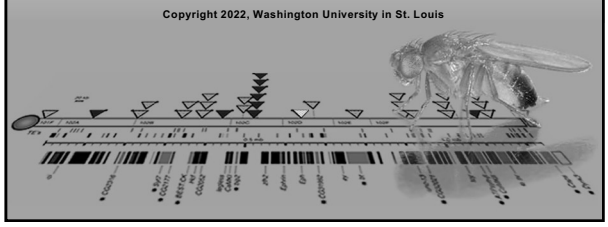
The Strange Case of the Drosophila Dot Chromosome

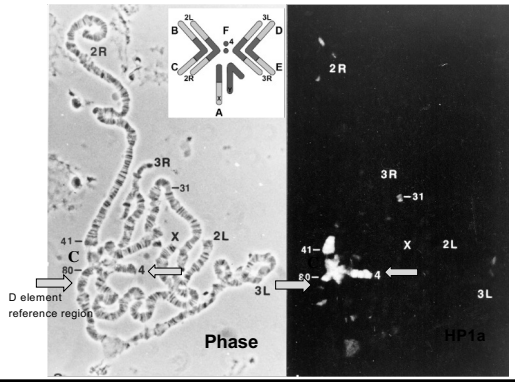
GEP 2022

Slides assembled with notes by
SCR Elgin, WUSTL



1

The *Drosophila melanogaster* fourth chromosome is largely heterochromatic, but the distal 1.35 Mb has a gene density equal to that of euchromatic arms

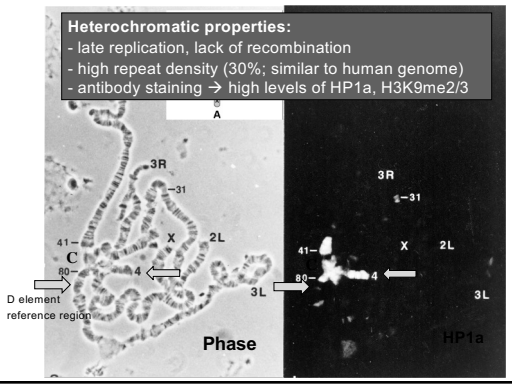


James et al, Mol Cell Bio 1986; Eur J Cell Biol 1989.

2

The *Drosophila melanogaster* fourth chromosome is largely heterochromatic, but the distal 1.35 Mb has a gene density equal to that of euchromatic arms

Heterochromatic properties:
- late replication, lack of recombination
- high repeat density (30%; similar to human genome)
- antibody staining → high levels of HP1a, H3K9me2/3



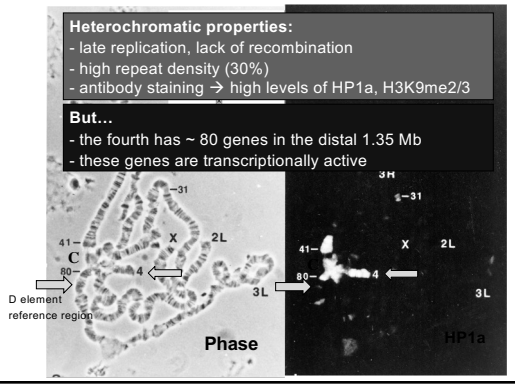
James et al, Mol Cell Bio 1986; Eur J Cell Biol 1989.

3

The *Drosophila melanogaster* fourth chromosome is largely heterochromatic, but the distal 1.35 Mb has a gene density equal to that of euchromatic arms

Heterochromatic properties:
- late replication, lack of recombination
- high repeat density (30%)
- antibody staining → high levels of HP1a, H3K9me2/3

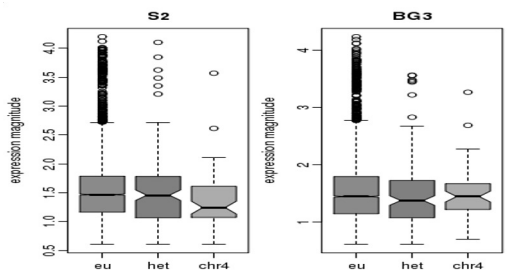
But…
- the fourth has ~ 80 genes in the distal 1.35 Mb
- these genes are transcriptionally active



James et al, Mol Cell Bio 1986; Eur J Cell Biol 1989.

4

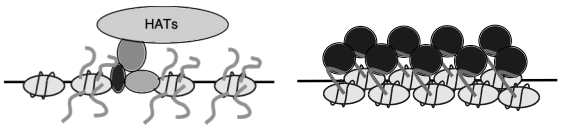
Expression levels of genes in the different
D. melanogaster genomic domains are similar



Adapted from Riddle et al 2012, PLoS Genetics 8: e1002954

5

The F element meets all criteria for heterochromatin

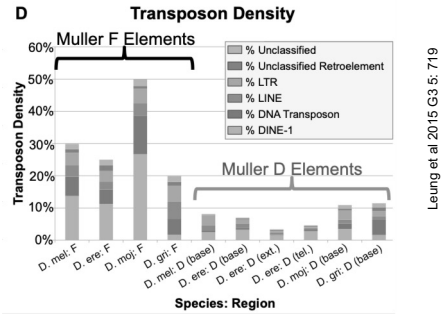


- Euchromatin
  - Less condensed
  - Chromosome arms
  - Unique sequences; gene rich
  - Replicated throughout S
  - Recombination during meiosis
- Heterochromatin
  - Highly condensed
  - Centromeres and telomeres
  - Repetitious sequences; gene poor
  - Replicated in late S
  - No meiotic recombination

Transcriptional activators
Hyper-acetylated histone tail
Heterochromatin Protein 1 complex (HP1)
Hypo-acetylated histone tail; methylated H3/K9

6

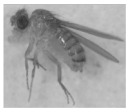
Muller F elements consistently have a higher repeat density than Muller D elements (heterochromatic vs euchromatic domains)

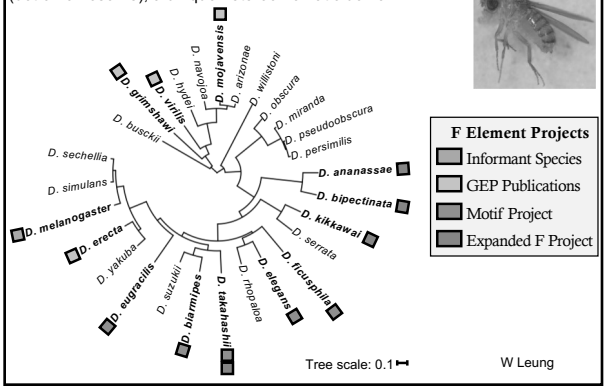


Leung et al 2015 G3 5: 719

7

Our research goal: to use *Drosophila* comparative genomics to understand the organization, evolution and gene function on the F element (dot chromosome), a unique heterochromatic domain

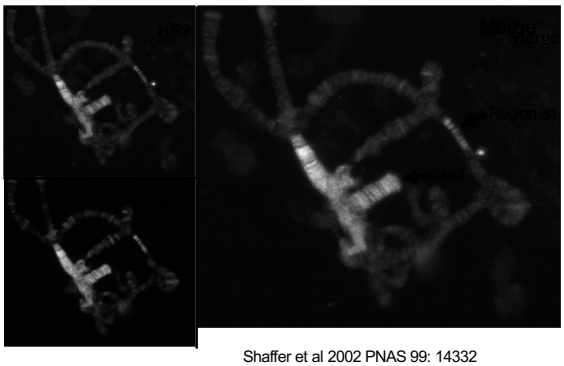




W Leung

8

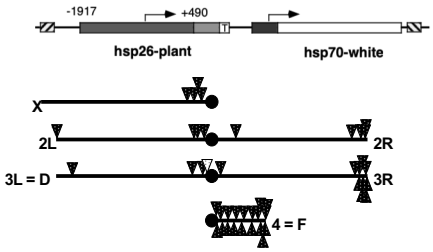
HP1a has a banded pattern on chromosome 4; could the genes lie in HP1-depleted domains?



Shaffer et al 2002 PNAS 99: 14332

9

Insertion sites resulting in a variegating phenotype identify heterochromatin domains



Silenced
1%

Active
99%

Are there **permissive domains** on F that allow full expression?

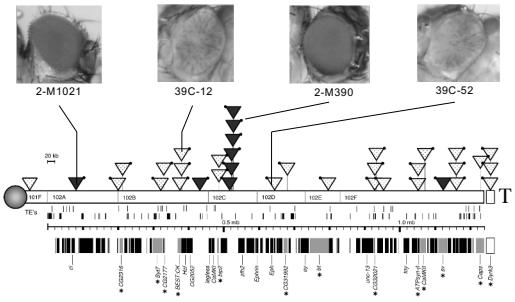
Screen for lines giving variegating or red-eye phenotype with insertion sites on the fourth chromosome.

(First red-eyed fourth chromosome insertion = #707).

Wallrath and Elgin, 1995

10

Most F element insertions variegate, including those inserted into genes. What are the underlying sequence characteristics?

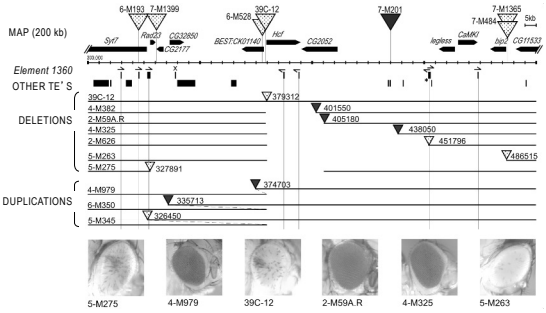


Sun et al 2004 Mol Cell Bio 24: 8210

Each triangle indicates a line carrying a single P element reporter inserted at that site, with the eye phenotype shown, red or variegating, indicating a permissive (euchromatic?) or silencing (heterochromatic) environment. Most insertions into genes are variegating.

11

Local genome deletions and duplications produce a switch in eye phenotype

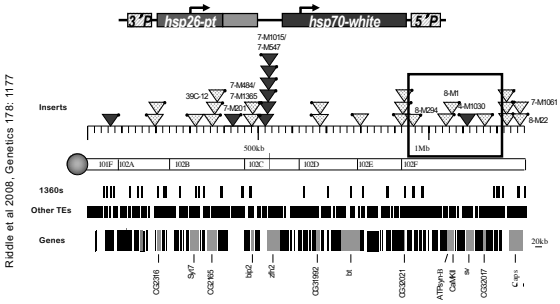


Sun et al 2004 Mol Cell Bio 24: 8210

Proximity (~10 kb) to *1360* (a TE) appears critical.
No evidence of distinct boundaries between silenced / active.

12

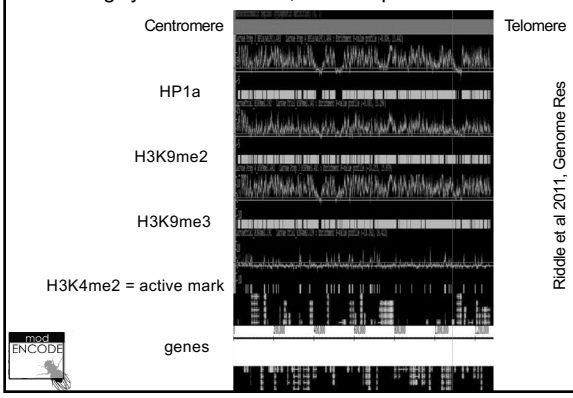
1360 is NOT the only target for heterochromatin formation on the F element, but not all TEs are targets



Conclusion: the bulk of the F element is packaged in heterochromatin, including the genes; a few permissive domains are detected.

13

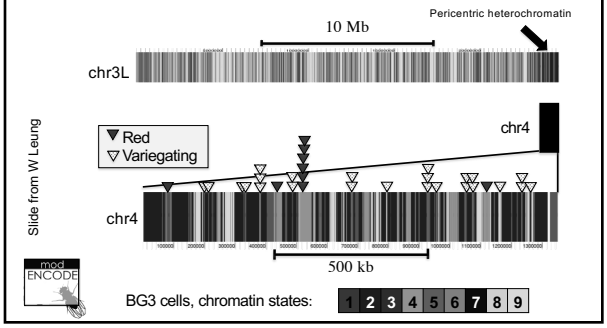
ChIP mapping shows that the fourth chromosome is largely heterochromatic, but has spikes of H3K4me2



14

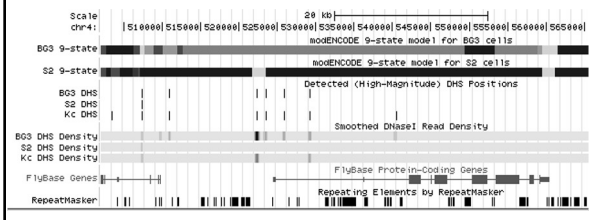
An expanded view of the fourth chromosome: variegating reporters lie in domains that are red (#1, TSS) or blue (#7/#8, heterochromatin).

Red-eye reporters lie in domains that are grey (#6, Polycomb) in at least one cell type; → #6 is a permissive state.



15

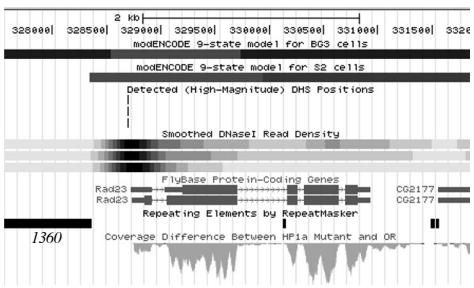
A subset of 4th chromosome genes is associated with Polycomb; these domains are permissive for reporter expression (red eyed fly).



It appears that Pc/H3K27me3 domains allow formation of DH sites, while HP1a/H3K9me2/3 domains do not.

16

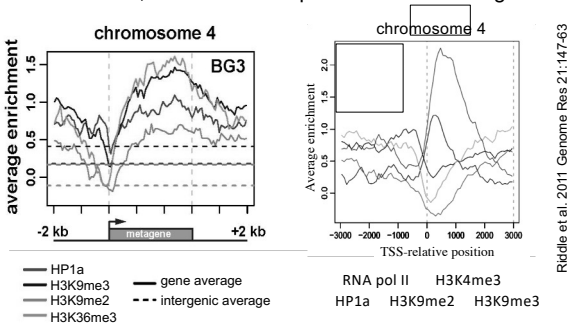
The fourth chromosome has many "red/blue" genes which are consistently expressed, *e.g., Rad23*



How do they manage? What features drives 4th chromosome gene expression that are absent from euchromatic genes (*hsp70-white*)

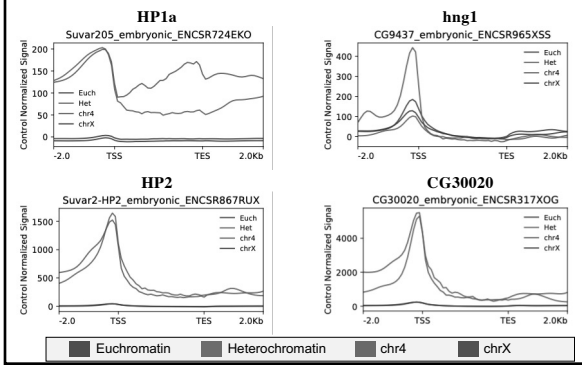
17

Active fourth chromosome genes show depletion of HP1a just after the TSS, but enrichment upstream & across the gene



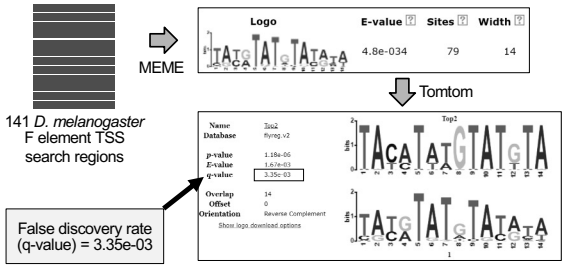
18

Metagene analysis of interesting candidates from the modERN ChIP-Seq data (*D. melanogaster*)



19

Fourth chromosome TSS regions are enriched for $(TATG)_3$, a putative Top2 binding site

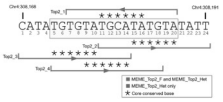


This motif has significant similarity to the reported binding site for Topoisomerase II (Top2).

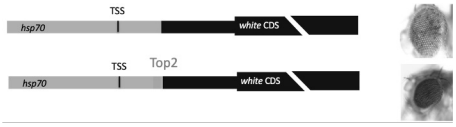
B French and W Leung, unpublished.

20

F element TSS regions are enriched for a putative Top2 binding site (D. melanogaster)



	F element TSS	chr2 and chr3 heterochromatic region TSS	D_base TSS
Number of input sequences	141	212	244
Number of significant motif occurrences (q-value < 0.05)	267	224	13
Number of TSS with at least one significant motif occurrence	70	104	9
Percentage of TSS with a significant motif occurrence	49.6%	49.1%	3.7%



B French, J Cantrell, W Leung, E Gracheva, S Elgin, unpublished observations

21

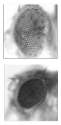
$(TATG)_3$ is found at ~50% of the heterochromatic gene promoters (+/- 300 bp of TSS), but only ~4% of euchromatic gene promoters

	F element TSS	chr2 and chr3 heterochromatic region TSS	D_base TSS
Number of input sequences	141	212	244
Number of significant motif occurrences (q-value < 0.05)	267	224	13
Number of TSS with at least one significant motif occurrence	70	104	9
Percentage of TSS with a significant motif occurrence	**49.6%**	**49.1%**	**3.7%**

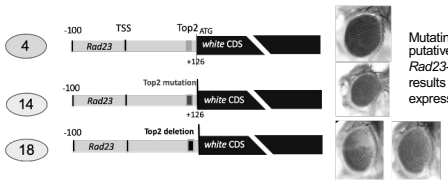
B French and W Leung, unpublished.

22

The $(TATG)_3$ site promotes gene expression on the 4th chromosome



Adding the putative Top2 site to the *hsp70-white* construct results in ca. full expression.



Mutating or deleting the putative Top2 site from *Rad23-white* constructs results in loss of expression.

E Gracheva, J Cantrell, S Elgin et al, unpublished

23

Conclusions

- TEs are essential to the evolution of eukaryotic genomes, but their activation is deleterious to the individual
- TEs are silenced by chromatin packaging; clusters of repeats appear as heterochromatin.
- The F element of *D. melanogaster* appears largely heterochromatic:
  - ~30% repetitious DNA (~80% in *D. ananassae*);
  - but ~80 genes, expressed at similar levels as euchromatic genes.
- Heterochromatin formation is driven by high repeat density:
  - TE *1360* is a target for silencing;
  - Other TEs, but not all, are targets for silencing.
- Gene/chromatin structure of F element genes:
  - most are "heterochromatic (red/blue)" genes, dependent on HP1a for optimal expression
  - active chromatin marks at TSS; see enrichment of het-specific TFs, possible Top2 motif;
  - **Further explorations through GEP "motif" project**

S Elgin

24

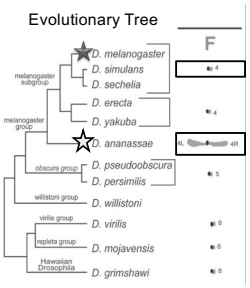
Project overview: F element expansion

- Drosophila F Element Expansion: A Window on the C-value Paradox
  - Sequence improvement and gene annotations of four *Drosophila* species: *D. ananassae*, ***D. bipectinata***, ***D. kikkawai***, and *D. takahashii*
- Scientific observations: Characterizing F element expansion for each species:
  - What factors (*e.g.*, transposons) contribute to F element expansion? What TEs are involved, what is their evolutionary history?
  - What is the impact on gene characteristics? Gene size? Distribution of repeats in relation to gene structure? First exon size? Codon bias? Tm?
  - Does high repeat density promote or allow other changes? Pseudogenes, mitochondrial genes, wanderer genes, fragments, tandem arrays?
- Scientific questions: Determining the impact of expansion on gene / genome evolution by comparing all four species (led by Dr. Chris Ellison).
  - Document rate and timing of TE acquisition;
  - Look at codon bias and substitution rate to contrast impact of centromere and of heterochromatin;
  - Ask whether the change in F element size is associated with a change in effective population size.
- Where are we now?
  - *D. ananassae*: estimates of properties (from 13 genes) available (Leung et al 2017).
  - high quality assembly now available (PacBio plus Nanopore assembly); RNAseq, RAMPAGE data available for all 4 species. Ready to explore!

25

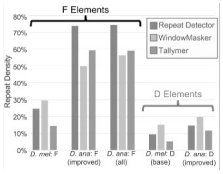
The *D. ananassae* F element (19.74 Mb) has significantly expanded in comparison to *D. melanogaster* (1.35 Mb)

Evolutionary Tree

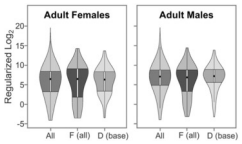


Schaeffer SW, et al. 2008. Genetics 179: 1601–1655

Repeat Densities on the *D. melanogaster* and *D. ananassae* F Elements

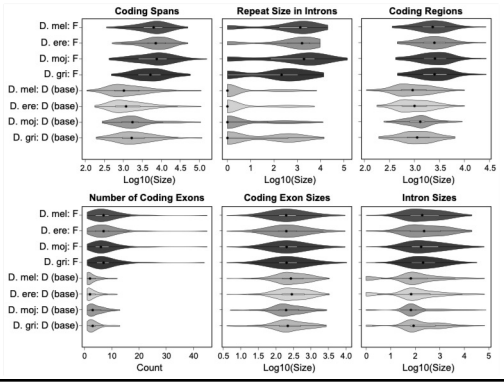


D. ananassae F element Relative Expression Patterns



26

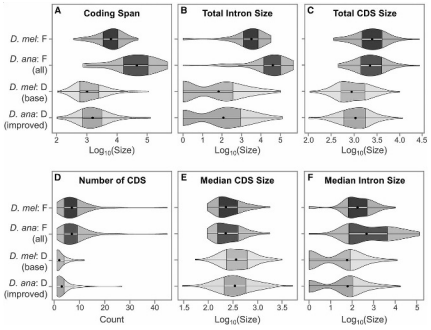
First results from study of four species: F element genes have larger coding spans, more repeats in introns, and more exons than D element genes



Leung et al 2015 G3 5: 719

27

D. ananassae F element genes show similar patterns, but are much larger than *D. melanogaster* F element genes.



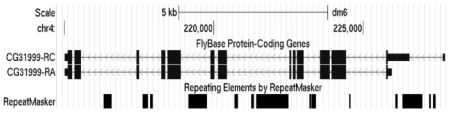
Wilson Leung et al. G3 2017;7:2439-2460

G3 Genes | Genomes | Genetics GSA

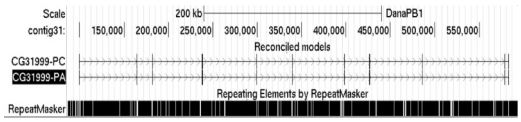
28

Impact of expansion on coding span size

***D. melanogaster* coding span = 10,730 bp**



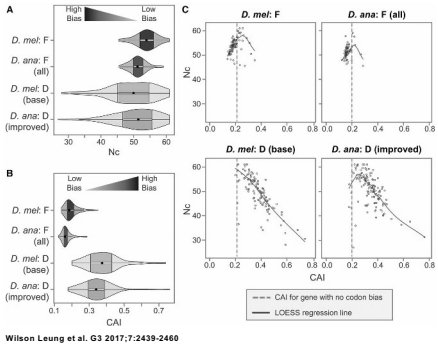
***D. ananassae* coding span = 483,573 bp**



Annabelle Laughlin and Meher Arora, unpublished

29

Codon bias in *D. ananassae* F-element genes can primarily be attributed to mutational biases instead of selection → very little recombination.

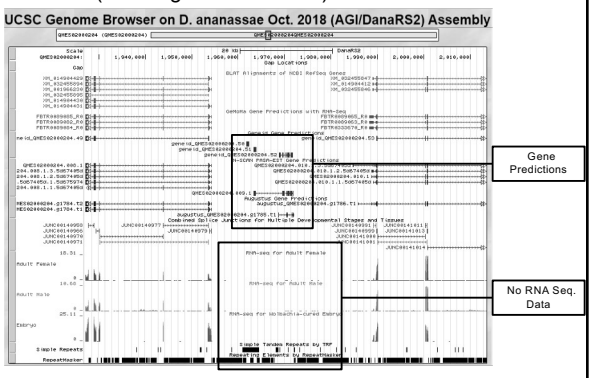


Wilson Leung et al. G3 2017;7:2439-2460

G3 Genes | Genomes | Genetics GSA

30

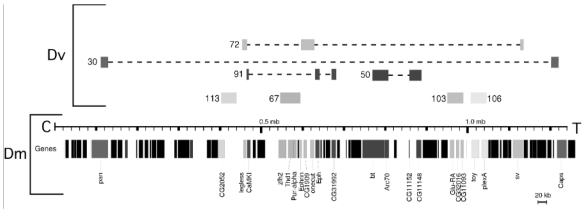
Higher frequency of paralogs, pseudogenes, retrogenes, insertions (including mitochon. DNA) in *D. ananassae* F element



Ethan Cordes and Martin Dalling, unpublished observations

31

Synteny: the order and orientation of the genes:
Initial analysis of the *Drosophila virilis* dot chromosome fosmids



Almost all of the same genes are present,

but rearrangements within the chromosome are common!

Slawson et. al., 2006 Genome Biology, 7(2):R15.

32

Challenges, Questions

What happens to all of these characteristics when the chromosome is larger due to repetitious elements? How similar, and how different, is *D. kikkawai* (2X) to *D. melanogaster* (1X) and *D. ananassae* (14X)?

- Scientific questions: Characterizing F element expansion for each species:
  - What is the impact on gene characteristics? Gene size? Distribution of repeats in relation to gene structure? First exon size? Position, characteristics [motifs] of the TSS? Codon bias?
  - What factors (*e.g.*, transposons) contribute to F element expansion? What TEs are involved, what is their evolutionary history?
  - Does high repeat density promote or allow other changes? Pseudogenes, mitochondrial genes, wanderer genes, fragments, tandem arrays?
- Available evidence
  - High quality genome assemblies, RNAseq & RAMPAGE data;
  - Prior results from *D. melanogaster* investigations;
  - Prior results from GEP student work.
- Current efforts: genes & TSS for *D. bipectinata, D. kikkawai.*
- Good luck, and good hunting!

S Elgin

33