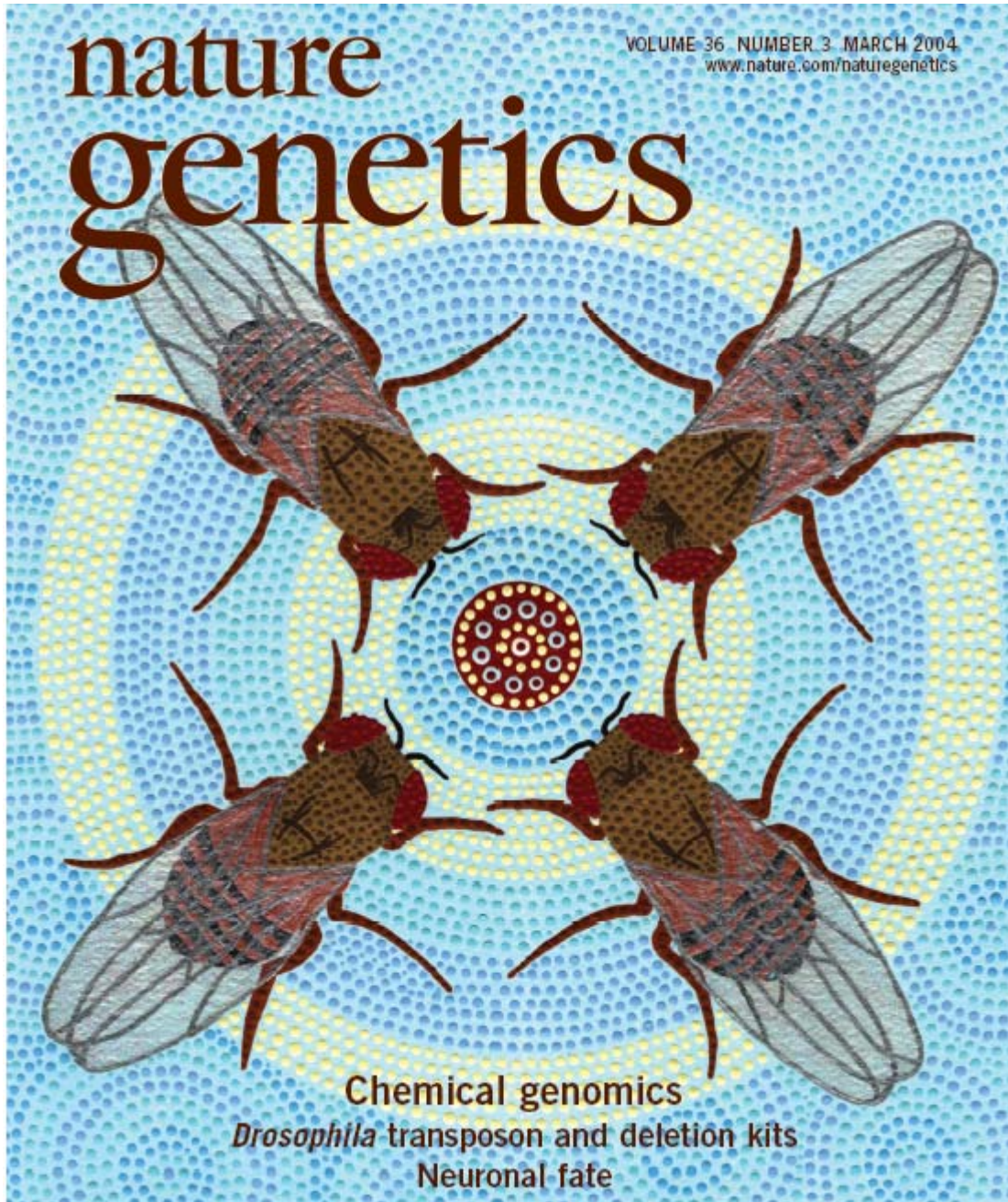


“Release the Flies” 2004



Exelixis

Bloomington Stock Center

Genome Disruption Project

Spyros Artavanis-Tsakonas



Dispersal of the Exelixis *Drosophila* Stock Collection

What fly stocks are being distributed?

- ◆ Isogenic transposon insertion collection (Thibault et al., 2004)
 - ◆ 2,200 inserts chosen as single gene tags for Genome Disruption Project (distributed by Bloomington Stock Center)
 - ◆ 14,300 additional inserts distributed throughout genome to provide an isogenic mapping and screening toolkit (Artavanis-Tsakonas/ Harvard)
 - ◆ Heterochromatic insertion working set (Lawrence Berkeley Labs)
- ◆ Deficiency stocks with molecularly defined endpoints, derived from EXEL insertions
 - ◆ 519 deficiencies (average 140 kb), covering 56% genome (Parks *et al.*, 2004, distributed by Bloomington)
 - ◆ Starting insertions containing FRTs to drive public deficiency kit 'completion' (Kevin Cook, Bloomington) of 7 gene tiling pattern
- ◆ 'Useful' transgenic lines
 - ◆ Tool stocks, eg. P transposase pB{ Δ 2-3}, CyO/Tft
 - ◆ Transgenics created to model interesting disease pathways (wt and mutants)

What else is being distributed?

- ◆ Sequence data associated with Tn insertions and Dfs (FlyBase, Genbank)
- ◆ DNA for transposon, tool and expression vectors (DGRC, Genbank)
- ◆ EST data from random primed libraries to complement public data (LBL, Genbank)



A Multi Functional Transposon Toolkit

How were Transposon Insertions Selected for Retention?

- ◆ Genome Disruption Project insert selection (2,200) at BSC
 - ◆ Selection based on genes previously untagged by P screens
- ◆ New collection at Harvard/MGH (Artavanis-Tsakonas)
 - ◆ Selected by spacing 'tags' throughout isogenic genome
 - ◆ Inserts retain differing functionalities (e.g. Df generation, overexpression capacity, in addition to simple gene disruption)



Exelixis Transposon Collection Profile

- ◆ All vectors mobilised in same isogenic background: ideal for screens and mapping
- ◆ Insulator and splice trap functions in vectors increase disruption efficacy (lethality)
- ◆ UAS sites (XP, WH) allow overexpression & antisense in presence of Gal4
- ◆ *piggyBac* vectors (PB, RB, WH) demonstrate different spectrum than *P*-elements

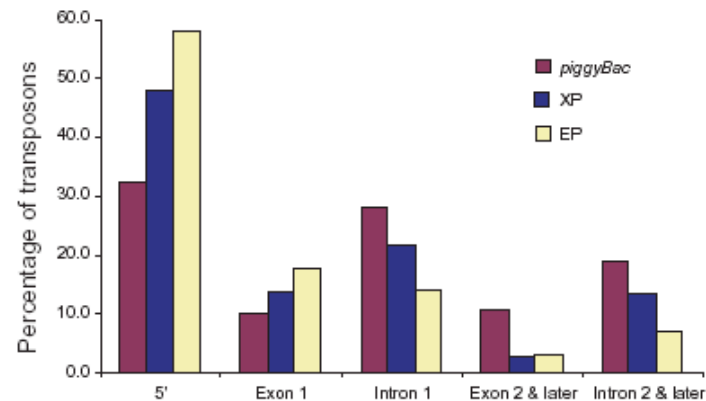


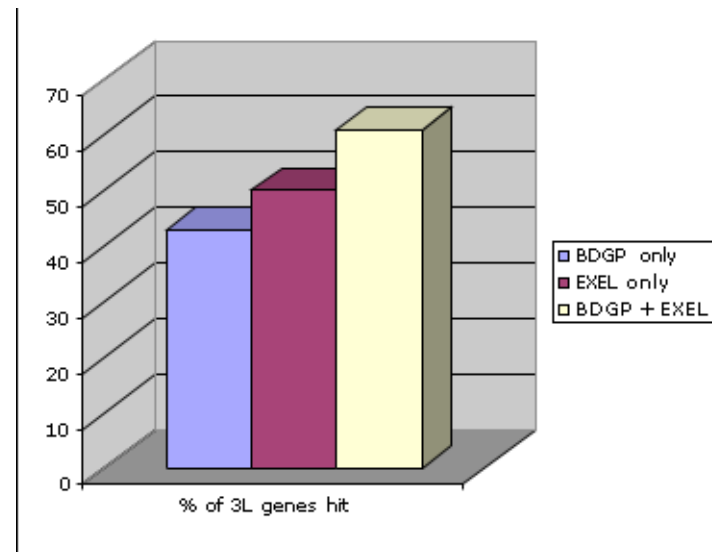
Figure 1 Distribution within genes of 6,196 *piggyBac* elements, 2,462 XP elements and 1,109 EP¹¹ transposons tagging DGC (r1.0) genes. 5' of start = 1,000-bp window upstream of the transcriptional start.

- ◆ *piggyBac* based transposons display clean, precise excision behaviour
PROS: less likely to result in unlinked lethals ('hit and runs')
CONS: can't be used to generate nulls through imprecise excision
- ◆ FRT sites allow creation of designer deficiencies (XP, RB, WH)

Insertions to Supplement Genome Disruption Project

- ◆ Genome Disruption Project (Spradling, Bellen, Hoskins, Hiesinger *et al.*)
- ◆ Compared the Exelixis transposon collection to their P disruption collection on chromosome arm 3L

BDGP alone: 40% of genes hit
EXEL alone: 46%
(EXEL algorithm estimate 53%)
BDGP + EXEL: 58%



- ◆ CGs which remain untagged by their P screens were searched for EXEL insertions
- ◆ One KO line to represent 2,200 new CGs was selected and stocked at Bloomington
 - Lethal insertions were favoured where available
 - No bias toward UAS, FRTs, strand etc was made in this selection

List and Order From Bloomington Center Website

<http://flystocks.bio.indiana.edu/exel-ins.htm>



Exelixis Insertions Chromosome Arm 2L



Updated 10 March 2004

All insertion site data are provided by the donor of the stock or taken from FlyBase updates. Please cite those who generated and analyzed these materials when publishing your own work with these and other Stock Center stocks.

	Stock #	Insertion Symbol	Phenotype	Site
1	10975	PBac{PB}CG2657[c02720]	viable	21A2
2	11335	PBac{PB}CG31973[c04017]	viable	21A3
3	18750	PBac{WH}net[f04249]	lethal (unverified)	21A5
4	19292	P{XP}CG3709[d08265]	lethal (unverified)	21B1
5	19214	P{XP}CG11490[d05023]	viable	21B4
6	18532	PBac{WH}CG3645[f02260]	lethal (unverified)	21B7
7	10033	PBac{PB}Plc21C[c00245]	viable	21C1
8	10720	PBac{PB}ds[c01777]	viable	21D1
9	18098	PBac{RB}CG11838[e02928]	viable	21D1
10	17883	PBac{RB}CG13689[e00773]	viable	21D1
11	18890	PBac{WH}CG13688[f05607]	lethal (unverified)	21D1
12	18678	PBac{WH}Ets21C[f03639]	viable	21D1
13	18988	PBac{WH}jush[f06555]	lethal (unverified)	21D1
14	18106	PBac{RB}Eaat2[e03003]	lethal (unverified)	21D2
15	18742	PBac{WH}CG13946[f04199]	viable	21D2

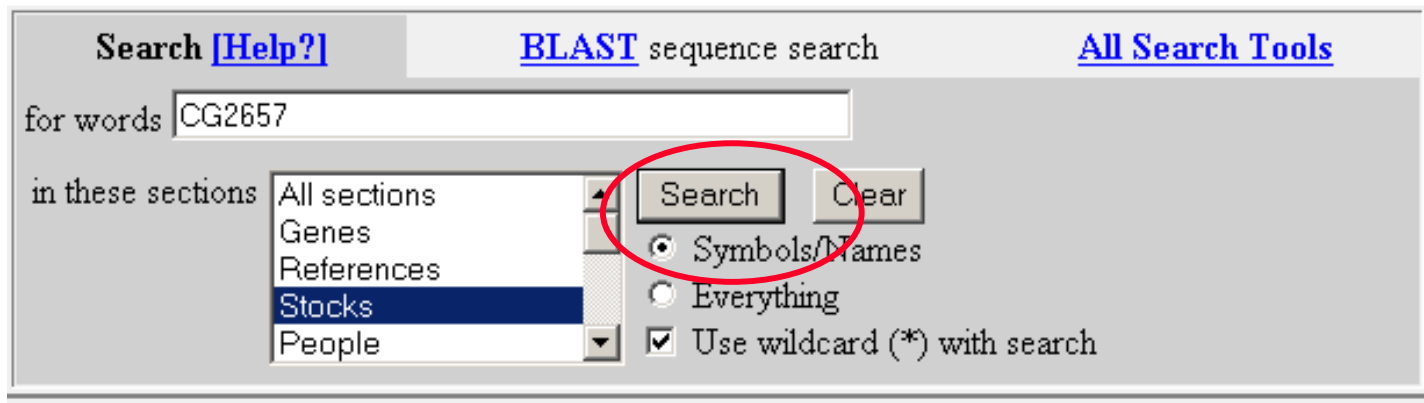
- ◆ Before obtaining these stocks you will be asked:
- ◆ To confirm that you represent a non-commercial institution (or are licensed by Exelixis)
- ◆ To agree to certain usage conditions (agreed by Fly Board, BSC and Exelixis)



Searching for An Insertion in Your Gene Via Flybase

<http://flybase.bio.indiana.edu/>

On FlyBase homepage, type in the gene of interest to query for stocks



Search [\[Help?\]](#) [BLAST](#) sequence search [All Search Tools](#)

for words

in these sections

- All sections
- Genes
- References
- Stocks**
- People

Symbols/Names
 Everything
 Use wildcard (*) with search

Your query returns the genotype, insert number etc with link to Bloomington

[FlyBase Next Generation](#)
[FlyBase](#) .. [Aberrations](#) .. [Anatomy](#) .. [BLAST](#) .. [Genes](#) .. [Annotation/Sequences](#) .. [Gene Products](#) .. [Maps](#) .. [People](#) .. [References](#) .. [Stocks](#) .. [Transgenes/Transposons](#) .. [cDNAs/ESTs](#) |
[Help](#) .. [Searches](#) .. [News](#) .. [Site](#)

FlyBase Stock Report (Bloomington Stock Center)

>>>
Stock Number: 10975
Old P Stock#: P975
Genotype: w[1118]; PBac(w[+mC]=PB)CG2657[c02720]
Chromosome(s): 1;2
Breakpts/Insertion: 21A2
Date added: 3/10/04
Donor: Exelixis, Inc.
Comments: May be segregating CyO, K.C.
Center: Bloomington Drosophila Stock Center Contact: flystocks@bio.indiana.edu

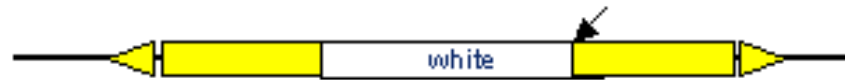
[Send comments](#) to us at flybase-help@AT.morgan.harvard.edu
FlyBase-NG uses [Argos: A Replicable Genome infOrMation System](#)



The EXEL Collection: More than Gene Disruptions

4 insertion vectors engineered with different functionalities
21,500 inserts generated in isogenic background, placed on genome
68% inserts are *piggyBac* based, 32% P element based

Simple *piggyBac* based insertion vector
Single splice trap (-orientation)
No UAS, or FRTs



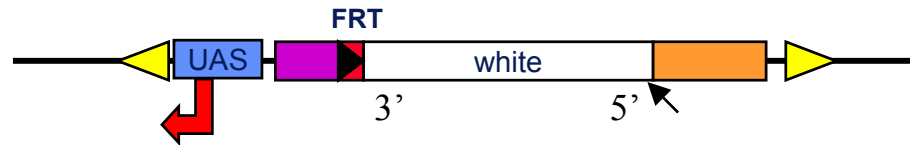
PB
4800 lines

piggyBac based insertion vector
Splice trap disrupts in both orientations
FRTs for deficiency generation



RB
3700 lines

piggyBac based insertion vector
Single splice trap (-orientation)
FRTs for deficiency generation
One Gal4/ UAS overexpression site



WH
6300 lines

P- based insertion vector
Single splice trap (-orientation)
FRTs for deficiency generation, UAS removal
Dual Gal4/ UAS overexpression sites



XP
6800 lines

Trimming the EXEL Collection for Harvard

Rules of Culling

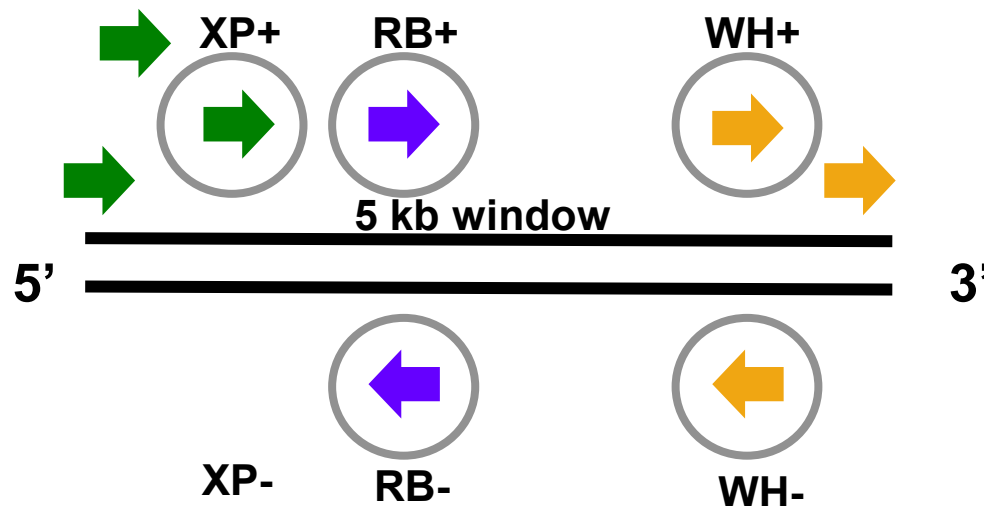
- 1) Generate a sliding window of 5 kb
- 2) Allow 8 max transposons per window

Choose one of each type of transposon on each strand, giving precedence to:

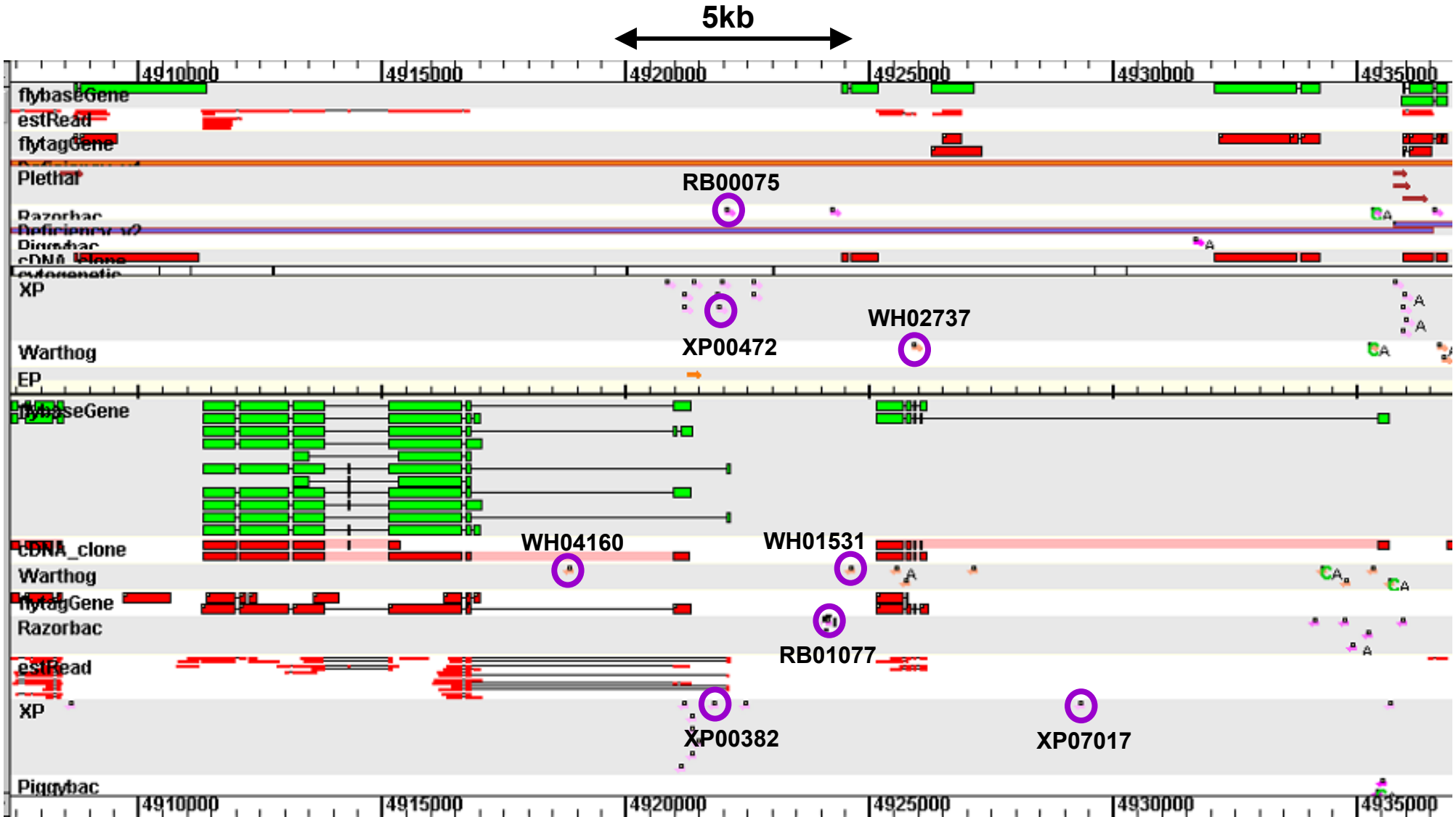
- Lethals
- Level of confidence in the placement (end sequence reads number and quality)
- Variety of transposon and strand (ie. Selection for Df generation, overexpression etc)

3) To the final list, added back

- All lines going to Bloomington
- All transposons which have been molecularly confirmed as clean single gene disruptions (confirmed from screens at EXEL)
- All transposons on the 4th chromosome



Culling Algorithm Example



Selected within 10kb interval:
 3 of 17 XPs
 2 of 3 RBs
 3 of 6 WHs



Culled Collection Selection for Harvard Stock Center

Insertion vector representation after culling 21,500 stocks to 16,500

<i>XP</i>		6800	3807
		32%	23%
<i>RB</i>		3700	3364
		17%	20%
<i>WH</i>		6300	5691
		29%	36%
<i>PB</i>		4800	3635
		22%	22%

Culled collection is 23% *P* vector-based, 77% *piggyBac* vector-based
 79% of insertions can be used for FLP mediated deletion generation
 59% of insertions carry a UAS overexpression element



Information and Distribution Information from Harvard

- Primary source of information : [http:// drosophila.med.harvard.edu](http://drosophila.med.harvard.edu)
 - ◆ Projected web launch date - April 15, 2004
 - ◆ Other sources will be FlyBase and Nature Genetics
 - ◆ Due to the tremendous number of requests already received and to streamline the operation, orders will only be accepted through the Harvard online system
 - ◆ If you have already sent in a request you will have to resubmit it via this system
- Once the website is running and the stocks are established at Harvard they will be available for order
 - ◆ Tentative distribution date is October 1 2004
 - ◆ When ready an announcement will be made via the Harvard website, FlyBase and Nature Genetics



Finding Your Stocks on the Harvard Website

- A FAQ and news section will be posted
- Three stock lists will be available
 - ◆ The complete Exelixis stock list, the Harvard subset and the Bloomington subset
 - ◆ The Exelixis list is relatively user-friendly for planning experiments, though stocks won't be available immediately
 - ◆ A file containing Tn flanking sequences will be posted
- Initially you will only be able to search by stock number:
 - ◆ Simplifying website will allow Harvard to start accepting and filling requests earlier
- Eventually, you will be able to search for stocks using several additional criteria
 - ◆ Text
 - ◆ CG number with range (in bp)
 - ◆ Chromosomal location (coordinate) with range (in bp)



How Can I Place My Order and When Will it Arrive?

- You will use an online ordering system modeled after an e-commerce site to search and order stocks
 - ◆ You must access an account which will require a password
 - ◆ Accounts will be created for labs in the name of the principal investigator (not for individuals)
 - ◆ A FedEx shipping account number must be provided
- Accurate prediction of lag time will require more extensive experience with the stocks and robots
- Orders will be placed in a queue
- Certain stocks will be more heavily requested and will obviously take longer to ship



A Molecularly Targetted Deficiency Kit

Molecularly Targetted Deficiencies

Isogenic background

Molecularly defined endpoint (co-ordinate)

Gene level resolution (~150kb+)

Capacity to flank haploinsufficients

vs

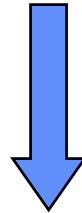
Old Public Deficiency Kit

Complex and variable backgrounds

Imprecise endpoints (cytology)

Cytologic resolution (Mb)

Haploinsufficient regions “missing”



Targetted FRT mediated deficiencies

EXEL Deficiency Kit

519 confirmed, available Dfs at BSC

Delete 56% of CG isoforms

Delete average of 25 CG isoforms

Average deletion size is 140kb

DrosDel Core Deficiency Kit*

176 deficiencies made, 100 confirmed

75 available stocks at Szeged

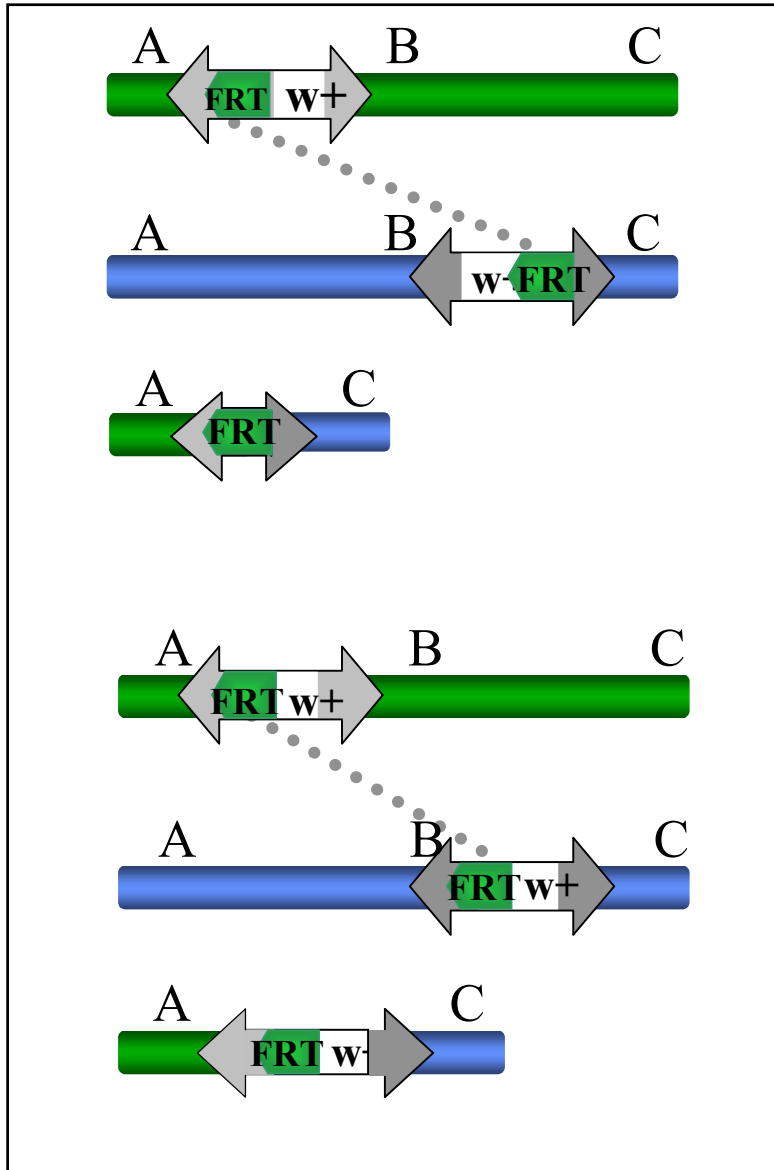
Delete ~50% of genome

Average deletion size is 467kb

*DrosDel collaborators made 94 additional Dfs, 10 stocked at Szeged



FRT Mediated Deficiency Construction



Deficiencies may be w^- or w^+

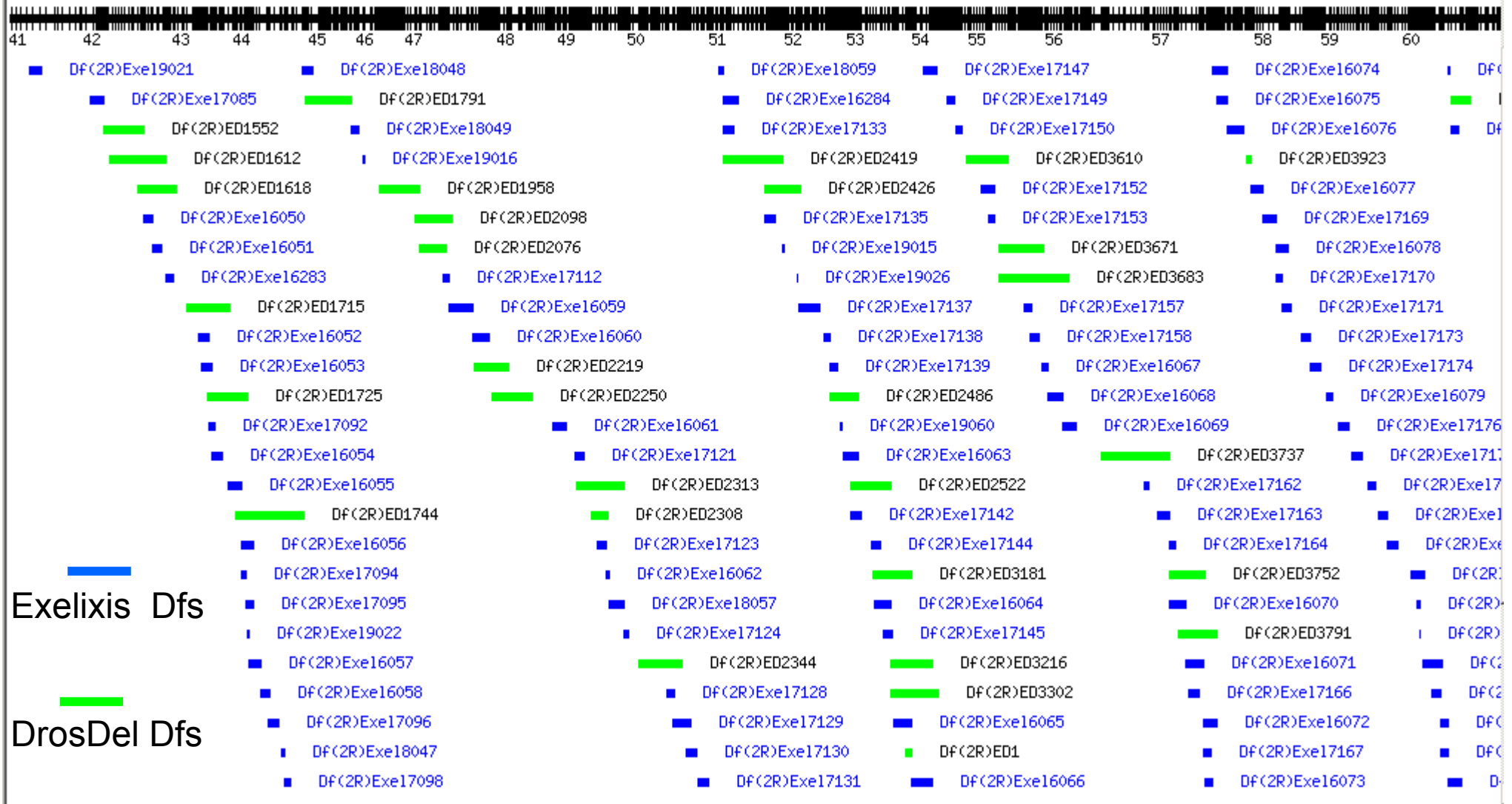
Exelixis methods are described in Parks *et al.*, 2004 (Nature Genetics)

Stocks posted on the BSC website
<http://flystocks.bio.indiana.edu/making-dels.htm>

Drosdel protocols are posted at
http://www.drosdel.org.uk/create_del.php

Coverage of 2R by Molecularly Defined Dfs: March 04

Chromosome 2R Deletions



Coverage maps courtesy of Ed Ryder see poster #907A
<http://www.drosdel.org.uk/coverage.php>



Coverage to Date and Coming Soon...

Exelixis deposited 519 deficiencies (average~140kb) 56% coverage, at Bloomington

Genome coverage 2004 by all molecularly defined deficiencies (EXEL+DrosDel)

<u>Arm</u>	<u>Coverage*</u>
2L	93.6%
2R	83.5%
3L	90.2%
3R	80.6%
X	50.2%
4	52.9%
Total	79.5%

* Coverage data was provided courtesy of Ed Ryder at DrosDel

Additional 1850 deficiencies planned by Bloomington using 3700 EXEL inserts


- ◆ Will 'complete' genome coverage at 15 gene deficiency tiling pattern
- ◆ Finish flanking haplo-insufficient regions
- ◆ Overlay new deficiencies onto existing, to attain 7 gene resolution coverage




Where Can I Obtain EXEL Deficiencies?

Exelixis deficiencies listed in Parks *et al.*, are available from Bloomington website

<http://fly.bio.indiana.edu/browse.htm>



Exelixis Deficiencies




Updated 10 March 2004


- Roughly half of the deletions described in Parks *et al.*, 2004, are listed below. The remaining deletion lines should be ready for distribution from the Bloomington Stock Center by April, 2004. The entire Exelixis deletion collection provides 56% coverage of the euchromatic genome. Unless otherwise noted in the stock description, Exelixis deletions were generated and are maintained in an isogenic background.
- Ed Ryder's (DrosDel Project) maps of deletion coverage may be viewed [here](#).
- Information on generating similar deletions is given at [Generating Deletions from Exelixis Insertions](#).
- Some [restrictions](#) apply to the distribution and use of deletion lines from Exelixis. Commercial researchers must obtain a use license from Exelixis to work with these stocks.

	Stock #	Deficiency	Breakpoints
1	7699	Df(1)Exel6221	1B4;1B8
2	7700	Df(1)Exel6223	1C4;1D2
3	7701	Df(1)Exel6224	1D2;1D4
4	7702	Df(1)Exel6225	1D4;1E3
5	7703	Df(1)Exel6226	1E3;1F2
6	7704	Df(1)Exel6227	1F2;2B1
7	7706	Df(1)Exel6231	3A2;3A3
8	7705	Df(1)Exel6230	3A2;3A4

Listed as Exelixis deficiencies, ordered by cytology



Chromosome 3 Deficiencies



Updated 10 March 2004

In a few cases other stocks with the same deficiency are available and can be identified by [searching](#) for the deficiency symbol.

	Stock #	Deficiency	Breakpoints
1	1452	R(3)S1	61At + 100Ft
2	2577	Df(3L)emc-E12	61A;61D3
3	5838	Df(3L)B71	61A1;61B
4	7562	Df(3L)Exel6083	61A6;61B2 ←
5	5837	Df(3L)7C	61B;61C
6	5836	Df(3L)2D	61B;>61C1
7	7563	Df(3L)Exel6084	61B2;61C1 ←
8	1478	Df(3L)Ar12-1	61C;61F
9	1479	Df(3L)Ar11	61C3-4;61E
10	7564	Df(3L)Exel6085	61C3;61C9 ←
11	439	Df(3L)Ar14-8	61C5-8;62A8
12	7565	Df(3L)Exel6086	61C9;61E1 ←

Listed among all deficiencies, ordered by cytology



Molecular Resolution Mapping Made Quick and Easy

1. Define your mutant region at 'low resolution' between cytological breakpoints or markers

2. Define molecular breakpoints on genomic scaffold with 'high resolution' deficiencies

Test all molecularly defined deficiencies (DrosDel and Exelixis) within region for phenotype

➤ If unlucky, complete your coverage of the region using EXEL or DrosDel FRT bearing insertions to fill in your gaps with a custom set (takes only 4 generations, highly efficient)

➤ When defined by molecularly boundaries (eg. Exel Df) you have resolved to an average 140kb, approx 24 genes. However if you are unlucky and the region is large, refine with a custom set using available FRT bearing transposon inserts in region

3. Test candidates between genomic co-ordinates

➤ Test alleles of CGs that fall within breakpoints: (GDP estimate, 58% of genes have been tags)

➤ If unlucky, cover gaps/ create nulls to create custom micro-deficiencies at gene resolution level using the FRT insertion set

SEQUENCE your candidate ORF!



Gathering Reagents for Molecular Resolution Mapping

- ◆ All EXEL Dfs and the 2,200 insertions selected by GDP are already here
- ◆ All deficiencies and transposon insertions stocked at BSC and Szeged stock centers can be retrieved by cytological search in FlyBase

[FlyBase](#) .. [Aberrations](#) .. [Anatomy](#) .. [BLAST](#) .. [Genes](#) .. [Annotation/Sequences](#) .. [Gene Products](#) .. [Maps](#) .. [People](#) .. [References](#) .. [Stocks](#) .. [Transgenes/Transposons](#) .. [cDNAs/ESTs](#) .. [Help](#) .. [Searches](#) .. [News](#) .. [Site](#)

FlyBase Next Generation

FlyBase Stocks query form

Select field name and enter values:

Stock number matches

and

Any field matches

A wild-card will automatically be added to the end of your search string unless you uncheck this box.

[\[Stocks Search Help\]](#) [\[Stocks Documentation\]](#)

Collection to search <input type="text" value="All"/>	CytoLocation (breakpoints and insertion sites) Min. <input type="text" value="43A"/> Max. <input type="text" value="43D"/>
---	---

Report up to records per page.

All DrosDel confirmed deficiencies, including those not yet available from Szeged at:
http://www.drosdel.org.uk/show_dels.php

Additional Exelixis stocks to generate custom deletions will be available from Harvard by October. Meanwhile you can search for insertions in Thibault *et al.*, Suppl. Table 3
<http://www.nature.com/ng/journal/v36/n3/extref/ng1314-S4.doc>

Genome Disruption Project stocks may be found by searching their database:
<http://flypush.imgen.bcm.tmc.edu/pscreen/>



Summary of Exelixis Stocks Being Distributed

- ◆ Transposon insertion lines
 - ◆ Bloomington has 2,200 insertions selected by GDP stocked
 - ◆ Harvard will distribute 16,500 insertions
- ◆ Bloomington also has general tool stocks for working with the transposon set
 - ◆ Marked, efficient *P* and *piggyBac* transposase mobilization sources on both autosomes for excisions
 - ◆ Isogenic ammunition lines for all EXEL vectors to generate new hops
- ◆ Deficiency sets
 - ◆ All 519 sent to Bloomington
 - ◆ Balancers and tool stocks in isogenic backgrounds for generating custom deletions
- ◆ A number of 'generally useful tools'
 - ◆ Bloomington will also take about 300 transgenics lines from various EXEL disease programs: Alzheimers, Metabolism, Vascular Biology, Oncology and Inflammation

The Fly Community Thanks a Cast of Many!!

Exelixis Fly People

Ferhad Ahmed
Marcia Belvin
Ross Bucholz
Dan Curtis
Maddy Demsky
Thierry Diagana
Nick Dompe
Kyle Edgar
Cathy Erickson
Bill Fisher
Helen Francis-Lang
Heather Gothot
Deanna Grant
Kim Greer
Stephanie Hartouni
Elizabeth Howie
Kari Huppert
Lisa Kadyk
Felix Karim
Linda Keyes
Tom Kidd
Keith Killpack
Niels Klitgord
Casey Kopczynski
Lynne L'Archeveque
Melissa Luning
Matt Mahoney

Marie Marcinko
Jon Margolis
Juile Mazzotta
Wes Miyazaki
Amanda Norton
Amanda Oudin
Annette Parks
Siobhan Roche
Ken Shaw
Matt Singer
Carol Singh
Ron Smith
Lynn Stevens
Christi Stuber
Mariano Tabios
Lory Tan
Doua Thao
Steve Thibault
Stanley Tiong
Sheryl Thomas
Richard Ventura
Kellie Whittaker
Meg Winberg
Chris Winter
Alesa Woo
Irena Zakrajsek
Lora Zhao

Exelixis Bioinformatics

Bob Fawcett
Brian Hawthorne
Dan Joo
Brett Milash
Wes Miyazaki
Stephanie Robertson

Exelixis Sequencing

Feng Chen
LaiMan Cheung
Angela Chong
Lakshme Jakkula
Alex Laufer
Candace Swimmer
Valentina Vysotskaia

Exelixis Other

Andy Buchman, Legal
Dan Elkes Operations

George Scangos CEO

KSE Smartwood

Doug Keller, robotics

Bloomington Stock Center

Kevin Cook

Kevin Bogart
Jennifer Deal
Megan Deal-Herr
Rachel Andrade
Kathy Matthews
Thom Kaufman

Genome Disruption Project

Hugo Bellen
Joe Carlson
Robin Hiesinger
Roger Hoskins
Allan Spradling

Mass. General, Harvard

Spyros Artavanis-Tsakonas

Doug Dimlich
Glenn Dougherty

Thanks:

Ruth Lehman (GSA)
Ed Ryder (Drosdel- coverage data)

