

Putative Enhancer Elements in Mouse: Computational Isolation and Experimental Testing

Steven Kuntz

Gene Background

- MOR 28 and 83 (mouse olfactory receptor)
 - Unknown promoter region
- Mrf4/Myf5 (Myogenic factors)
 - Muscle differentiation
- Hes1 (Hairy/enhancer of split)
 - Somite development
- Hey1 (Hairy/enhancer of split with YRPW motif 1)
 - Somite development
- Lfng (Lunatic fringe or β -1,3-N-

Part I: Computation

Background

- Comparative genomics
- Search for putative enhancer elements necessary for regulation
- Screen for interesting enhancer elements

Odorant Receptors

- Largest gene family in vertebrates
- Exclusive expression of given odorant receptor (OR) protein in any olfactory sensory neuron (OSN)
- Each group of OSNs expressing an identical OR extend axons to the same glomerulus (relay center)
- Mechanism of exclusive expression unknown, but several theories exist
 - Serizawa, et al. 2003 suggests negative feedback regulation
 - Chromatin structure modifications have been suggested
- Activating promoter not identified for OR families
 - Small upstream region appears to be crucial for expression

Family Relations Analysis

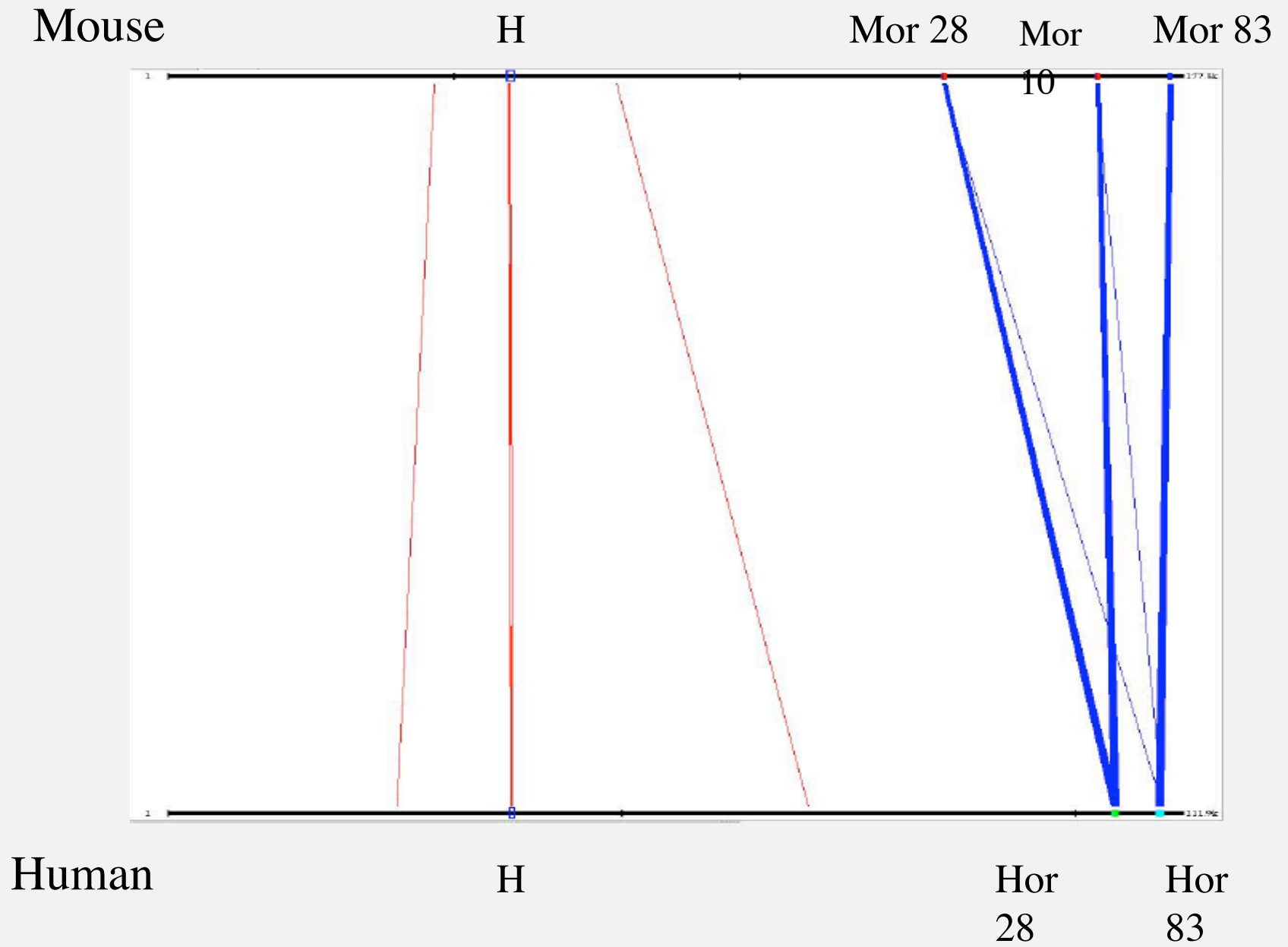
Serizawa, et al. 2003

- Mouse olfactory receptor conserved enhancer element



- Conserved in mouse, human
 - No known conservation in chicken
 - May be artifact of sequencing or avian olfaction
 - Birds rarely have behaviorally functional olfaction, suggesting that the genetic basis may be modified from mammals as well

87% Threshold

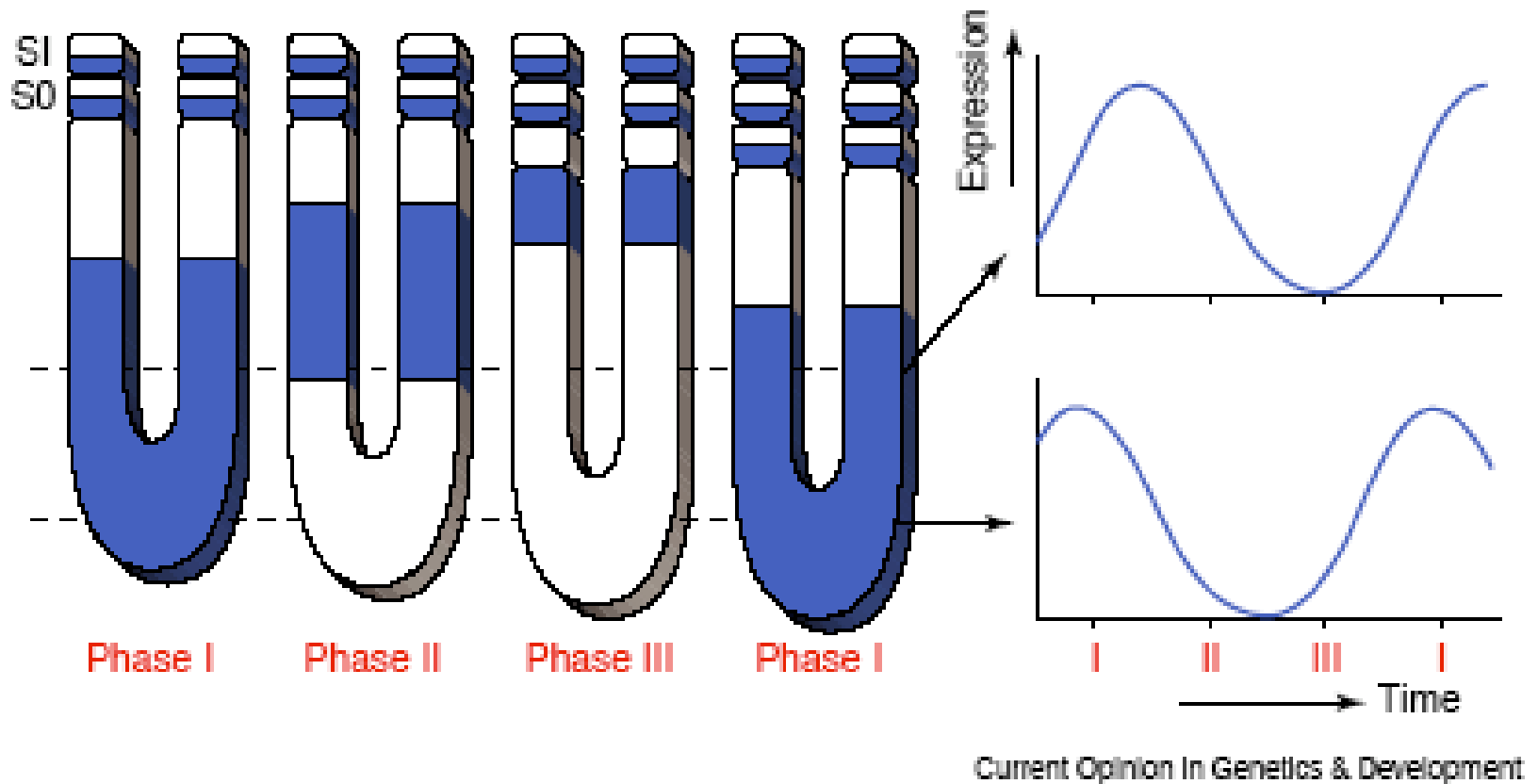


Somite Development

Palmeirim, et al. 1997

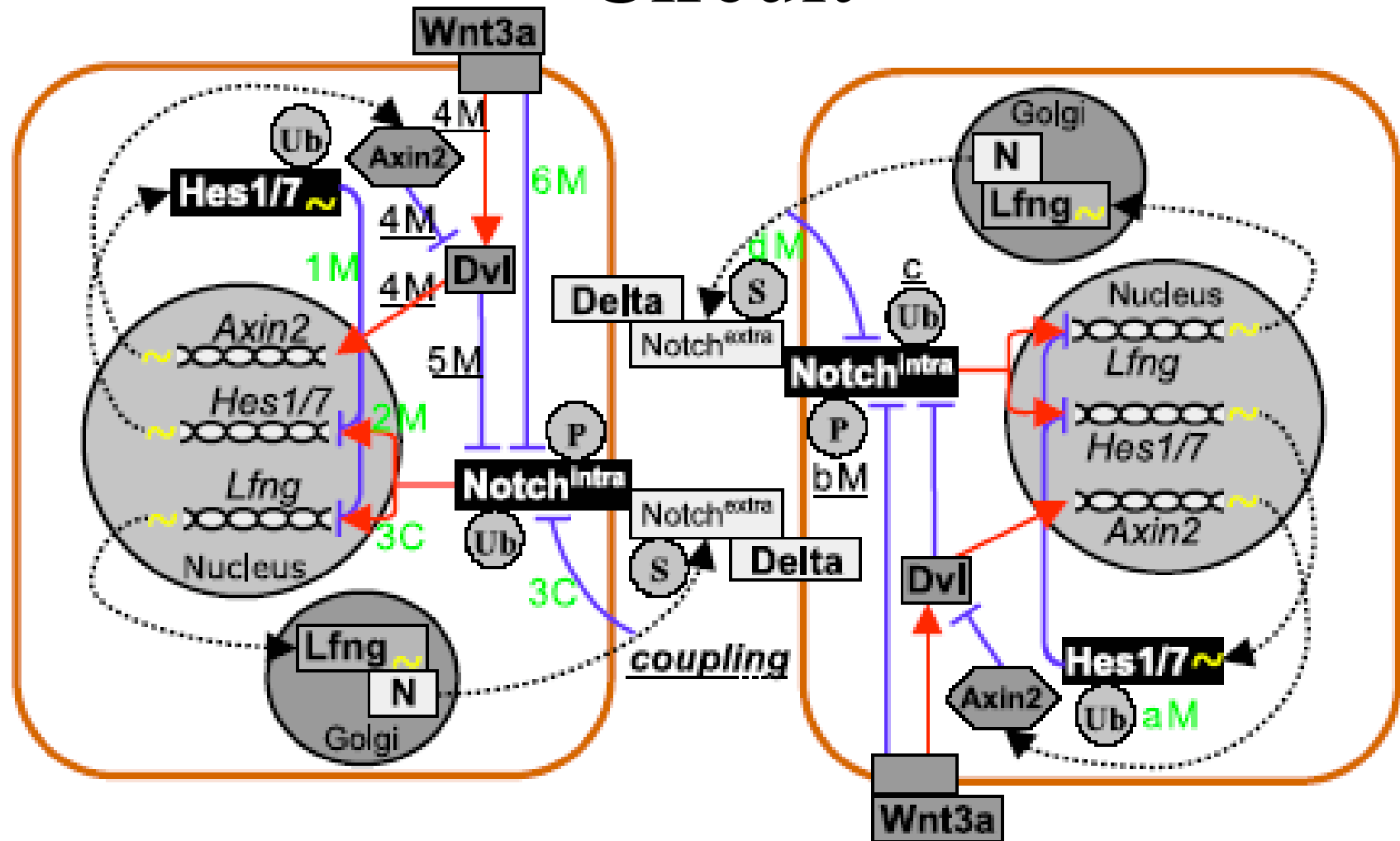
- Cyclic expression of chicken *hes1* (c-hairy) is necessary in somitic development
- 90 minute periodicity corresponds with somite formation in the PSM
- Multi-cellular gene circuit initiated by Notch signaling regulates *hes1* expression (along with lunatic fringe, *axin2*, *hes7*, etc)

Oscillating Expression



Bessho Y, Kageyama R. *Current Opinion in Genetics and Development*, 13:379-384 2003.

Complex Oscillatory Feedback Circuit



Rida P, Minh N, Jiang Y. *Developmental Biology* 265:2-22 2004.

MUSSA Analysis

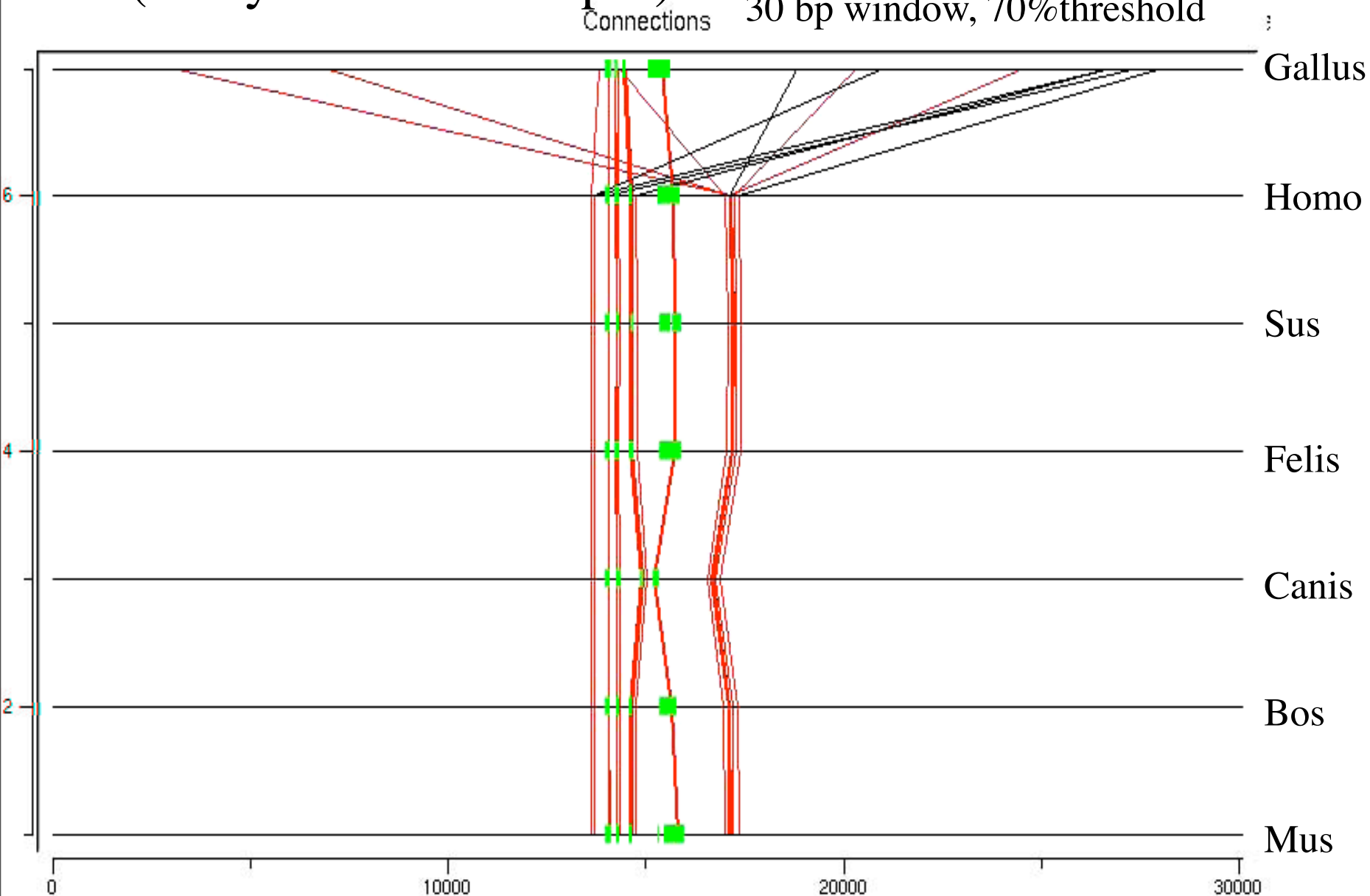
- Smaller scale analysis than FRII, but higher dimension
- Chicken loci determined using FRII analysis
 - Prior to sequence annotation
 - Utilized conservation of nearest neighbors
- Gene coding regions approximate
 - Based on sequence blast and FRII using mouse coding region
- Hes1
 - 7 vertebrates available
- Hey1

Hes1

(Hairy/enhancer of split)



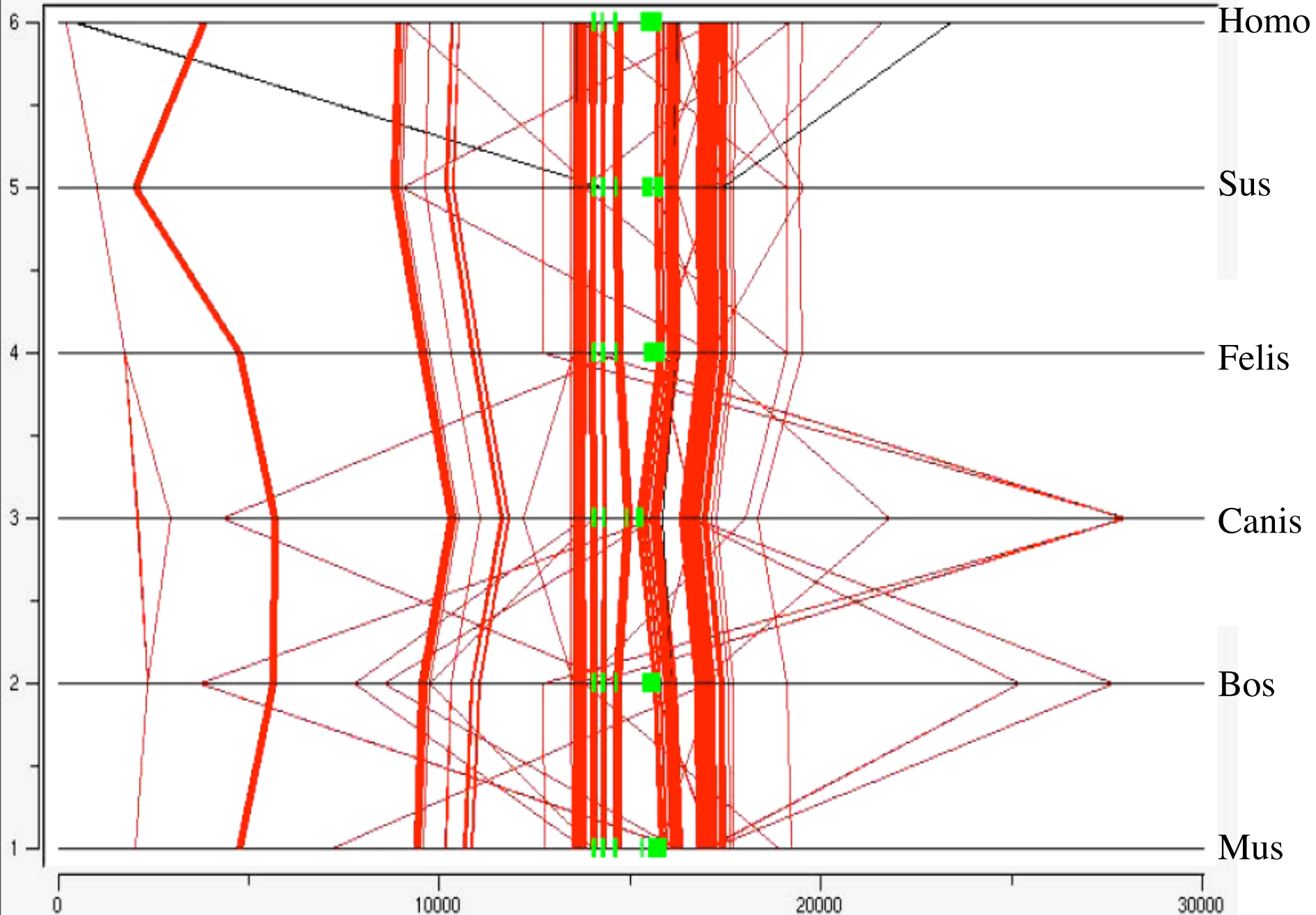
7-way comparison,
30 bp window, 70% threshold



Hes1

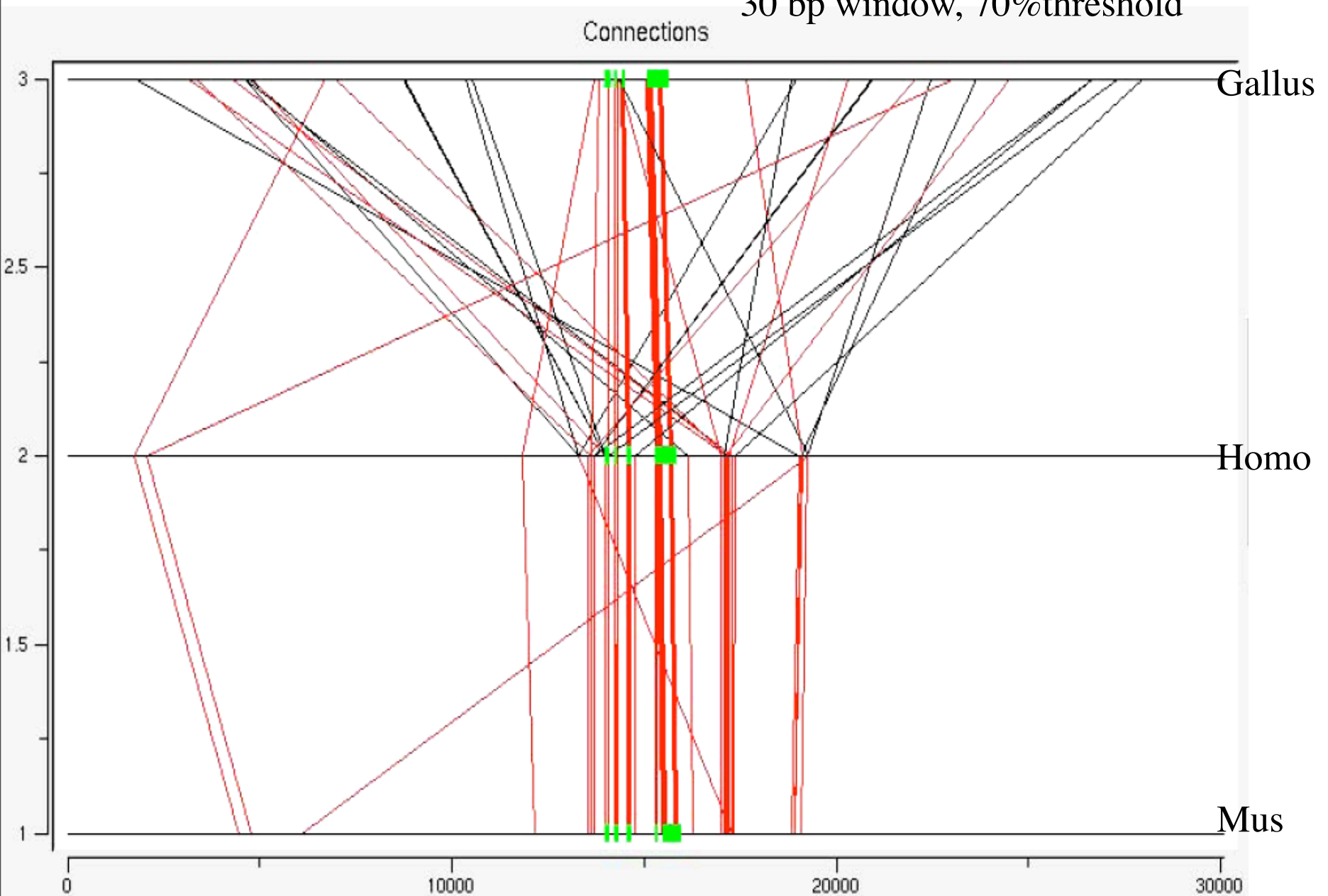
Connections

6-way mammalian comparison,
30 bp window, 70%threshold



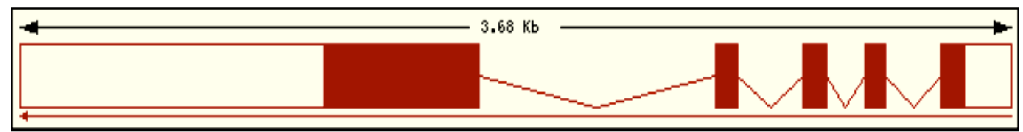
Hes1

3-way comparison,
30 bp window, 70% threshold

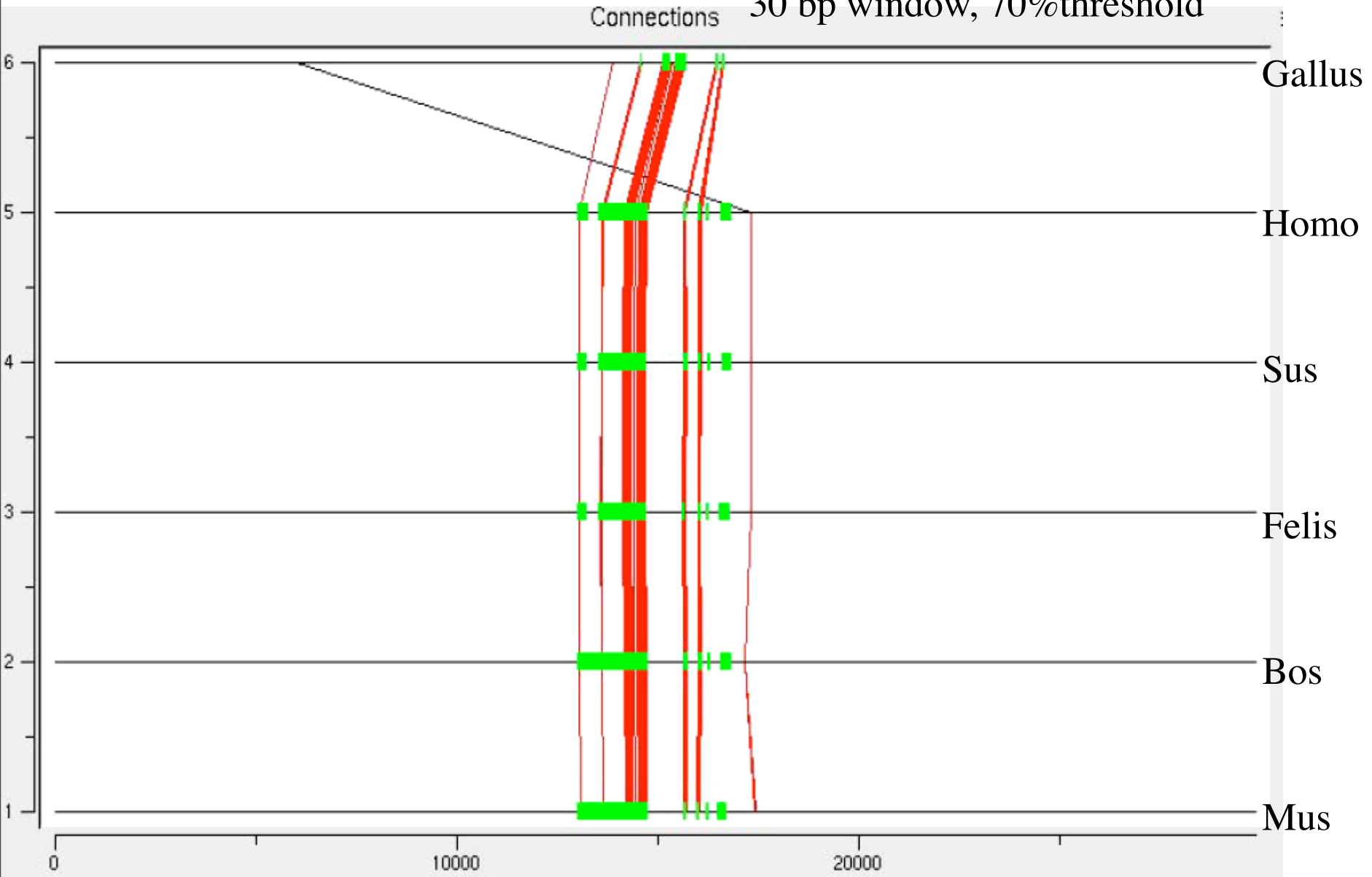


Hey1

(Hes with YPRW motif 1)



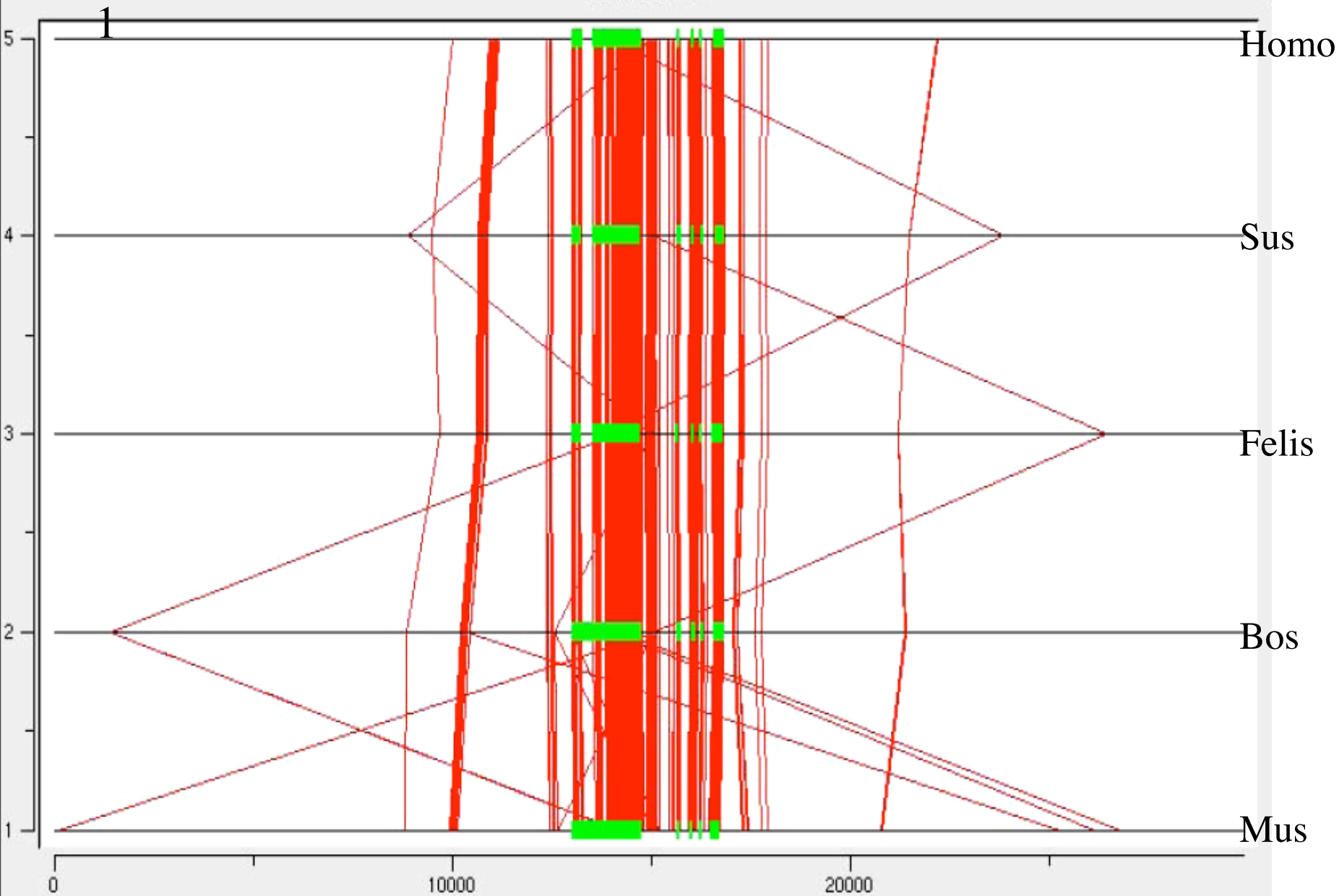
6-way comparison,
30 bp window, 70% threshold



Hey

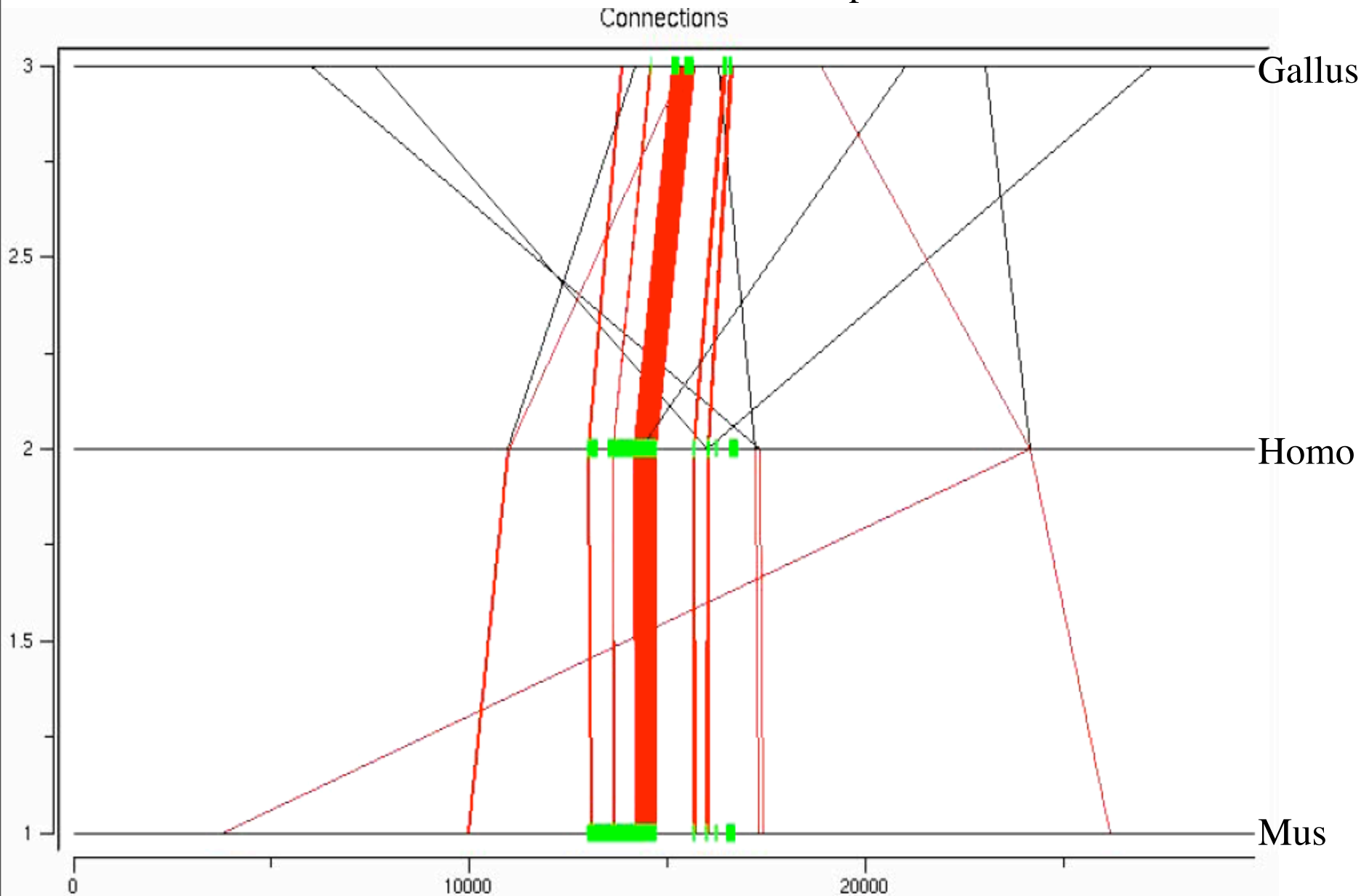
5-way mammalian comparison,
30 bp window, 70%threshold

Connections



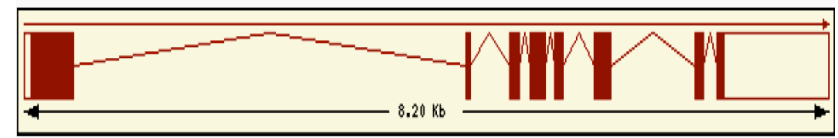
Hev1

3-way comparison,
30 bp window, 70%threshold

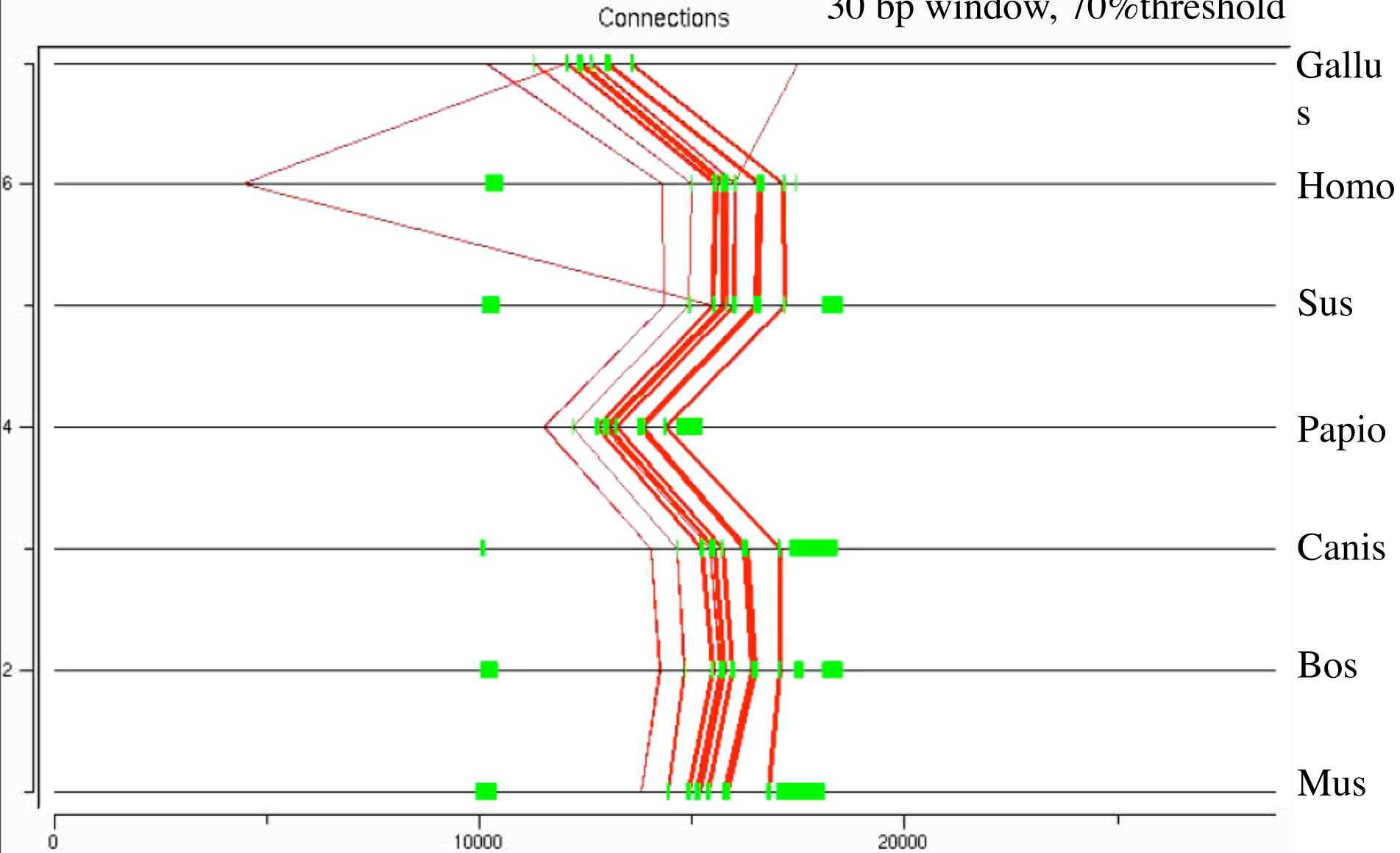


Lunatic Fringe

(β -1,3-N-Acetylglucosaminyltransferase)

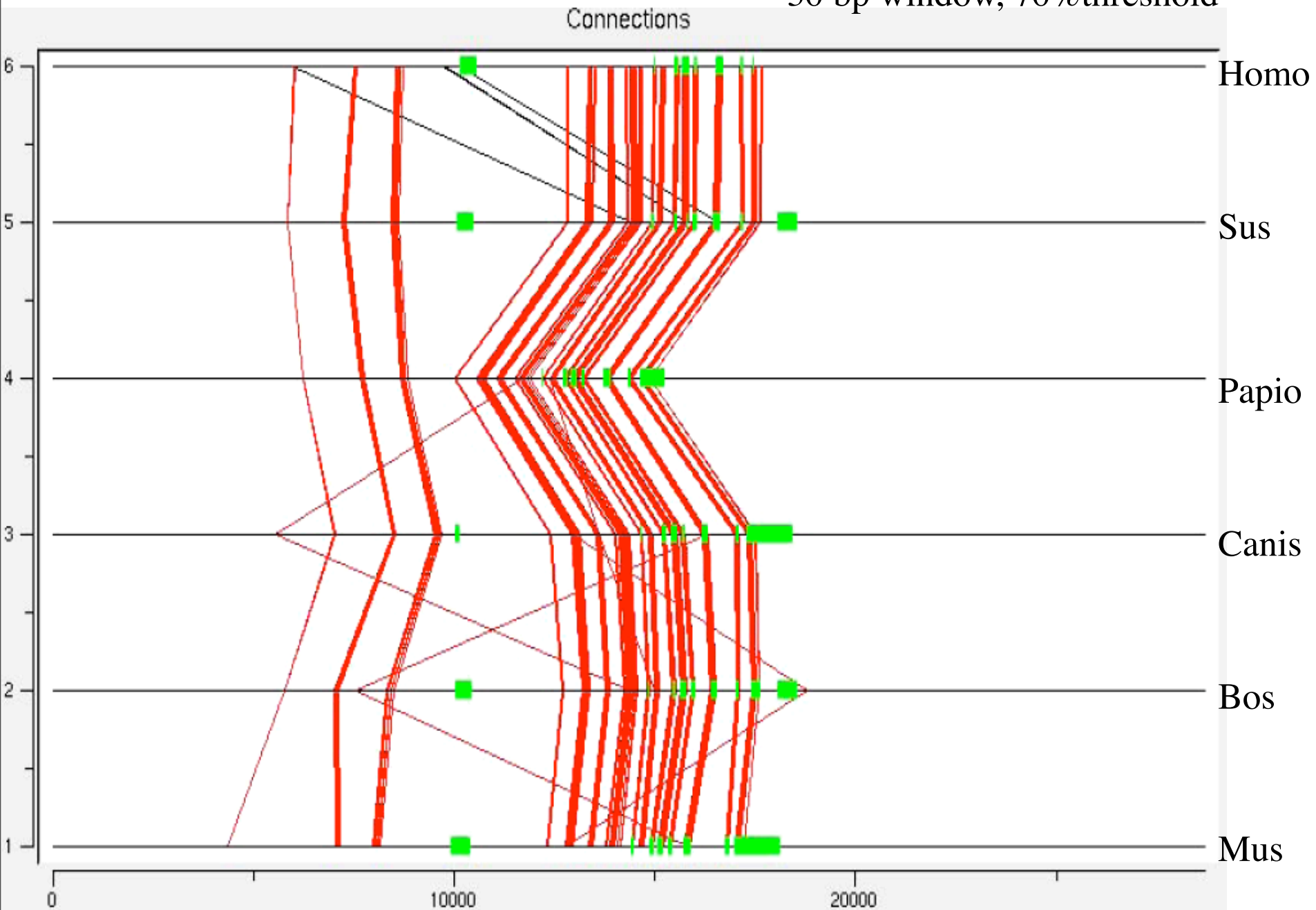


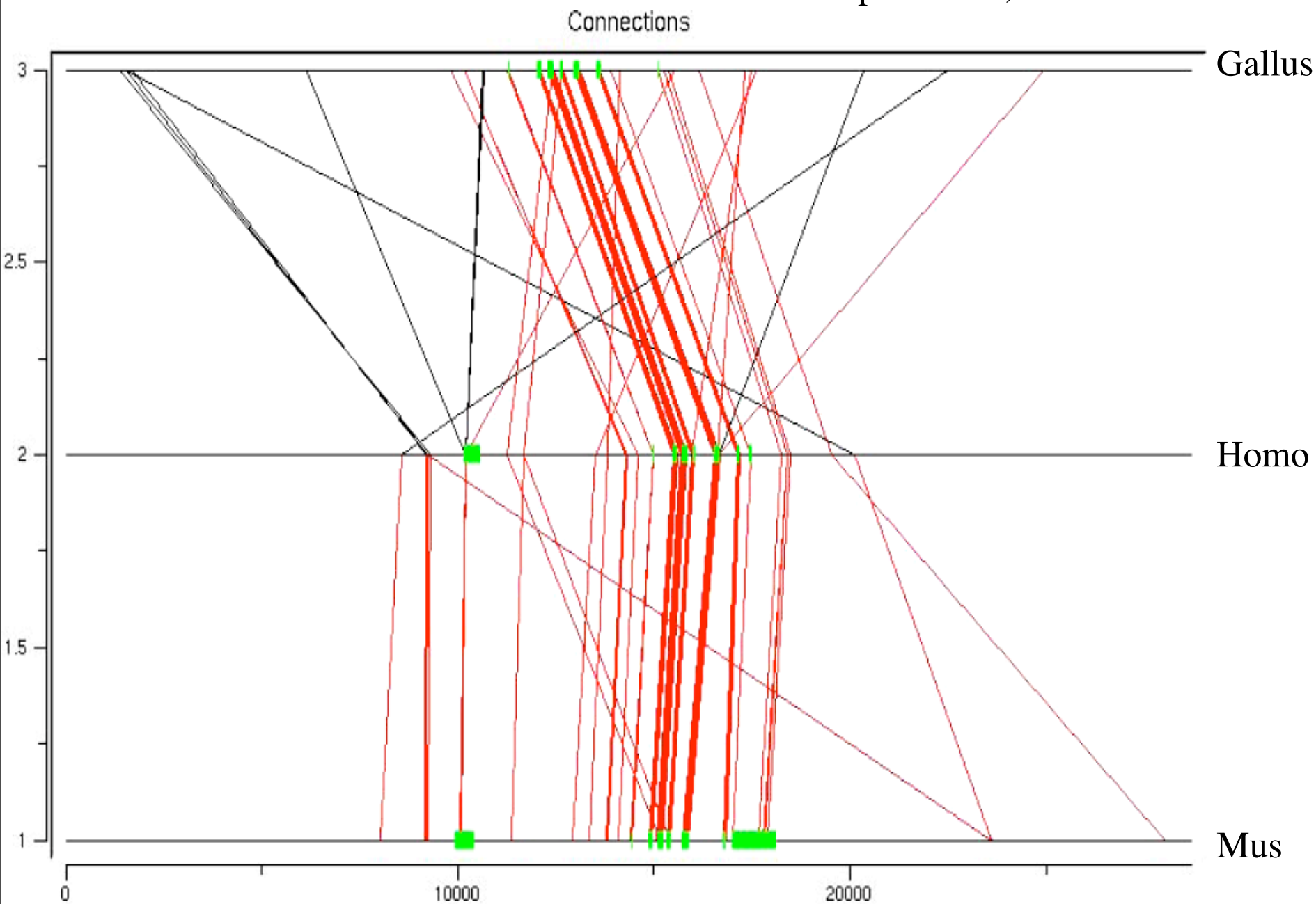
7-way comparison,
30 bp window, 70% threshold



Lunatic Fringe

6-way mammalian comparison,
30 bp window, 70%threshold





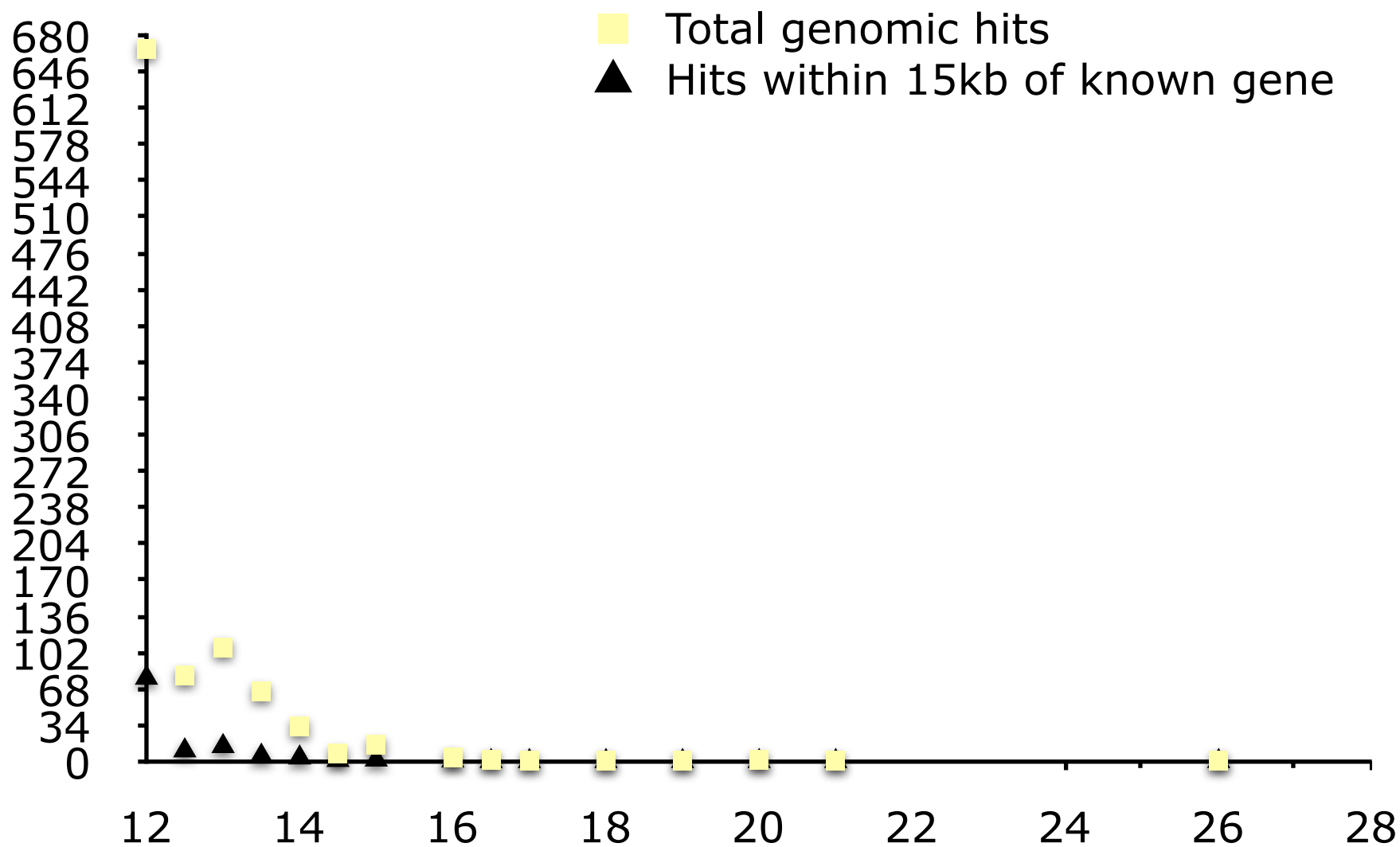
Cistematic Analysis -

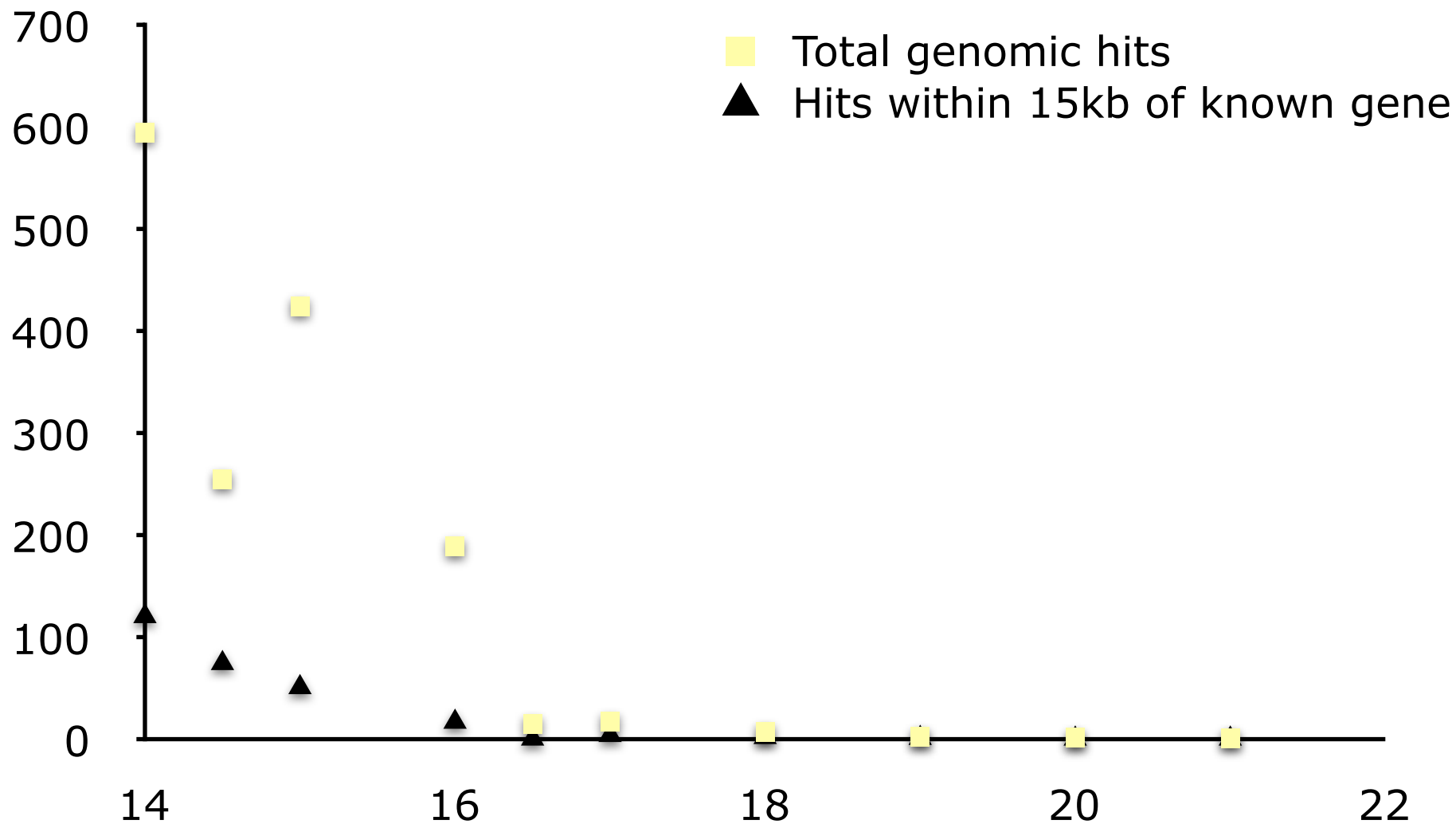
- Different screening techniques
 - Full sequence (12-20 bp)
 - Truncated sequence (13-15 bp)
 - Single mismatch (14-16 bp)
 - Single mismatch on truncated sequence (13-15 bp)
- Frequency of hits
 - Over whole genome
 - Multiple hits per gene
 - Multiple genes per hit
 - Drastic changes depending on technique
 - Dependence on possibility of coding
- Interesting hits
 - Mrf4/Myf5 correlation

Cistematic Analysis - Sample Results for Hes1

Trial 1 for Hes1 (mammals comparison) : Full Sequences

MotifID	Match ID	Match Genom	Match Chrom	Match Locatic	Match Sense	Gene Start	GeneStop	Gene Sense
0-LOC-HES1M	5	mouse	1	121774549	F	121785565	121826795	R
0-LOC-HES1M	6	mouse	1	158634208	R	158619750	158664166	F
0-LOC-HES1M	6	mouse	1	158634208	R	158619750	158664166	F
0-LOC-HES1M	6	mouse	1	158634208	R	158619750	158664166	F
0-LOC-HES1M	6	mouse	1	158634208	R	158619750	158664166	F
0-LOC-HES1M	6	mouse	1	158634208	R	158619750	158664166	F
0-LOC-HES1M	14	mouse	2	142674324	R	142681722	142682160	F
0-LOC-HES1M	15	mouse	2	161333406	R	161345757	161346842	F
0-LOC-HES1M	18	mouse	3	95639800	F	95637165	95639352	R
0-LOC-HES1M	18	mouse	3	95639800	F	95650229	95652079	R
0-LOC-HES1M	2	mouse	5	20549444	F	20546314	20550938	F
0-LOC-HES1M	7	mouse	5	66380016	F	66387784	66469033	R
0-LOC-HES1M	9	mouse	5	103366869	R	103364018	103375463	F
0-LOC-HES1M	9	mouse	5	103366869	R	103364018	103375463	F
0-LOC-HES1M	9	mouse	5	103366869	R	103364018	103375463	F
Genome	Gene Locus	Match Seque	Gene Name	Gene Descrip	Gene Alterna	Gene GO	Gene GO Des	Gene GO Type
mouse	383556	TCTCGCagcataaggacctgt						
mouse	58244	TCTCACAGCA	Stx6	syntaxin 6	2310039E05F	GO:0016020	membrane	C
mouse	58244	TCTCACAGCA	Stx6	syntaxin 6	2310039E05F	GO:0016021	integral to m	C
mouse	58244	TCTCACAGCA	Stx6	syntaxin 6	2310039E05F	GO:0008565	protein trans	F
mouse	58244	TCTCACAGCA	Stx6	syntaxin 6	2310039E05F	GO:0006810	transport	P
mouse	58244	TCTCACAGCA	Stx6	syntaxin 6	2310039E05F	GO:0006886	intracellular p	P
mouse	329527	tgtctcagtttcatcacatgt						
mouse	12547	TATCCCAGAGTGAATACATGT						
mouse	66195	TCTCTCAGCATCAAGACATGT						
mouse	67718	TCTCTCAGCATCAAGACATGT						
mouse	269628	TTTCCCAGCATAAGGACgtgt						
mouse	77569	TCTCCCAGTCTCAGAACATGT						
mouse	13406	TTTCTCAGAA	Dmp1	dentin matrix	PP serine rich	GO:0005578	extracellular	C
mouse	13406	TTTCTCAGAA	Dmp1	dentin matrix	PP serine rich	GO:0005615	extracellular	C
mouse	13406	TTTCTCAGAA	Dmp1	dentin matrix	PP serine rich	GO:0001503	ossification	P
mouse	208118	TCTCTCAGAGTCAAAACATGT						





Cistematic Analysis- Results for Hey1

Computational Conclusions

- 2 additional conserved sites of interest near Mor28
- 25 conserved upstream elements from Hes1
 - At least 5 of particular interest (interesting matches or interesting behavior)
- 20 conserved upstream elements from Hey1
- 6 conserved upstream elements from Lfng
- Need improved screening techniques
 - Cross-genome conservation of hit neighborhoods
 - Distribution of probable sequence necessary to

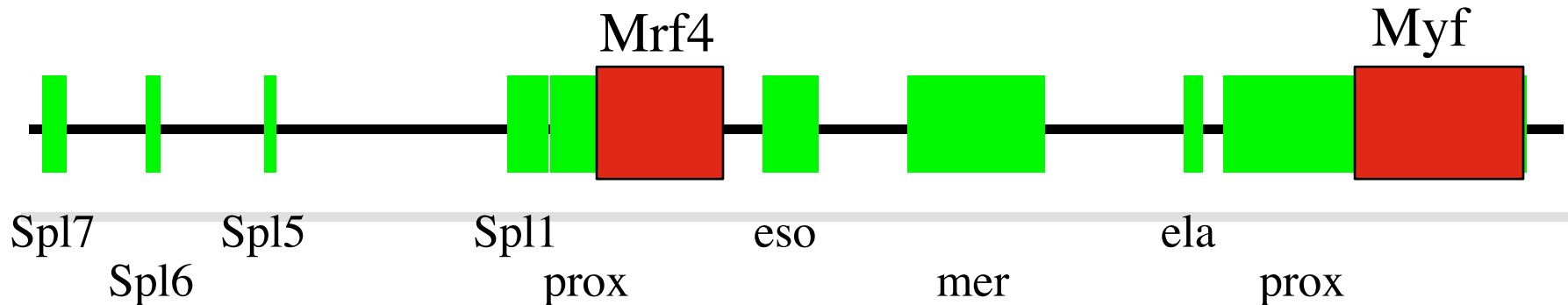
Part II: Wet Lab Experiments

- Mrf4/Myf5 elements
- Hey1 elements
- β -galactosidase vector with β -globin promoter
 - Requires enhancer for significant expression
- Utilization of lentivirus for mammalian infection
- Future experiments
 - Hes1
 - Lfng
 - Hox

Plasmid Constructs

- Standard mouse BAC (from *bacpab.chori.org* and prepared by Hiroaki Shizuya)
- Vector from Libera
- Successfully cloned and inserted into vector
 - Mrf4: Mrf4prox, Spl1, Spl5, Spl6, Spl7
 - Myf5: prox, proxA, proxB, ela, eso
 - Hey1: c9 (2 varieties)
 - The c7 and c67 elements may have faulty primers

Mrf4/Myf Element Distribution



Location on Chromosome 10

	Start	Stop
mrf4_spl7	109995845	109995523
mrf4_spl6	109994545	109994362
mrf4_spl5	109993081	109992921
mrf4_spl1	109990085	109989567
mrf4_prox	109989553	109989246
Mrf4	109989365	109987424

myf5_eso	109986922	109986233
myf5_mcr	109985123	109983442
myf5_ela	109981488	109981714
myf5_prox	109980455	109981217
myf5_proxA	109980794	109981217
myf5_proxB	109980460	109980788
Myf5	109977476	109980702

Lentiviruses

- Lentivirus
 - Modified HIV
 - Single infection
 - Three plasmids for formation
- Grown in 293T cells (human kidney, epithelial cells)
- Infected C2C12 cells (murine muscle (myoblast), fibroblast cells)
- Viral constructs:
 - Mrf4prox, Spl1, Spl5, Myf5ela

Viral Infection Results

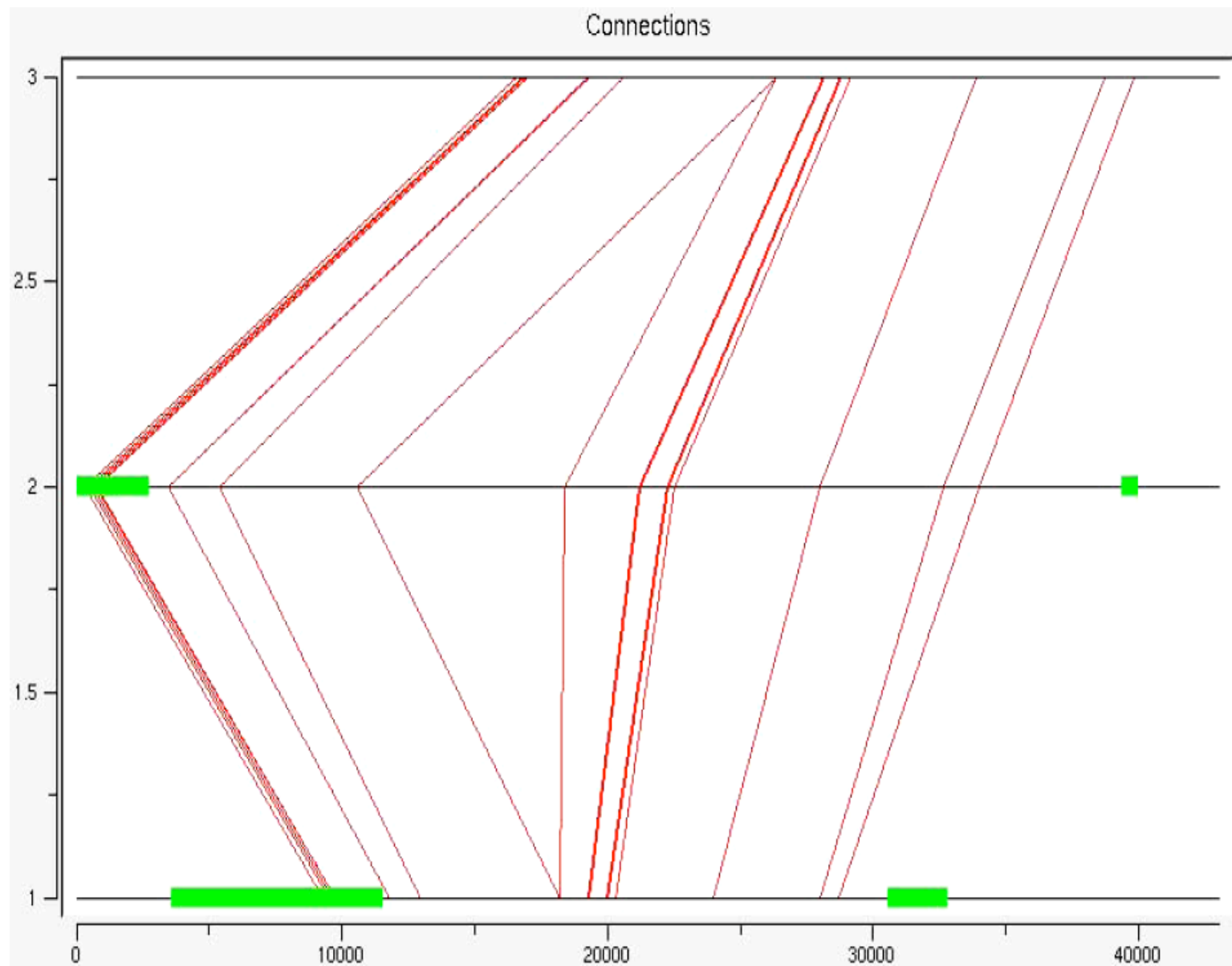
- C2C12 cells appear to have differentiated successfully
- No successful β -galactosidase expression
- Require a positive control to narrow problem to the vector rather than the virus or cells
- Potential problems with expression
 - Poor viral production/isolation
 - Failure of cells to express Mrf4/Myf5 during differentiation
 - Failure of conserved sequence to act as a independent enhancer element sufficient for β -gal expression
- Potential solutions
 - Improve production/isolation technique
 - Screen virus in embryonic infections
 - Attempt combinatorial conserved sites

Future Directions in Comparative Genomics

- Further somitic gene targets: Hes7, Axin2, Lfng, Homeobox family
 - Obtain putative enhancer elements
- Improved Screening Techniques
- Mouse vs. Worm Developmental Regulation
 - Reallocation of developmental pathways in distant species
 - Hes1 (lin-22 in worm) correlation
 - Fails to appear during a Blast search
 - Hey1 not identified in worm

MUSSA and cistematic in worm

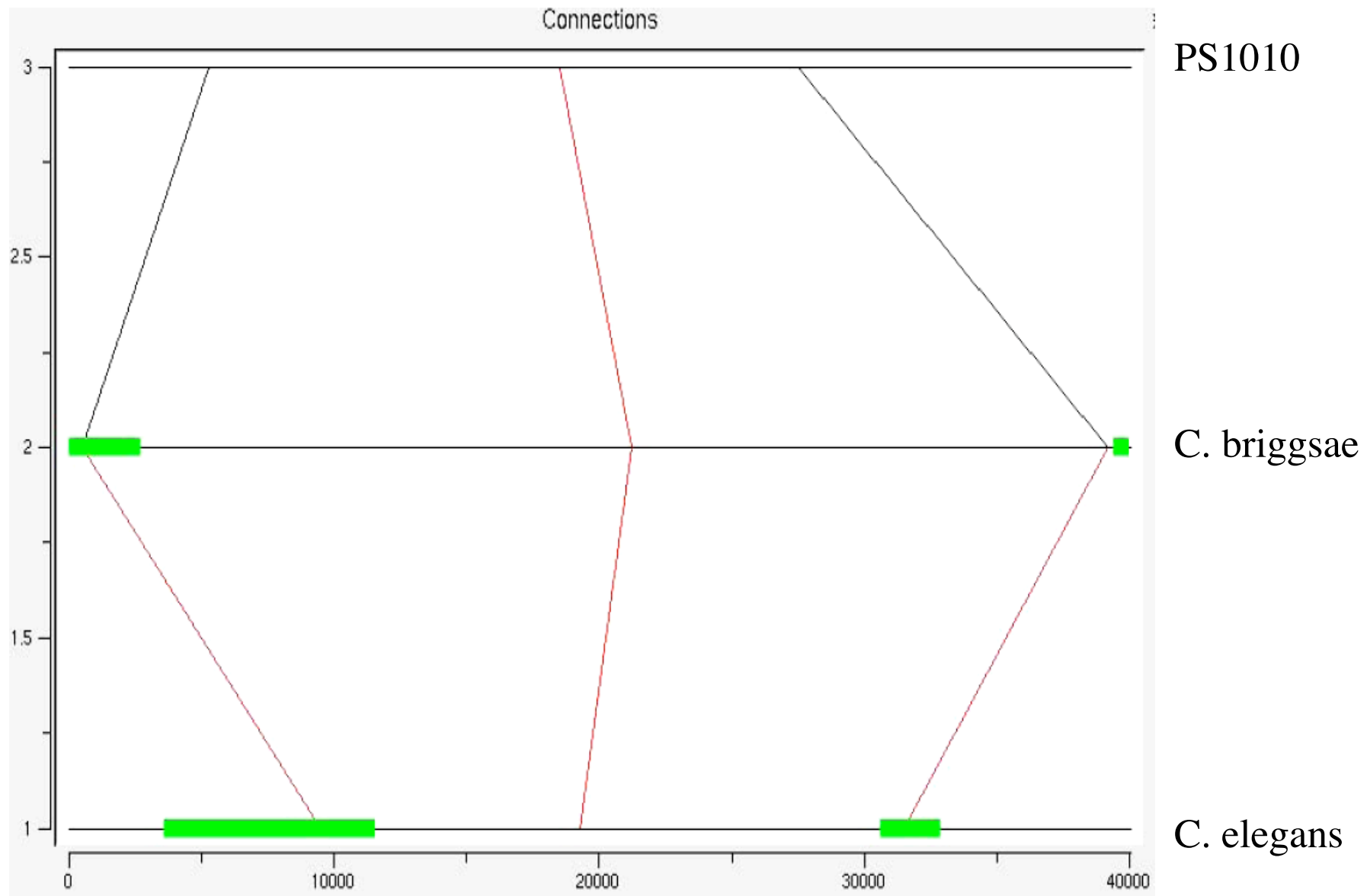
- Mussa analysis image (C. elegans vs. C. briggsae vs. CB5161)
- Relate hox with worm and vertebrate
 - Hox very poorly studied in worm
- Lin-39 : Ceh-13 :: Hoxa5 : Hoxd1
- Hes1 (lin-22 in worm, some research has been done on it, causes some neural development problems) and other somite-development associated genes in worm



CB5161

C. briggsae

C. elegans



Acknowledgments

- The Wold Lab
- Wold Lab associates

