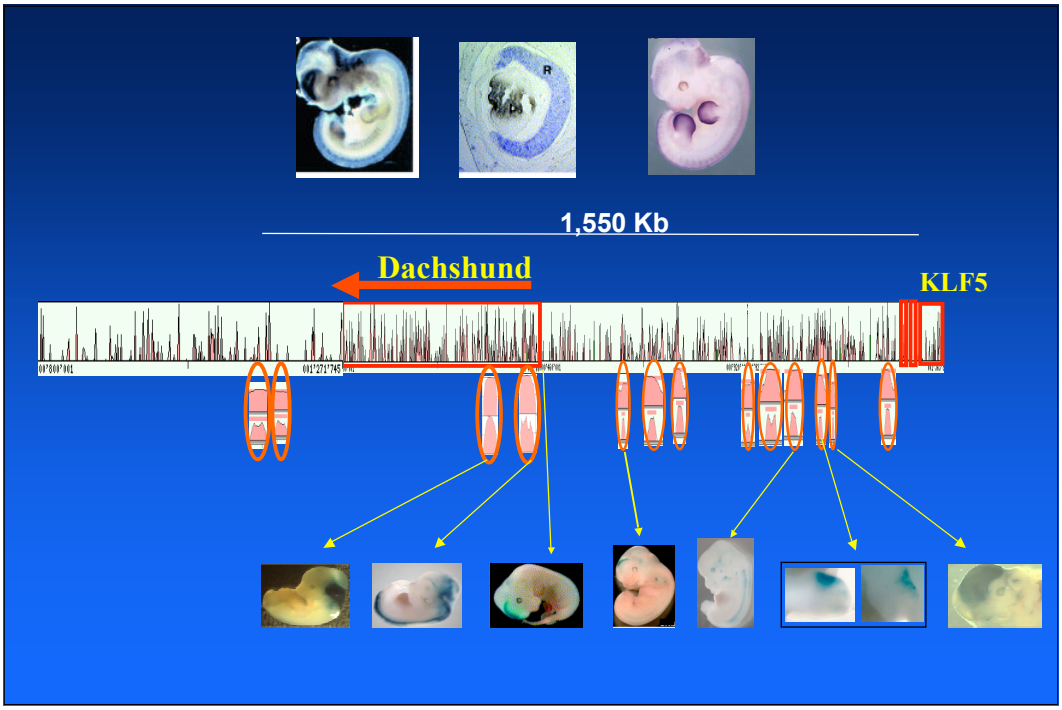
Identification of a Dach1 enhancer using lacZ transient transgenics



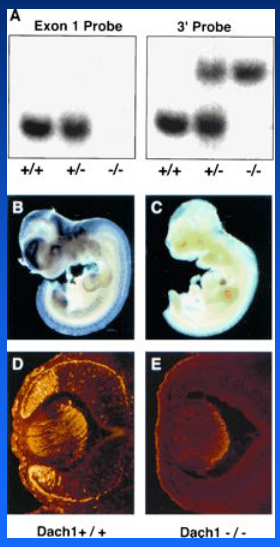
Testing the function of CNSs: Transient transgenics



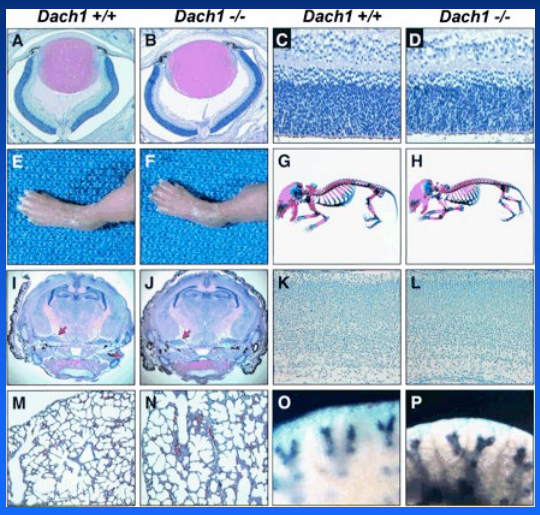
Inject into fertilized eggs → Stain for embryonic LacZ

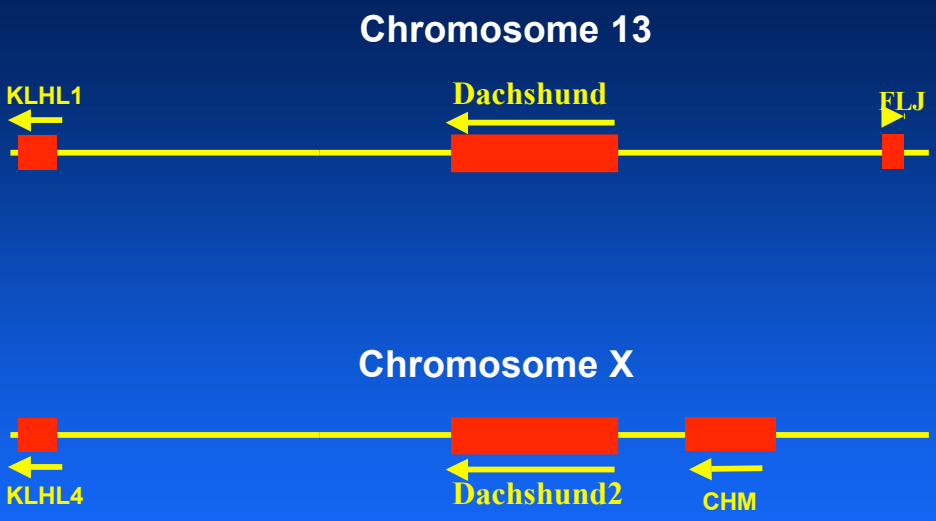


Dach1 Expression in wt and ko mice

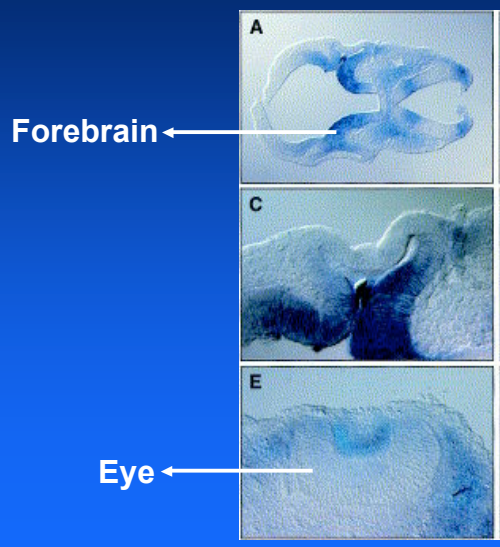


Dach1 ko does not result in gross morphologic/morphometric alterations





The expression pattern of Dach1 and Dach2 overlap

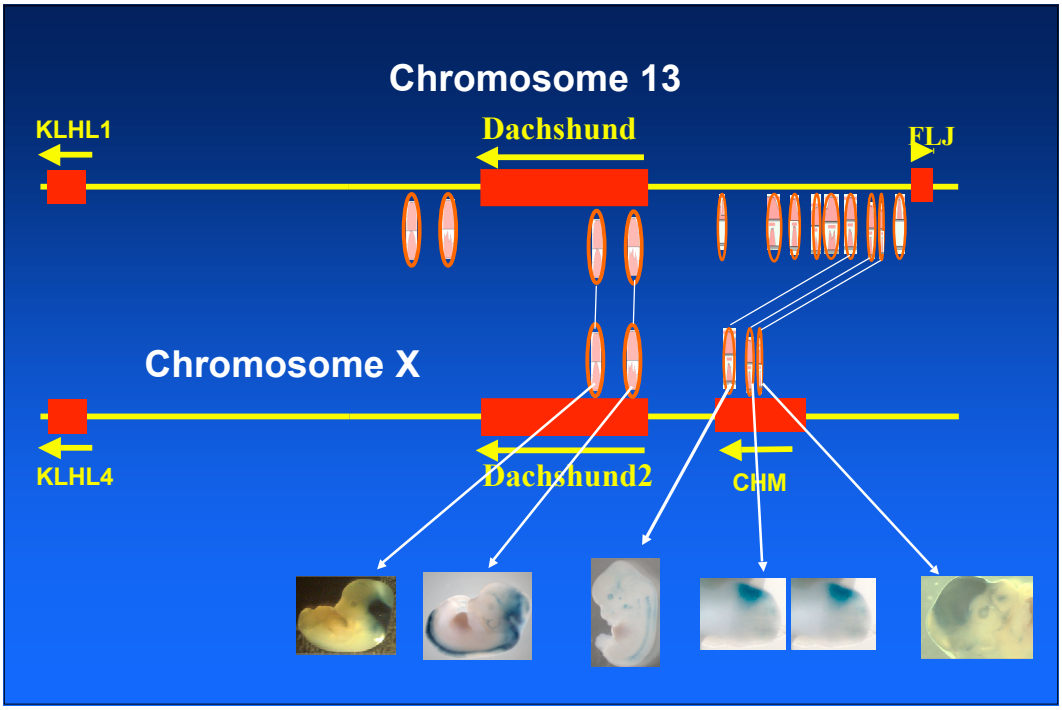


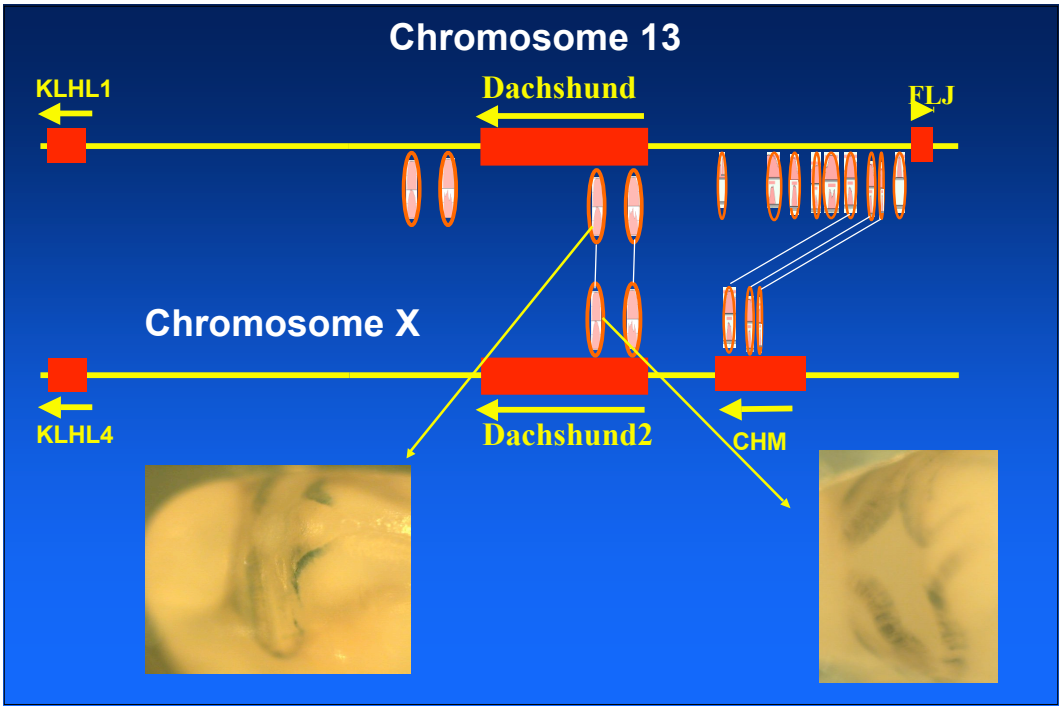
Forebrain

Hindbrain

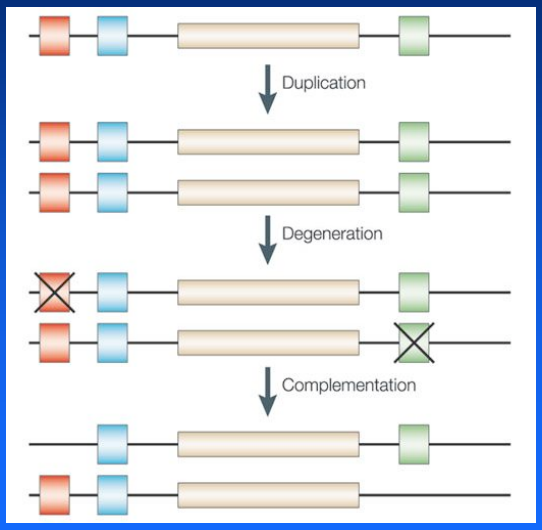
Eye

Limb buds





Fate of Duplicated Genes: Degeneration and Complementation



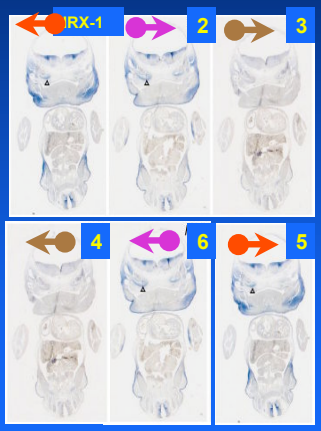
Duplication of regulatory elements: common theme or oddity?

Duplicated: **90** of 1739 noncoding

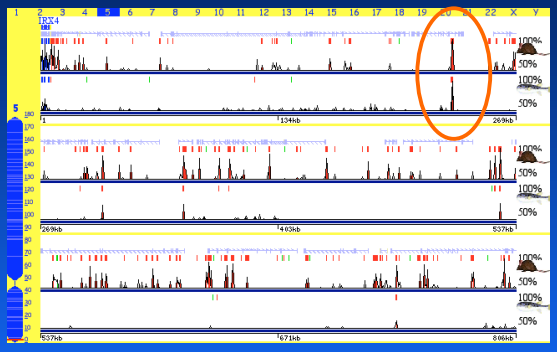
27 out of 842 intronic

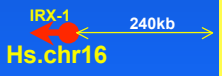
~ 4%

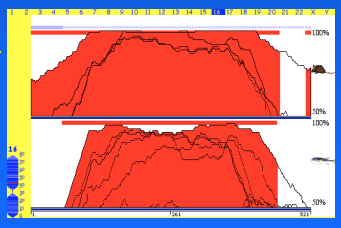
Iroquois Homeobox Gene Clusters



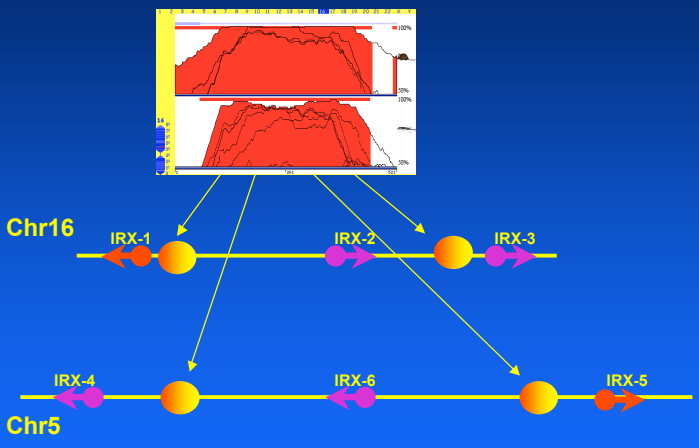
Duplication of a CNS Upstream of IRX-4



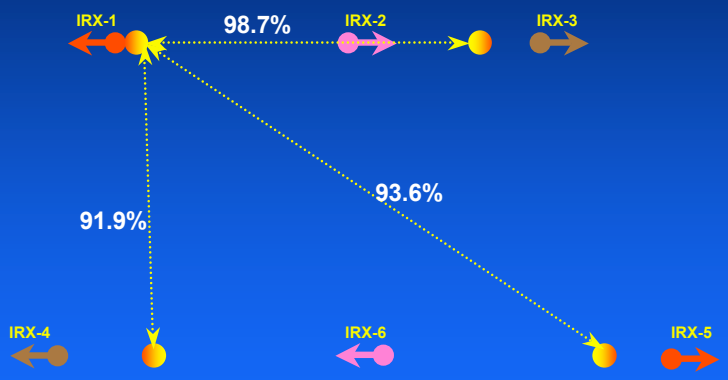




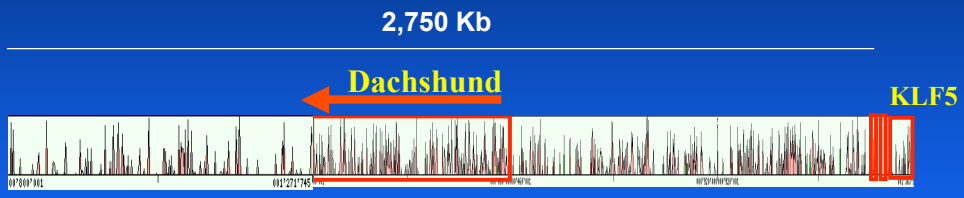
Duplicated CNS coregulating IRX genes?



Sequence-function Relationship of Duplicated CNSs



What are the gene deserts surrounding Dacshund?

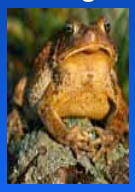
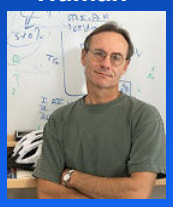


- 1-2% of DNA is translated
- 25-30% of DNA is transcribed
- 60-70% of the genome contains non-genic DNA

Why do we have a 2,900,000,000 bp Genome?

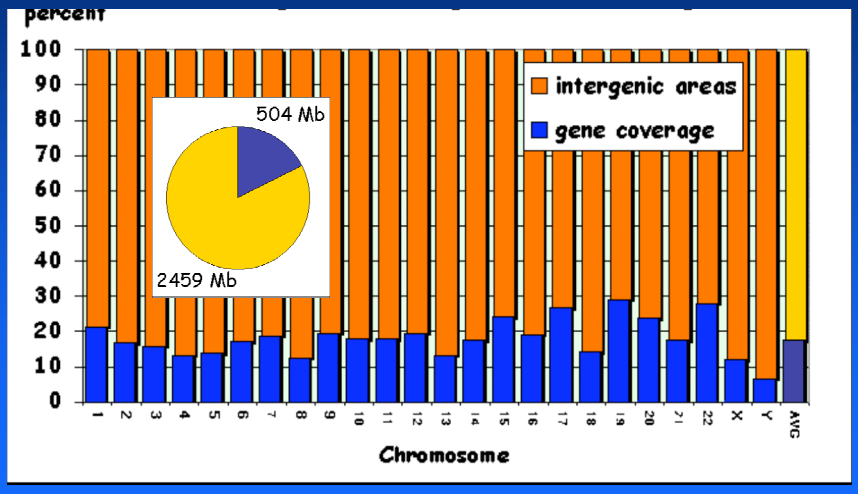
C-value Paradox

Genome size does not correlate with organismal complexity

Amoeba	Fruit-fly	Grasshopper	Frog	Mouse	Human
					
620,000	180	180,000	6,900	2,500	2,900

(Million bp)

Gene coverage of the human genome



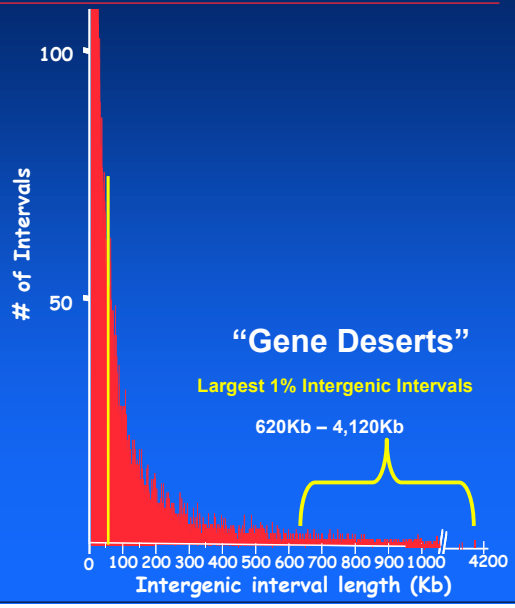
Defining "Gene Deserts"

1 – Calculate Intergenic Lengths

EMSEMBL- 21,978 genes

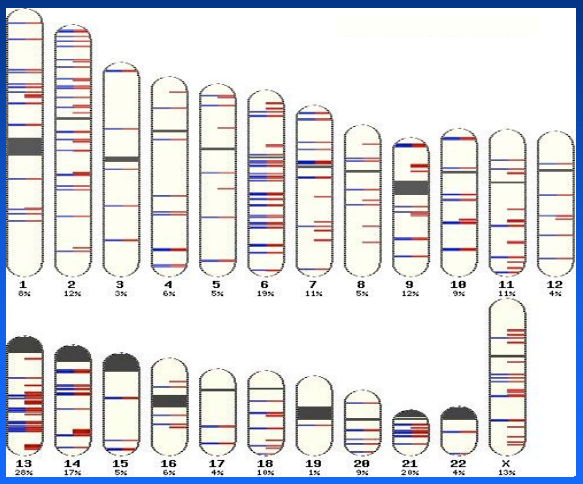
REFSEQ Annotation- 12,439 genes

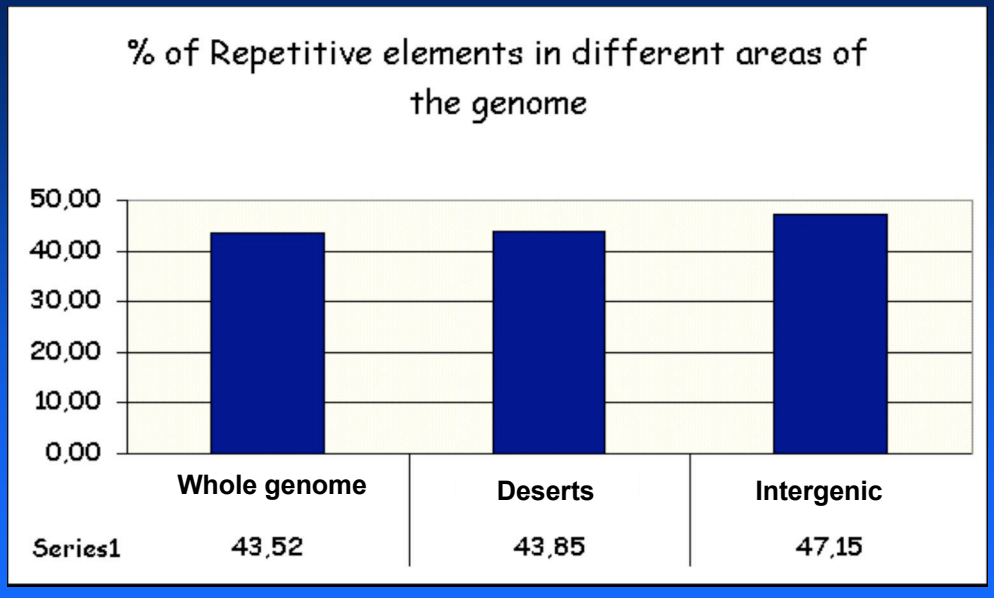
2 – Exclude Heterochromatic DNA & Clone Gaps.



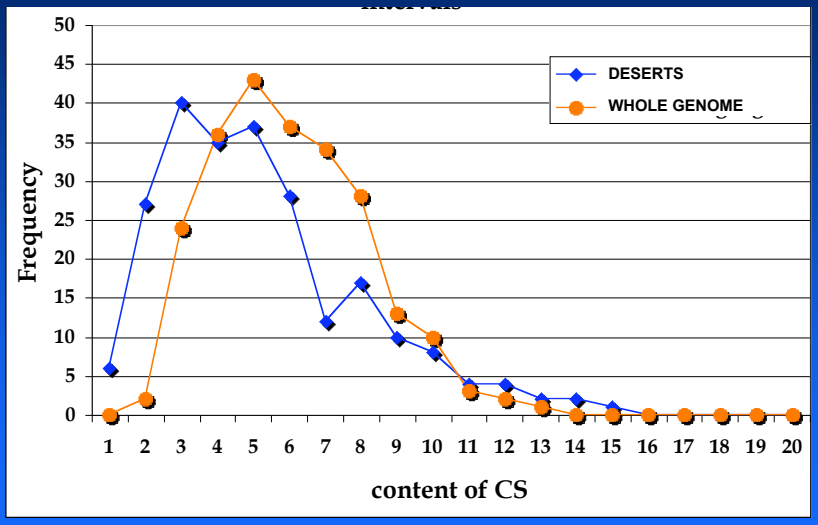
Human and Mouse Gene Deserts

- HUMAN
  - 234 Gene Deserts
- Ortholgous Mouse Comparison
  - 178 (74%) are also Deserts

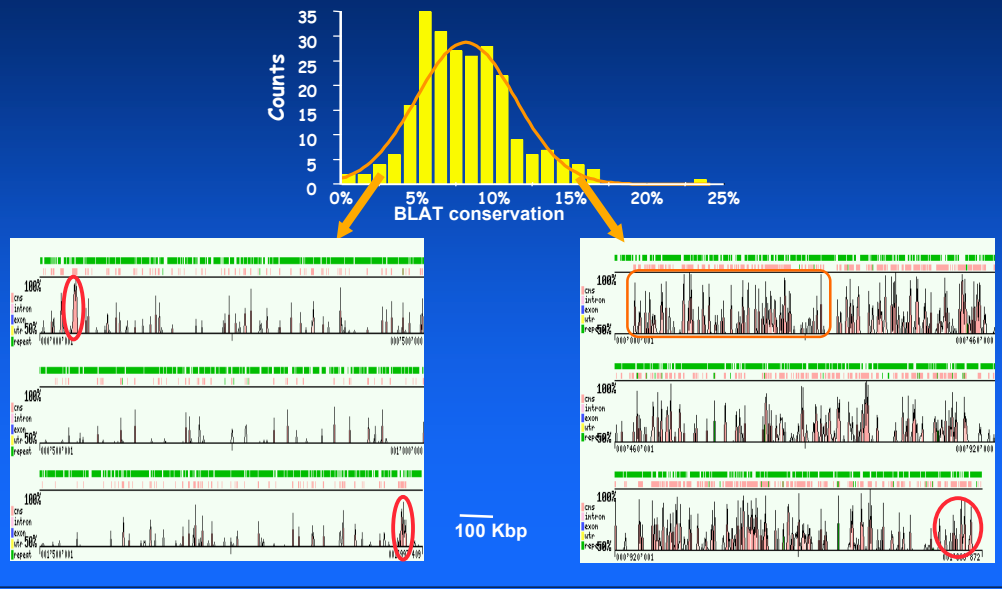




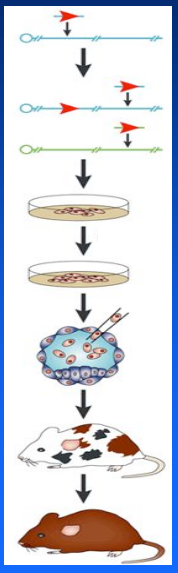
Conservation in Gene Deserts Compared to Whole Genome



Human-Mouse Conservation in Gene Deserts



Generation of Chromosomal Deletions in Mice



Insert loxP site into first end point

Insert loxP site into second end point on same chromosome

Select and isolate double-targeted clones

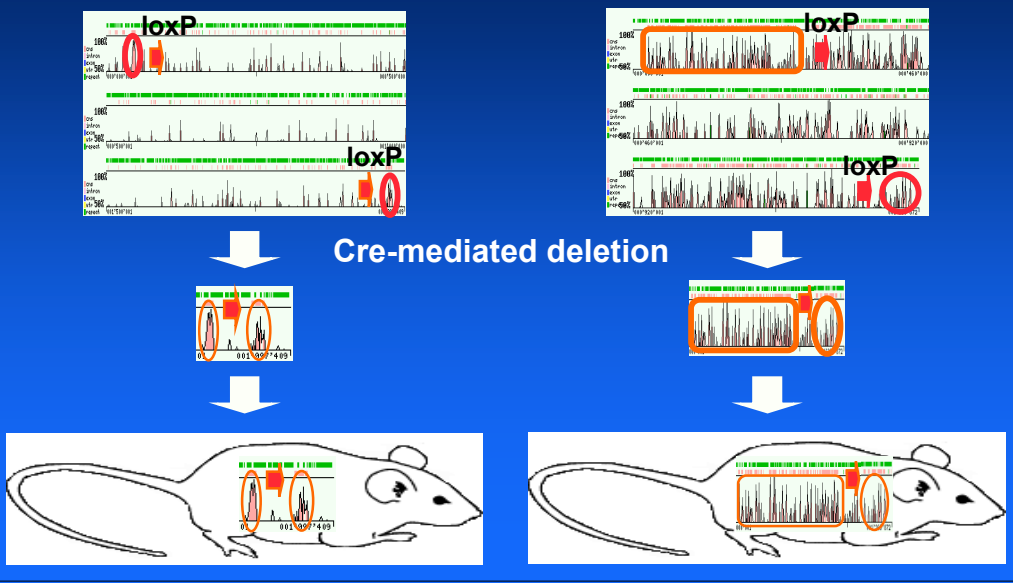
Select and Isolate the ES-cell clones with the desired rearrangements

Inject ES cells into blastocyst

Identify chimeric mice

Derive progeny carrying the engineered chromosome

Determining function of Gene Deserts



http://nemo.lbl.gov/~ovcharen/

LBNL – Rubin Lab.

- **Eddy Rubin**
- Ivan Ovcharenko
- Yiwen Zhu
- Gabby Loots
- Veena Afzal
- Ingrid Plajzer-Frick
- Michael Collier
- Jody Schwartz
- Len Pennacchio

UCSF
- Brian Black

Stanford University
- Catherine Gunther

Vanderbilt University
- Doug Mortlock