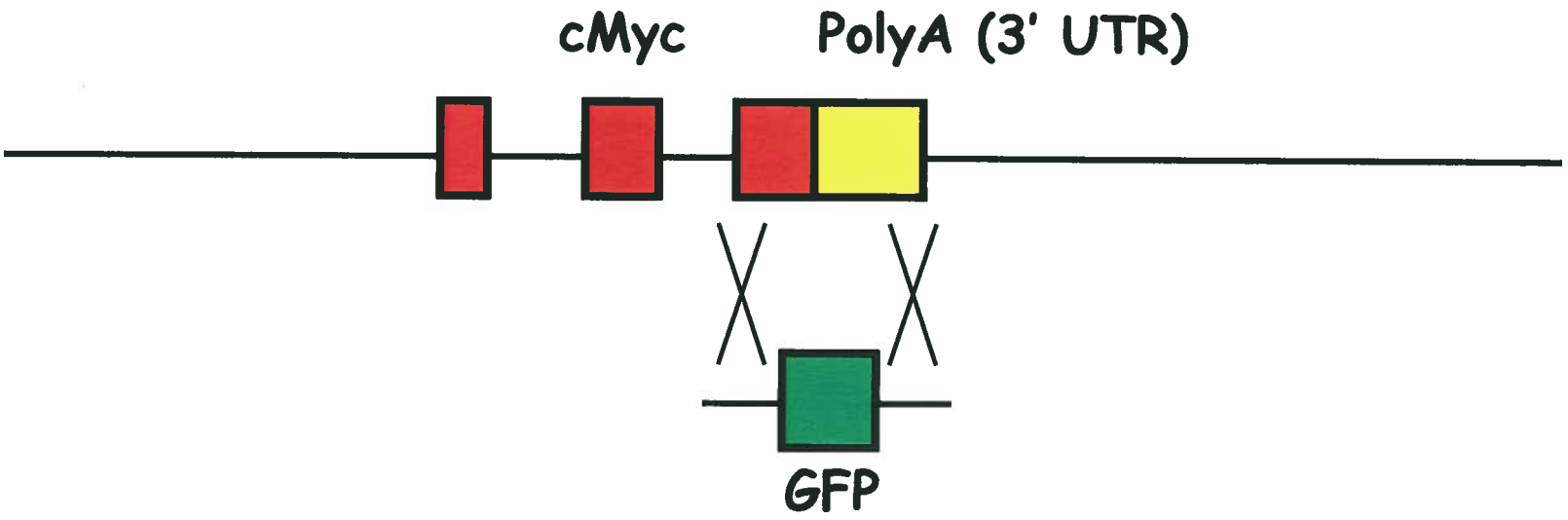
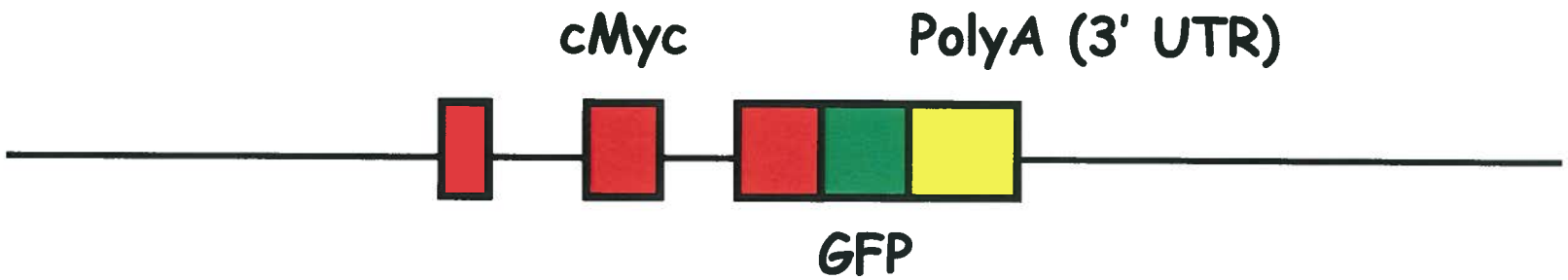
# Bacterial Artificial Chromosome (BAC) Recombineering

Generate cMyc-GFP fusion protein




Homologous Recombination

Generate cMyc-GFP fusion protein




ES cells (knock-in)

Bacterial Artificial Chromosome
(BAC transgene)

Homologous Recombination

# Choosing the right BAC

# BACs

Mouse (Mus musculus) Genome Browser Gateway

Back Forward Reload Stop History Bookmarks Print New Window Home

http://genome.ucsc.edu/cgi-bin/hgGateway

Google

Most Visited ChangePassword Mouse repository UCSC Genome Brows... Ensembl Genome Br... NIAMS Genome Bro... General Conference Keji's analysis MICER and 129 BACs Yellow Sheet

Home Genomes Blat Tables Gene Sorter PCR Session FAQ Help

## Mouse (*Mus musculus*) Genome Browser Gateway

The UCSC Genome Browser was created by the Genome Bioinformatics Group of UC Santa Cruz.
Software Copyright (c) The Regents of the University of California. All rights reserved.

| clade | genome | assembly | position or search term | image width | |
|---|---|---|---|---|---|
| Mammal | Mouse | July 2007 | chr15:62,711,312-62,739,195 | 1500 | submit |

Click here to reset the browser user interface settings to their defaults.

manage custom tracks | configure tracks and display | clear position

## About the Mouse July 2007 (mm9) assembly (sequences)

The July 2007 mouse (*Mus musculus*) genome data were obtained from the Build 37 assembly by NCBI and the Mouse Genome Sequencing Consortium.

### Sample position queries

A genome position can be specified by the accession number of a sequenced genomic region, an mRNA or EST, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the Mouse genome. See the User's Guide for more information.

# BACs

Mouse (Mus musculus) Genome Browser Gateway

Back Forward Reload Stop History Bookmarks Print New Window Home

http://genome.ucsc.edu/cgi-bin/hgGateway

Google

Most Visited ChangePassword Mouse repository UCSC Genome Brows... Ensembl Genome Br... NIAMS Genome Bro... General Conference Keji's analysis MICER and 129 BACs Yellow Sheet

Home Genomes Blat Tables Gene Sorter PCR Session FAQ Help

## Mouse (*Mus musculus*) Genome Browser Gateway

The UCSC Genome Browser was created by the Genome Bioinformatics Group of UC Santa Cruz.
Software Copyright (c) The Regents of the University of California. All rights reserved.

| clade | genome | assembly | position or search term | image width | |
|---|---|---|---|---|---|
| Mammal | Mouse | July 2007 | Myc | 1500 | submit |

Click here to reset the browser user interface settings to their defaults.

manage custom tracks | configure tracks and display | clear position

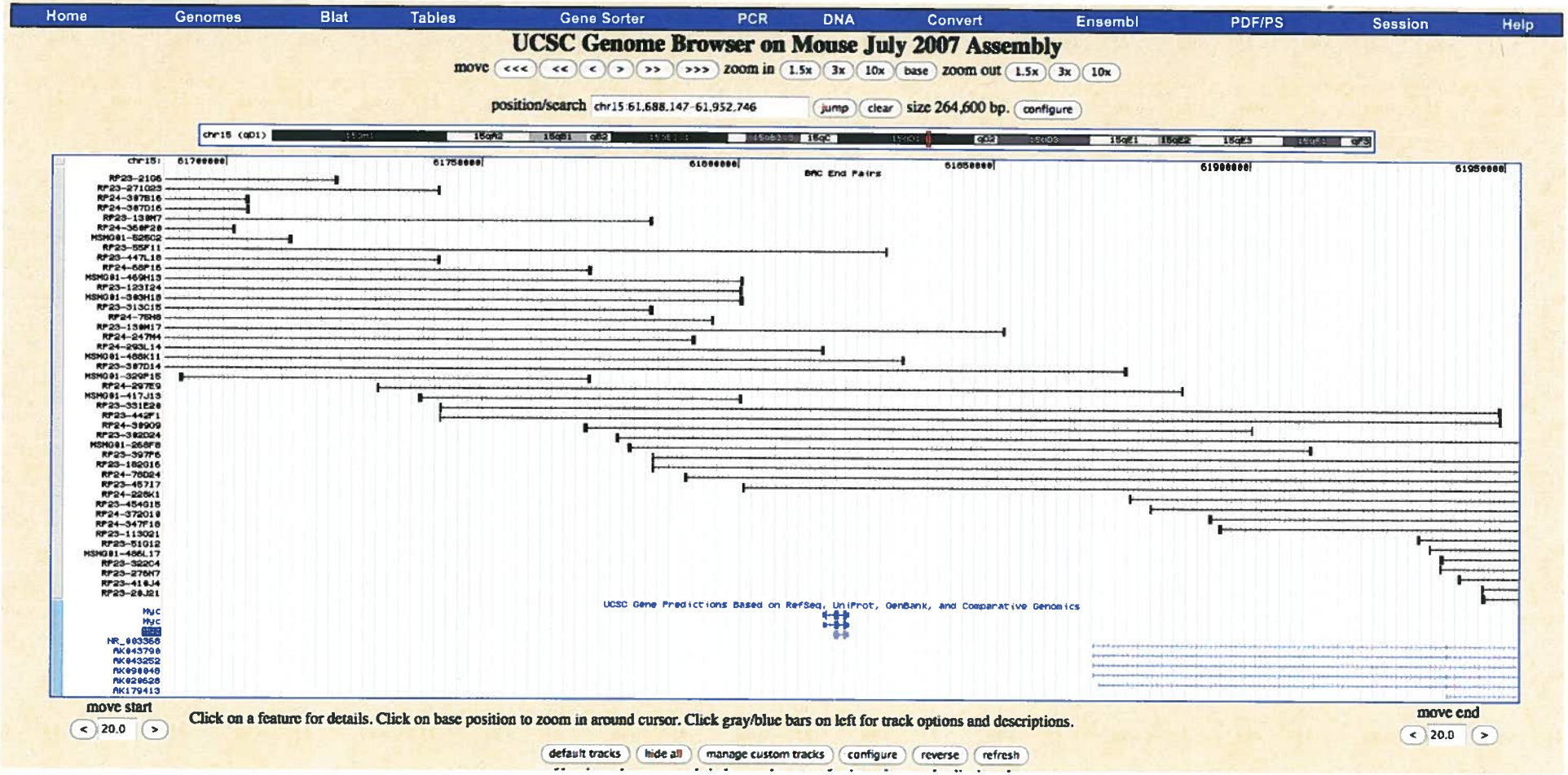
## About the Mouse July 2007 (mm9) assembly (sequences)

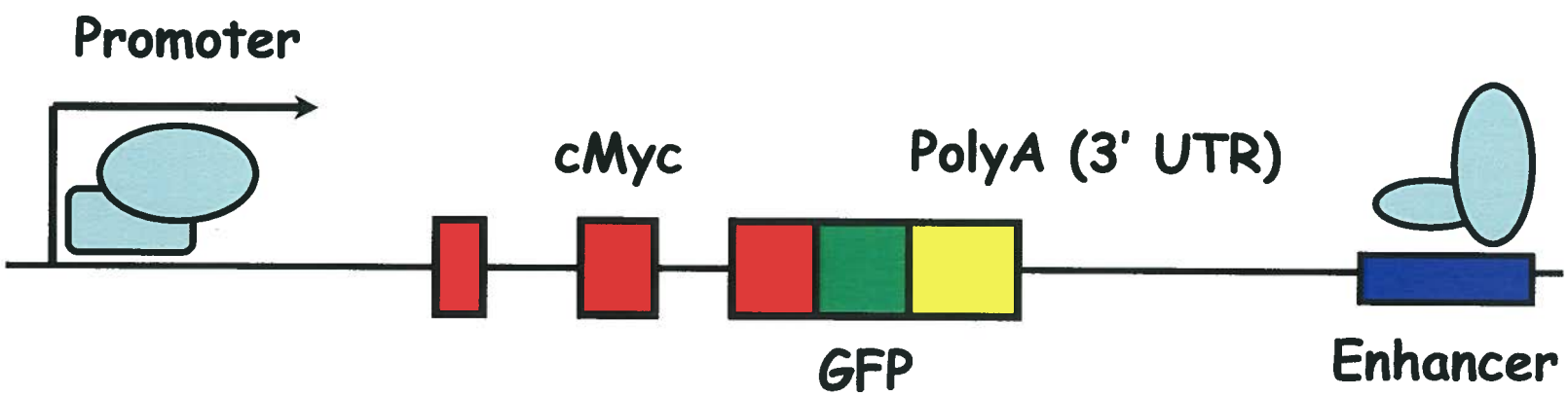
The July 2007 mouse (*Mus musculus*) genome data were obtained from the Build 37 assembly by NCBI and the Mouse Genome Sequencing Consortium.

### Sample position queries

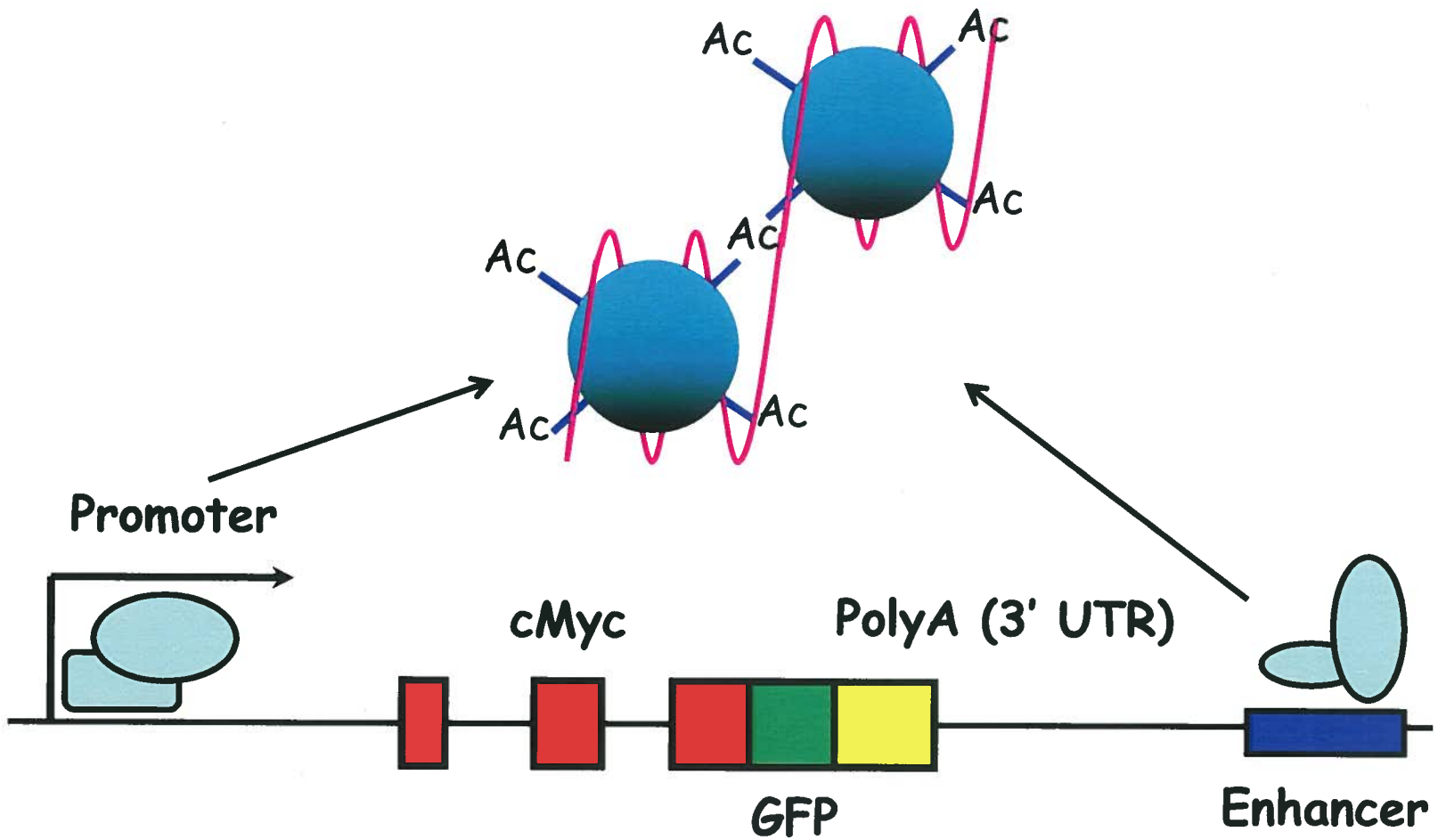
A genome position can be specified by the accession number of a sequenced genomic region, an mRNA or EST, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the Mouse genome. See the User's Guide for more information.

# Which one?


Home
Genomes
Blat
Tables
Gene Sorter
PCR
DNA
Convert
Ensembl
PDF/PS
Session
Help
UCSC Genome Browser on Mouse July 2007 Assembly
move
zoom in
zoom out
position/search
chr15:61,688,147-61,952,746
jump
clear
size 264,600 bp.
configure
BAC End Pairs
UCSC Gene Predictions Based on RefSeq, UniProt, GenBank, and Comparative Genomics
move start
move end
Click on a feature for details. Click on base position to zoom in around cursor. Click gray/blue bars on left for track options and descriptions.
default tracks
hide all
manage custom tracks
configure
reverse
refresh

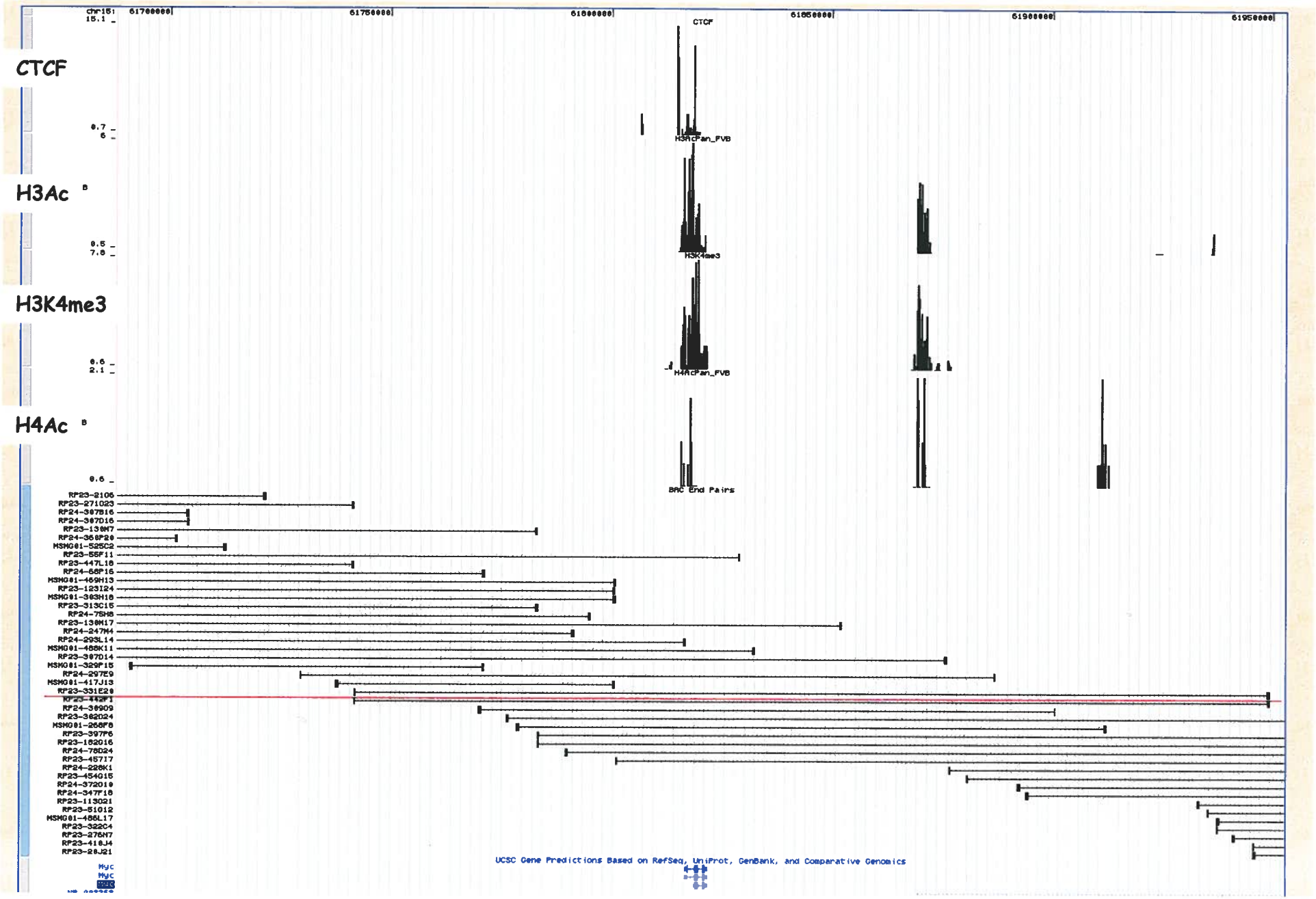



Requirement: All regulatory elements must be present




Histone modifications delineate regulatory elements

Is there a genomic map of histone modifications?


CTCF
H3Ac
H3K4me3
H4Ac
chr15:
61700000
61750000
61800000
61850000
61900000
61950000
CTCF
H3AcPan_FVB
H3K4me3
H4AcPan_FVB
BAC End Pairs
UCSC Gene Predictions Based on RefSeq, UniProt, GenBank, and Comparative Genomics

Home Genomes Genome Browser Blat Tables Gene Sorter PCR Session FAQ Help

## Location of RP23-331E20 using BAC end sequences

### RP23-331E20

**Chromosome:** chr15
**Start:** 61741832
**End:** 61949323
**Length:** 207492
**Strand:** -
**Score:** 1000
**Band:** 15qD1

**This is the only location found for RP23-331E20**

## Genomic alignments of BAC ends:

### AQ973628

| SIZE | IDENTITY | CHROMOSOME | STRAND | START | END | QUERY | START | END | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| 330 | 98.2% | 15 | + | 61741832 | 61742163 | AQ973628 | 1 | 333 | 333 |

View details of parts of alignment within browser window.

### AQ993182

| SIZE | IDENTITY | CHROMOSOME | STRAND | START | END | QUERY | START | END | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| 655 | 99.1% | 15 | - | 61948668 | 61949323 | AQ993182 | 1 | 655 | 655 |

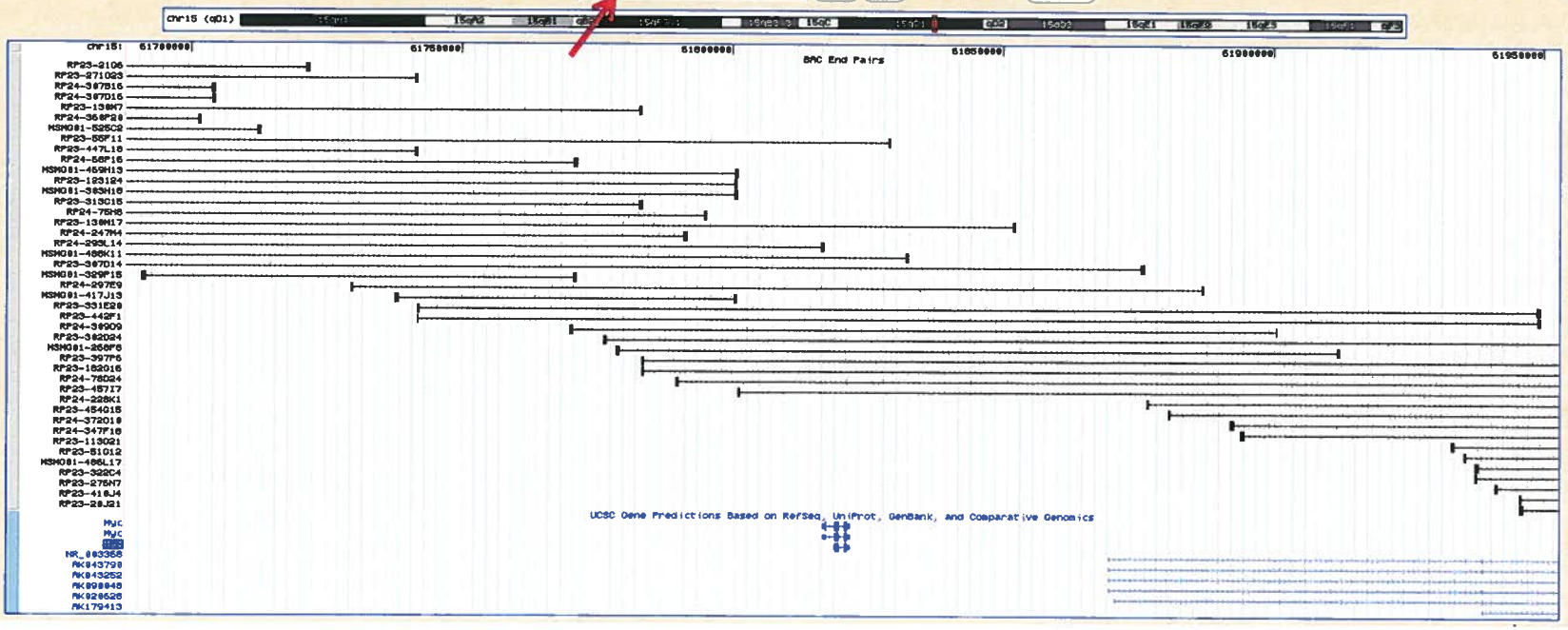
View details of parts of alignment within browser window.

Home | Genomes | Blat | Tables | Gene Sorter | PCR | DNA | Convert | Ensembl | PDF/PS | Session | Help

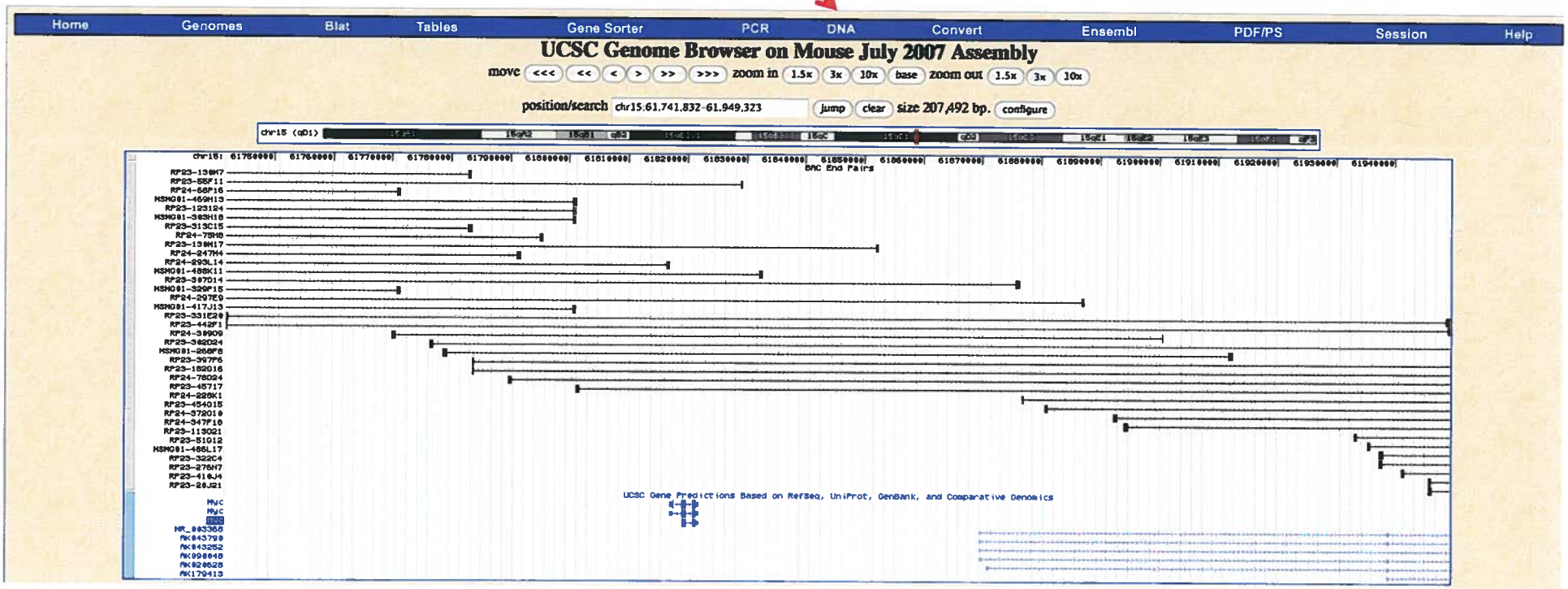
# UCSC Genome Browser on Mouse July 2007 Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

position/search chr15:61741832-61949323 jump clear size 264,600 bp. configure



Get BAC's DNA


Home
Genomes
Blat
Tables
Gene Sorter
PCR
DNA
Convert
Ensembl
PDF/PS
Session
Help
UCSC Genome Browser on Mouse July 2007 Assembly
move
zoom in
zoom out
position/search
chr15:61,741,832-61,949,323
jump
clear
size 207,492 bp.
configure
BAC End Pairs
UCSC Gene Predictions Based on RefSeq, UniProt, GenBank, and Comparative Genomics

Mozill

Back Forward Reload Stop History Bookmarks Print New Window Home http://genome.ucsc.edu/cgi-bin/hgc?hgsid=1185

Most Visited ChangePassword Mouse repository UCSC Genome Brows... Ensembl Genome Br... NIAMS Genome Bro... General Conference

```
>mm9_dna range=chr15:61741782-61949373 5'pad=50 3'pad=50 strand=+ repeatMasking=none
ATTAATTTCTCTGGGCCTCCATTTACCCTTTTTTTAAAATTGAGTCAATTC
CACGCAGACACTTGACCCAACTCACTCTCTAGAACTCTCTACTATAAGCA
TGAGGCTTACTAATGGAACTGATAACTCTGAAGCCCATGTGGGAATATGC
ACCCTTTCAGGTAGCCCTAGCAAATAGTATTTATTAAACATGTGGATGGT
TACTGCTCACCTGGTTCTGGCTTCTGGGGACTTTGCTAGGGGCCATCTCA
TAGCAGCAAAACAATAATGATGATGATGATGAAAATGATGATCGCACTAT
TTTCAATTCATTTTCATCAGGAAAATACAGGTGAGTGGTGACCTCTCGGG
CCATTTCCAAGGCTGTGTTTGTGTAAAGACATGACAGGCCTGTGAGCTCA
TTTTCTTAAGCAACGAGATTATGAATTTCAAAGAAAGAAAAGAGGTGTTA
TGTAAAAGTCACCTCTTATCATGAACTTTCTCCAGTGCTATTTCTTTGCT
TTCATGTTGTGCTTCTTCTGCATGTTTGAACTTGCCCTGGAACTCTTACT
CACTCAAAACACAAACATATTGCAAAGCAAGTTTGCTTCCCCTGAGTTCC
ACAAGGTAAGGTATCTGAGTTTTTTCTTCCTTTCTTCCCTTCTTTTCTCA
TTCTGTCTGGTCTAAGCCCACTGTACCTTTTGAAATGTAATCTCTCAACC
CACTGATTCTTTTAAGCCTCTGTCTTTCAAGGATTTCAAGTGAAGTACTC
CAACAGCGGCAACAACATTCGGCAACATCTTTGATTGCCTGTATATGACA
GTCCGTTTCTTTCCTCCGTCATCTTCTCACATTAACTAAAGTCTCTCTCA
GAGAGCATCAGCCAGCCAACTCCATGGGGCTATGATAAGAAGCACCTTTC
CACAAACTGGACACAACATCTCTACTTGCTGAGTCTCATCACAAGCAGCA
GGAGAGGACTGCAAGAGCACATGACTCTTTGGACCCCACCCATGCTCCCT
CATTACAAACAGAGAAGGGAAAGCTGCAGAGGGCATGACCATTAGAAGAT
AGCATCTAACGCTGAGAATTTAACTGATGCTAAGATGCTTACCTCACAAG
CACAAGGCCCAGGGGTCGATCCCCAAGACCATACACAAAAGTAAAAGACA
ATCGCTCCCTTTGTTATTTTAAACCCCATTCCCTACATGTTAATTCTCAA
AGCTCAGCCAAGAAGTTGATTCCAAAACTTCATATTCTTTATGATGCACC
ACTCAGTTCTTTTATGCTGTTGCCAAAATCAGTTGCTGGGGTTTTTTTTT
TCTCAAAGATTATGTTATCAGTTGTATATTAATAGTGCACCTCCCCAAGG
AACTATCAGAGACAATATTCCTAGGTTACACAAATAGGATCTGGATGGAA
AATAAGACTTTCATGGATAATCAAATTAAGATTACATCACATGTAATTAA
AATGGGCTCCAATCCAATGCGATTGATATCGTTATAAGAAGAGTCACAGA
CAGACACAAAGAGGCAGAAAGAGAATGTGTCCACAAGCCAAGGAAACTCA
AGAACTTCAAGTAATGCCAACATGCTCAGAGAACAGGGACAAACCAGATT
TACCCTAAAGTCTTGTCTGGGAGCCTGGCCCTGACAACACTTGCATCTTG
AGTTTCTGCCCTTCTACATTATGGTTATTTTGGCCATTCCAATTGTGGTT
CTTTATTGCTGAAACCCTAAGAACTTAACAGACCGTTCTTACTGTTCTTC
CCTCCCTCTGTTCCTTTCACTGGGACAACCCTTCTGCTTGTGTCACCACA
CTGTGACTGACTCTCCTATATATTCAGGGTGTAGAGACAGTGACCTTGGG
TTCCCAAATGTCTGCTTCTCCCAACTTAGCACTCACCGAGATGCACTCGG
ATCATCTCGCTCACAAGACTGTGTGTTAGTCATCTGTTGATACTTGCTTG
GCAGACACTCAATAAATGATGAAAAACCTGAAGGGAACATTCAGATAGGT
GGCACTTGAAGACAGAGAGATGAGTAAGGAACCAACCTTACCCTGGAGGA
ATATGGTAATTTTCAAAGGGAAAAGTGGTAACAAAGACGTAGCTCAACAT
CTTACCTCAGTATGCTGATCTCTTTCATTGGCTTTCCAGAGTCACCATCC
ATGACAGTCTCTCAAATGTTCACACTATAATCCACTTCATAAGGAGTTCA
AACTATGGAGCCCATCAGTGGAGTGGTAGCAGGGAAAAAATACTAAGATG
CCTGGAATAAGTCCTCTTGTGATGCTAGCAAAGACTTTTGACTAAATGTT
CCAAATGCATTCATCATTTAGAGTGAACATAGCATAAGTCCAATCTCCTG
ATCACTAAATATAGAACAGCCACTCCTAAGTACCCTTCTAGGCTCAGTTA
TGGCTGCATCCTTGGGAAATTGTTCCTGGCCCCTAAATGCATTACACCCC
ACTGTTTCTGTCTATTACATCTGAACCCACTATCATCTGGGACCCACATC
TTACCCTTGTGAAATCACAGGACTCATTGTAACCACTAAGGGTCTATTTA
ATCAAATGGACATGAGTATAGGAGAGAAACTCAGGTCATGCCTGTCATGC
CTTTGAGACTGAAGCATTTGAGAGCTCTCAACAATTCAGAAAGTAAATCT
TTGGAAAATTCAGACATCCTGAGACCACACATAAACCCATAATTTTTTTT
AAATCTAATCACTAGACAGAAGCAGAGAAGAGGCAGCCATTTTTCCATTA
TGAAGCTCATACCTGTGTGAAATTCAATTCCCCACTGAAGCCTTAAAAAA
GTTGAGTAACAAAGAAGCTAAAGGGATCTGCAACCCTATAGGTGCAACAA
CATTATGAACTAACCAGTACCCCGGAGCTCTTGACTCTAGCTGCATATGC
ATCAAAAGATGGCCTAGTCGGCCATCACTGGAAAGAGAGGCCCATTGGAC
ACGCAAACTTTATATGCCCCAGTACAGGGGAACGCCAGGGCCATAAAAGG
GGAGTGGTGGGTAGCGGAGAGGGGTGGGTGGCTATGGGGGACTTTTGG
TATAGCATTGCAAATGTAAATGAGCGAAATACCTAATTAAAAAAAAAAAA
AAAACAAAAAAAAAGTTGAGTAATTCTTGACCAAACTATGGTTTTTTAGCT
TCCTCAGCTTCTCCCCTAGACAGATGCCTTATCATATGCAGATGCTTATC
TTAATAGATAATAGTTATTGACTGGATCAATCATTCCCTCTACCTTCCAG
CCTGGCTTTGGCCCACATCTCCCGCCTTCTCCTTGCAGGCCCCTGATTCA
TGCCTTATGAAACCTGACAGCAAGGGATCTTGACCTGGTTCCCAGGGAAC
GACTGGGCAGCCTCAGAGAACAAGACAGAGAGCCAGGGTGGAGGCTCACA
CAGCCTGGAGGCCAGTAGGGAGGGAGGTTTCCCTGGAAGACAGACGTAAG
CACTGCTCTGGGCATACACACTGACTGTTTACCTGGCAGACGGGCACAGC
ATGTTGCAACACCTCCTTGCCGACACATTATAGCCTGATTATTACTCAGC
AGTGGCAGGCCTAGATGAGATGACCTACAATGGTGAGCCTTAACAAGAAA
GCTTGTCTGCCACACCCAGCGTACACCTTTTACCCCGGGAACTTGCCTGA
AGCCAGTTCATATAGCAGCTGCCAATAGCAGCTGGAAGCTATATAGGGAA
AGGTTGCTGGCCCAAGCAGCGGCAGAAGGTTCTAGGACTTGCTTCTGCTA
```

Done

NCBI **Clone Registry**

PubMed Entrez BLAST OMIM Taxonomy Structure

Search CloneRegistry for Go

Clone Registry Home Page

Clone Nomenclature

BAC End Report

Distributor Information

**Clone-ID:** RP24-68I7 **CurrState:** Free
**Comment:** CloneInfo processing

**Buying the BAC**

**FingerPrint Information:** Yes

**End-Sequence Information**

| Accession | Seqlen(bp) | Repeat | Hit | End |
|---|---|---|---|---|
| AZ712347 | 334 | No | No | SP6 |
| AZ712350 | 566 | No | No | T7 |

**Library Information**

**Name:** RPCI mouse BAC library 24
**Organism:** Mus musculus **Type:** BAC
**Approved:** Unknown **Sex:** Male
**Source:** spleen/brain **Rsite:** MboI

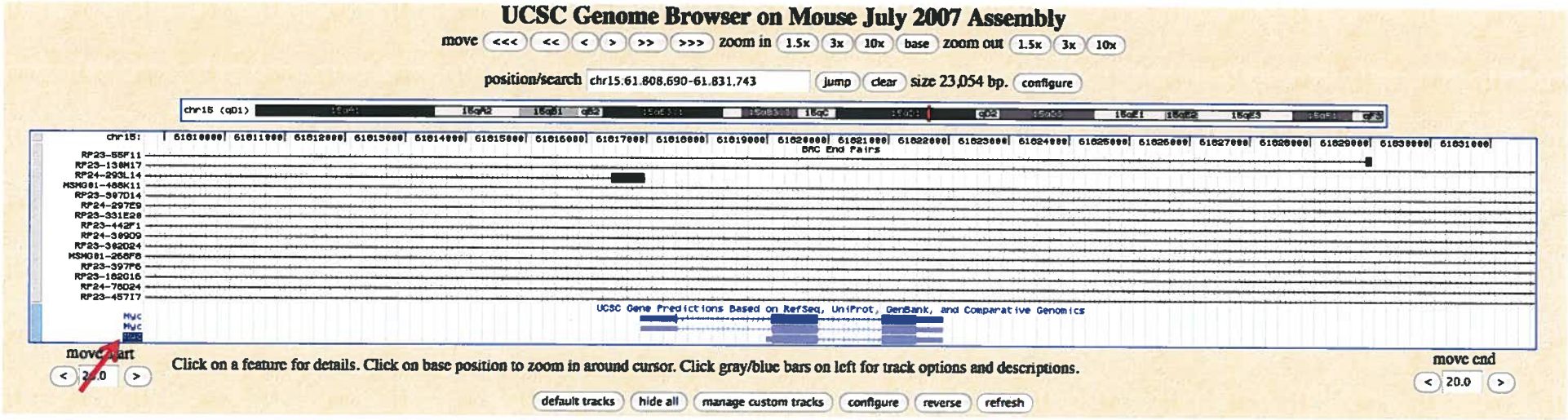
**Distributor Information**

Available from clone library
**Name:** Children's Hospital Oakland Research Institute (CHORI)
**Contact:** Pieter J. de Jong
**Email:** BACPACorders@chori.org
**Web:** http://bacpac.chori.org
**Phone:** 510-450-7917

**Transaction History**

| G. Center | Trans-State | Chrom | Result | Submit-date |
|---|---|---|---|---|
| SERVER | Free | | Accept | 03/01/2001 |

# Designing the recombineering construct



## Mouse Gene Myc (uc007vyi.1) Description and Page Index

**Description:** myc proto-oncogene protein
**RefSeq Summary (NM_010849):** The protein encoded by this gene is a multifunctional, nuclear phosphoprotein that plays a role specific target genes. Mutations, overexpression, rearrangement and translocation of this gene have been associated with a variety translation initiations from an upstream, in-frame non-AUG (CUG) and a downstream AUG start site result in the production of tv deprivation, there is a specific and dramatic increase in the synthesis of the non-AUG initiated protein, suggesting its importance i
**Strand:** + **Genomic Size:** 2940 **Exon Count:** 2 **Coding Exon Count:** 2

| | | | | | |
|---|---|---|---|---|---|
| Page Index | Sequence and Links | UniProt Comments | CTD | Microarray | RNA Structure |
| Protein Structure | Other Species | GO Annotations | mRNA Descriptions | Other Names | Model Information |
| Methods | | | | | |

## Sequence and Links to Tools and Databases

| | | | | | |
|---|---|---|---|---|---|
| Genomic Sequence (chr15:61,818,977-61,821,916) | | | mRNA (may differ from genome) | | Protein (439 aa) |
| Gene Sorter | Genome Browser | Protein FASTA | Proteome Browser | VisiGene | Table Schema |
| Allen Brain Atlas | Ensembl | Entrez Gene | ExonPrimer | GeneCards | Gepis Tissue |
| PubMed | Stanford SOURCE | UniProt | | | |

## Comments and Description Text from UniProt (Swiss-Prot/TrEMBL)

**ID:** MYC_MOUSE
**DESCRIPTION:** Myc proto-oncogene protein (c-Myc) (Transcription factor p64).
**FUNCTION:** Participates in the regulation of gene transcription. Binds DNA both in a non-specific manner and also specifically t
**SUBUNIT:** Efficient DNA binding requires dimerization with another bHLH protein. Binds DNA as a heterodimer with MAX. In
**SUBCELLULAR LOCATION:** Nucleus.
**DEVELOPMENTAL STAGE:** Expressed in the proliferating cells of the developing CNS and the epidermis. In the spinal cord a differentiating cells at the ventral portion of the intermediate zone. Also detected in the roof plate and in the neural crest. At 14.5 c epidermis. At 17 dpc, expression is confined primarily to the proliferative malphigian layer of the epidermis and to the dermal pap
**PTM:** Phosphorylated by PRKDC (By similarity).
**DISEASE:** Overexpression of C-Myc is implicated in the etiology of a variety of hematopoietic tumors.
**SIMILARITY:** Contains 1 basic helix-loop-helix (bHLH) domain.

## Comparative Toxicogenomics Database (CTD)

Genomic Sequence Near Gene

Back Forward Reload Stop History Bookmarks Print New Window Home http://genome.ucsc.edu/cgi-b Google

Most Visited ChangePassword Mouse repository UCSC Genome Brows... Ensembl Genome Br... NIAMS Genome Bro... General Conference

Home Genomes Genome Browser Blat Tables Gene Sorter PCR Session FAQ Help

**Genomic Sequence Near Gene**

# Get Genomic Sequence Near Gene

Note: if you would prefer to get DNA for more than one feature of this track at a time, try the Table Browser using the output format sequence.

## Sequence Retrieval Region Options:

- [x] Promoter/Upstream by 1000 bases
- [x] 5' UTR Exons
- [x] CDS Exons
- [x] 3' UTR Exons
- [x] Introns
- [x] Downstream by 1000 bases
- (•) One FASTA record per gene.
- ( ) One FASTA record per region (exon, intron, etc.) with 50 extra bases upstream (5') and 50 extra downstream (3')
  - [ ] Split UTR and CDS parts of an exon into separate FASTA records

Note: if a feature is close to the beginning or end of a chromosome and upstream/downstream bases are added, they may be truncated in order to avoid extending past the edge of the chromosome.

## Sequence Formatting Options:

- (•) Exons in upper case, everything else in lower case.
- ( ) CDS in upper case, UTR in lower case.
- ( ) All upper case.
- ( ) All lower case.
- [ ] Mask repeats: (•) to lower case ( ) to N

submit

Mozilla Firefox

Back Forward Reload Stop History Bookmarks Print New Window Home http://genome.ucsc.edu/cgi-b Google

Most Visited ChangePassword Mouse repository UCSC Genome Brows... Ensembl Genome Br... NIAMS Genome Bro... General Conference

```
>mm9_knownGene_uc007vyi.1 range=chr15:61817977-61822916 5'pad=0 3'pad=0 strand=+ repeatMasking=none
gaccgacacggaggtgcgtcccgcccgccaatccccggcggcgatcgcaa
cccgtccctgagccttttaagaagttgctattttggctttaaaaatagtg
atcgtagtaaaatttaagcctgacccccgcggcactaggacttgatgttg
ggctagcgcagtgaggagaagcaaaattgggacagggatgtgaccgattc
gttgacttgggggaaaccagagggaatcctcacattcctacttgggatcc
gcgggtatccctcgcgcccctgaattgctaggaagactgcggtgagtcgt
gatctgagcggttccgtaacagctgctaccctcggcggggagagggaaga
cgccctgcacccagtgctgaatcgctgcagggtctctggtgcagtggcgt
cgcggtttagagtgtagaagggaggtgtctcttattatttgacacccccct
cccctttatttcgagaggcttgtgatagccggagactgagctctctcct
ccaagtcagcaatcggaaagaaaagccggcaaaggaaggaaggggcgcg
ctgggggtggagaaagaggagggcggagaggggcggcggcgccggctggg
taggagcgcggcgacggcgcgaatagggactcggacccggtcggcggcgc
agagagccggcacacgggaggggggccgagcgacgcggcgcctctcgcctt
tctccttcaggtggcgcaaaactttgcgcctcggctcttagcagactgta
ttccctacagtcgcctccctcagcctctgaagccaaggccgatggcgatt
cctgggcgtctgcagggctaagtccctgctcgaaggaggcggggactcgg
agcagctgctagtccgacgagcgtcactgatagtagggagtaaaagagtg
catgcctcccccccaaccacacacacacacacacacacacacacacacac
acacacacacacacttggaagtacagcacgctgaaaggggagtggttcag
GATTGGGTACGCGCTGCGCCAGGTTTCCGCACCAACCAGAGCTGGATAA
CTCTAGACTTGCTTCCCTTGCTGTGCCCCCTCCAGCAGACAGCCACGACG
ATGCCCCTCAACGTGAACTTCACCAACAGGAACTATGACCTCGACTACGA
CTCCGTACAGCCCTATTTCATCTGCGACGAGGAAGAGAATTTCTATCACC
AGCAACAGCAGAGCGAGCTGCAGCCGCCCCGCGCCCAGTGAGGATATCTGG
AAGAAATTCGAGCTGCTTCCCACCCCGCCCCTGTCCCCGAGCCGCCGCTC
CGGGCTCTGCTCTCCATCCTATGTTGCGGTCGCTACGTCCTTCTCCCCAA
CGGAAGACGATGACCGCGCGCGGTGGCAACTTCTCCACCGCCGATCAGCTG
GAGATGATGACCGAGTTACTTGGAGGAGACATGGTGAACCAGAGCTTCAT
CTGCGATCCTGACGACGAGACCTTCATCAAGAACATCATCATCCAGGACT
GTATGTGGAGCCGTTTCTCAGCCGCTGCCAAGCTGGTCTCGGAGAAGCTC
GCCTCCTACCAGGCTGCGCGCAAAGACAGCACCAGCCTGAGCCCCGCCCG
CGGGCACAGCGTCTGCTCCACCTCCAGCCTGTACCTGCAGGACCTCACCG
CCGCCGCGTCCGAGTGCATTGACCCCTCAGTGGTCTTTCCCTACCCGCTC
AACGACAGCAGCTCGCCCAAATCCTGTACCTCGTCCGATTCCACGGCCTT
CTCTCCTTCCTCGGACTCGCTGCTGTCCTCCGAGTCCTCCCCACGGGCCA
GCCCTGAGCCCCTAGTGCTGCATGAGGAGACACCGCCCACCACCAGCAGC
GACTCTCgtaagctaccccattcacagcagggtaggaagcgagaggttgg
atggacctccttctccaccactcattggcattaattcaattggcctccgg
ggctccccctttctttcccttctgtctaagagctcttcatccctggattc
ccgtgcttcagctatccctcccctcccctcccccccccccccccagaccgc
cctggttgtcaccccccaccccccaccccttcccaccccctcaggaatttcat
ttgggttttaaatcttctggcttatctttcagctccatccactccctttt
acccctcctaagcattttaattaccctgggaagggtgtgaatgaggataa
aggaactgatctggagggggggtgaattacctgcttcttttcttgactgcc
agaagaatatttgaattaatggatacgtttgtctaagacccaagagaag
caatgacagaagctgggacagcctttatagccttagagccaggcactagt
gaaagttcctaaagaattgaagagctgggctcttttgggtgttttgtttt
gctttgtttttcgtcctcttttgaacacttcaaagcaaattctgttcaa
tttggacttctcccccccgtcccaacactcccccaacaccaggacgtttgg
caaagctgcaagactttttttttttttttttttttttaattgtgcttcca
gtaaaatagggagttgctaaagtcatagcaagagatttgcagctatccct
cacgggacctgaaaggttctcggtaaagtcccttaaaaataggaggtgct
tgggaaatgtgctttgctttgggtgtgtctgaagcctcattaaatcttag
gtaagaattggcaaggataccatatcctggtacatggtaattttctcacc
tgtgccctaaccctgttctgcctttctgggagaagggaagatggtgtctg
gatctgattcttactttcttcccttccaacttggtatttggatagcatc
ggtcaaatcctatgtatagcgtccgggattcaggaggcgtggctaactgt
gatcttccacttcctcccttacagAAGAAGAGCAAGAAGATGAGGAAGAA
ATTGATGTGGTGTCTGTGGAGAAGAGGCAAACCCCTGCCAAGAGGTCGGA
GTCGGGCTCATCTCCATCCCGAGGCCACAGCAAACCTCCGCACAGCCCAC
TGGTCCTCAAGAGGTGCCACGTCTCCACTCACCAGCACAACTACGCCGCA
CCCCCCTCCACAAGGAAGGACTATCCAGCTGCCAAGAGGGCCAAGTTCGA
CAGTGGCAGGGTCCTGAAGCAGATCAGCAACAACCGCAAGTGCTCCAGCC
CCAGGTCCTCAGACACGGAGGAAAACGACAAGAGGCGGACACACAACGTC
TTGGAACGTCAGAGGAGGAACGAGCTGAAGCGCAGCTTTTTTGCCCTGCC
TGACCAGATCCCTGAATTGGAAAACAACGAAAAGGCCCCCAAGGTAGTGA
TCCTCAAAAAAGCCACCGCCTACATCCTGTCCATTCAAGCAGACGAGCAC
AAGCTCACCTCTGAAAAGGACTTATTGAGGAAACGACGAGAACAGTTGAA
ACACAAACTCGAACAGCTTCGAAACTCTGGTGCATAAACTGACCTAACTC
```

Done

Location of RP23-331E20 using BAC end sequences

Back Forward Reload Stop History Bookmarks Print New Window Home http://genome.ucsc.e GooQ

Most Visited ChangePassword Mouse repository UCSC Genome Brows... Ensembl Genome Br... NIAMS Genome Bro...

Home Genomes Genome Browser Blat Tables Gene Sorter PCR Session FAQ Help

## Location of RP23-331E20 using BAC end sequences

### RP23-331E20

**Chromosome:** chr15
**Start:** 61741832
**End:** 61949323
**Length:** 207492
**Strand:** -
**Score:** 1000
**Band:** 15qD1

**This is the only location found for RP23-331E20**

## Genomic alignments of BAC ends:

### AQ973628

| SIZE | IDENTITY | CHROMOSOME | STRAND | START | END | QUERY | START | END | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| 330 | 98.2% | 15 | + | 61741832 | 61742163 | AQ973628 | 1 | 333 | 333 |

View details of parts of alignment within browser window.

### AQ993182

| SIZE | IDENTITY | CHROMOSOME | STRAND | START | END | QUERY | START | END | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| 655 | 99.1% | 15 | - | 61948668 | 61949323 | AQ993182 | 1 | 655 | 655 |

View details of parts of alignment within browser window.

NCBI Sequence Viewer v2.0

Back Forward Reload Stop History Bookmarks Print New Window Home http://www.ncbi.nlm.

Most Visited ChangePassword Mouse repository UCSC Genome Brows... Ensembl Genome Br... NIAMS Genome

Location of RP23-331E20 usin... NCBI Sequence Viewer v2.0

NCBI Nucleotide

PubMed Nucleotide Protein Genome Structure PMC Taxonomy

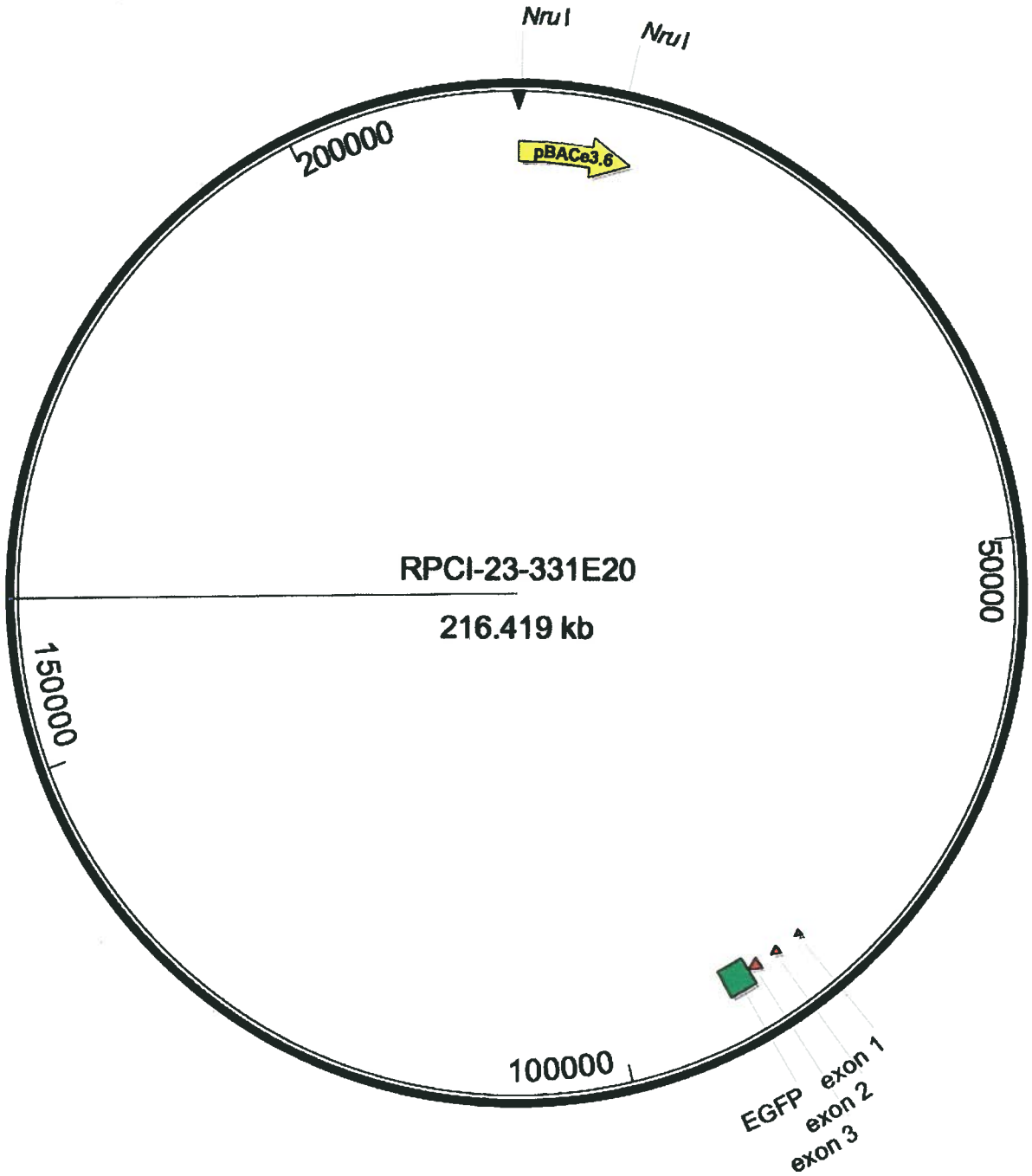
Search GSS for AQ973628 Go Clear

Display GenBank Show 20 Send to Hide: sequence all but gene, CDS and mRNA feature

Range: from begin to end Reverse complemented strand Features: + Refresh

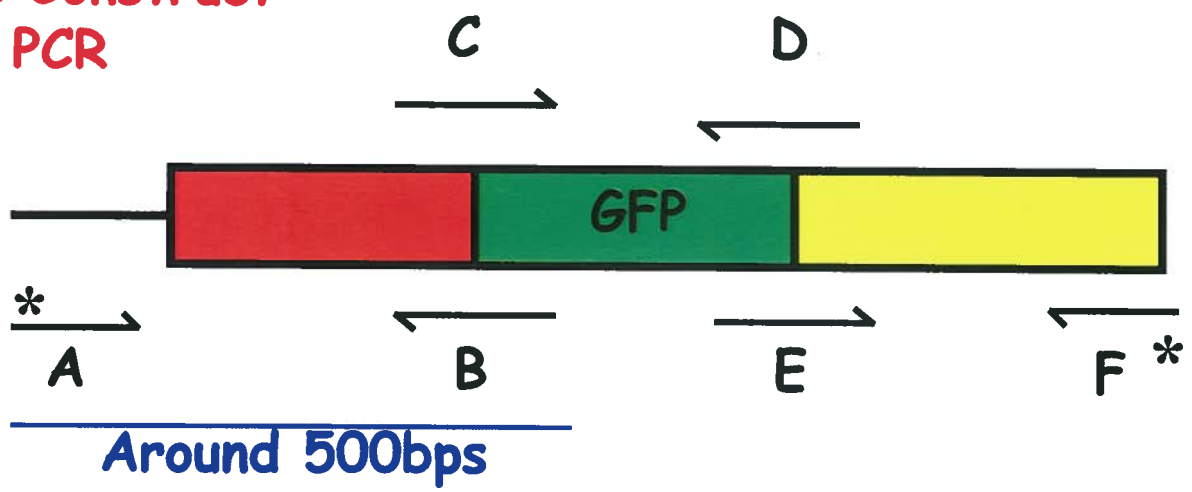
1: AQ973628. Reports RPCI-23-331E20.TJ...[gi:6804081]

```
LOCUS       AQ973628                 333 bp    DNA     linear   GSS 28-JAN-2000
DEFINITION  RPCI-23-331E20.TJ RPCI-23 Mus musculus genomic clone RPCI-23-331E20
            , genomic survey sequence.
ACCESSION   AQ973628
VERSION     AQ973628.1  GI:6804081
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 333)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 331  row: I  column: 20
            Seq primer: SP6
            Class: BAC ends
FEATURES             Location/Qualifiers
     source          1..333
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-331E20"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN
        1 cacgcagaca ctgacccaa ctcactctct agaactctct actatatagc atgaggctga
       61 ctaatggaac tgataactct gaatcgcatg tgggaatatg cacccttca ggtagcccta
      121 gcaaatagta tttattaaac atgtggatgg ttactgctca cctggttctg gctctgggg
      181 actttgctag ggggcatctc atatcagcaa aacaataatg atgatgatga tgaaaatgat
      241 gatggcacta ttttcaattc atttcatca ggaaaataca ggtgagtggt gacctctcgg
      301 gccatttcca aggctgtgtt tgtgtnanga gat
//
```

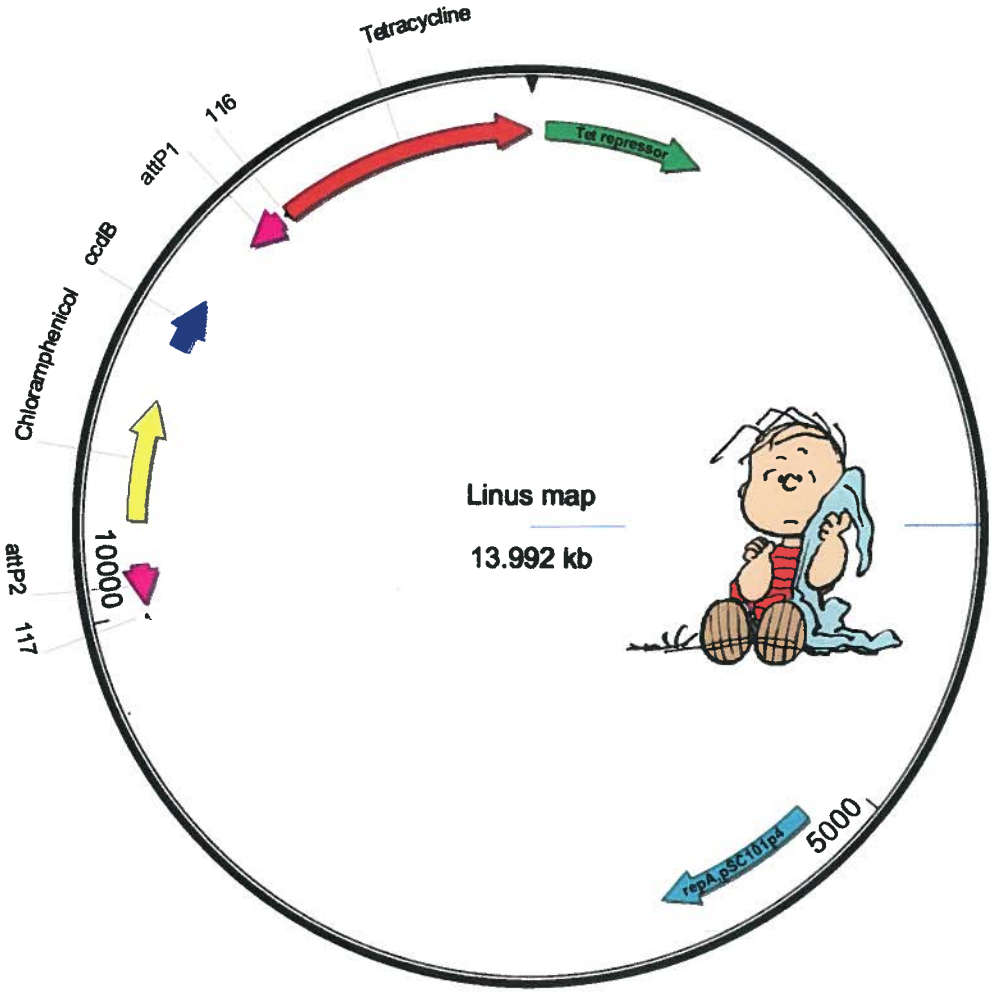

NruI
NruI
pBACe3.6
200000
RPCI-23-331E20
216.419 kb
50000
150000
100000
EGFP
exon 1
exon 2
exon 3

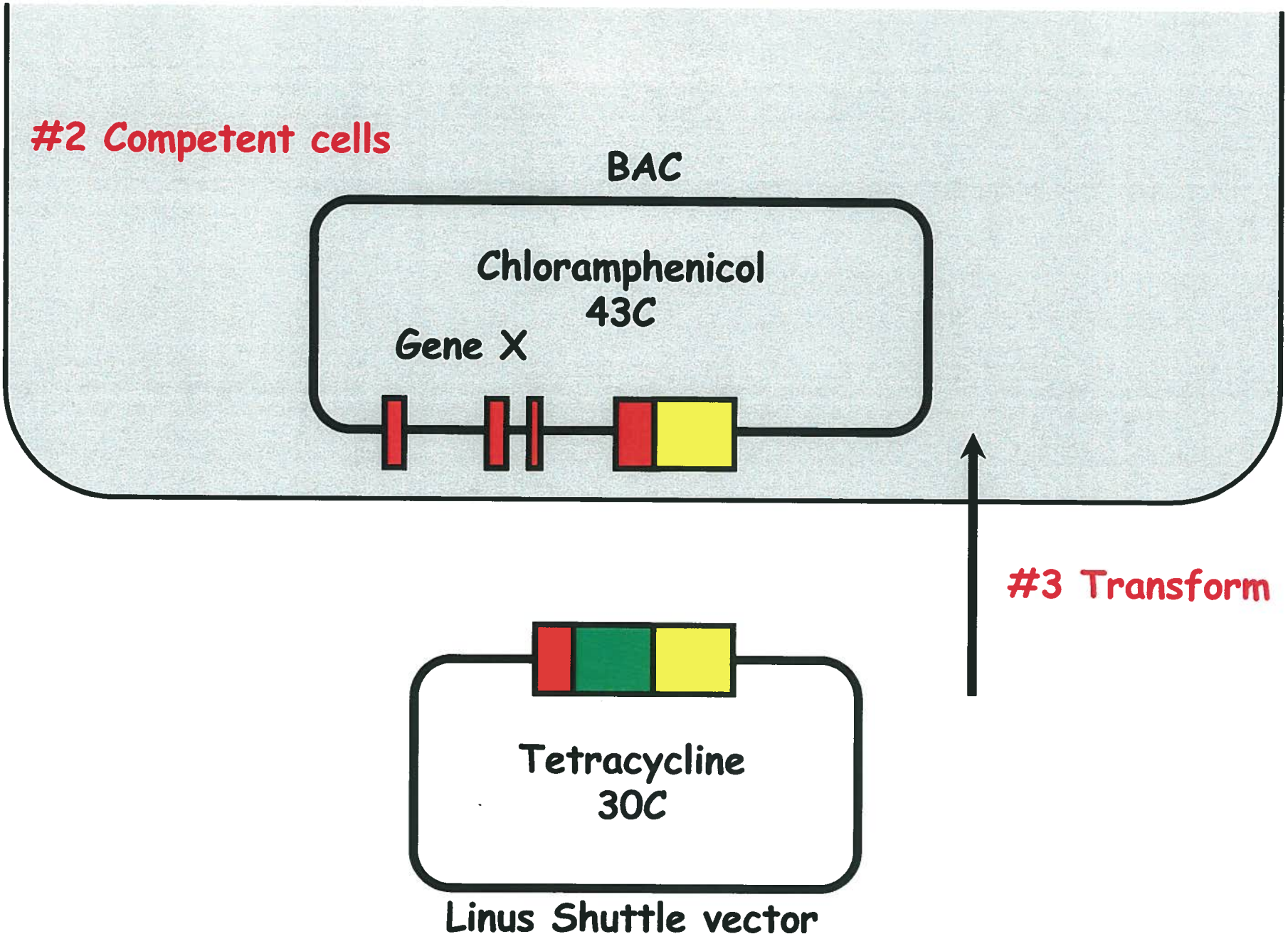
# Modifying the BAC

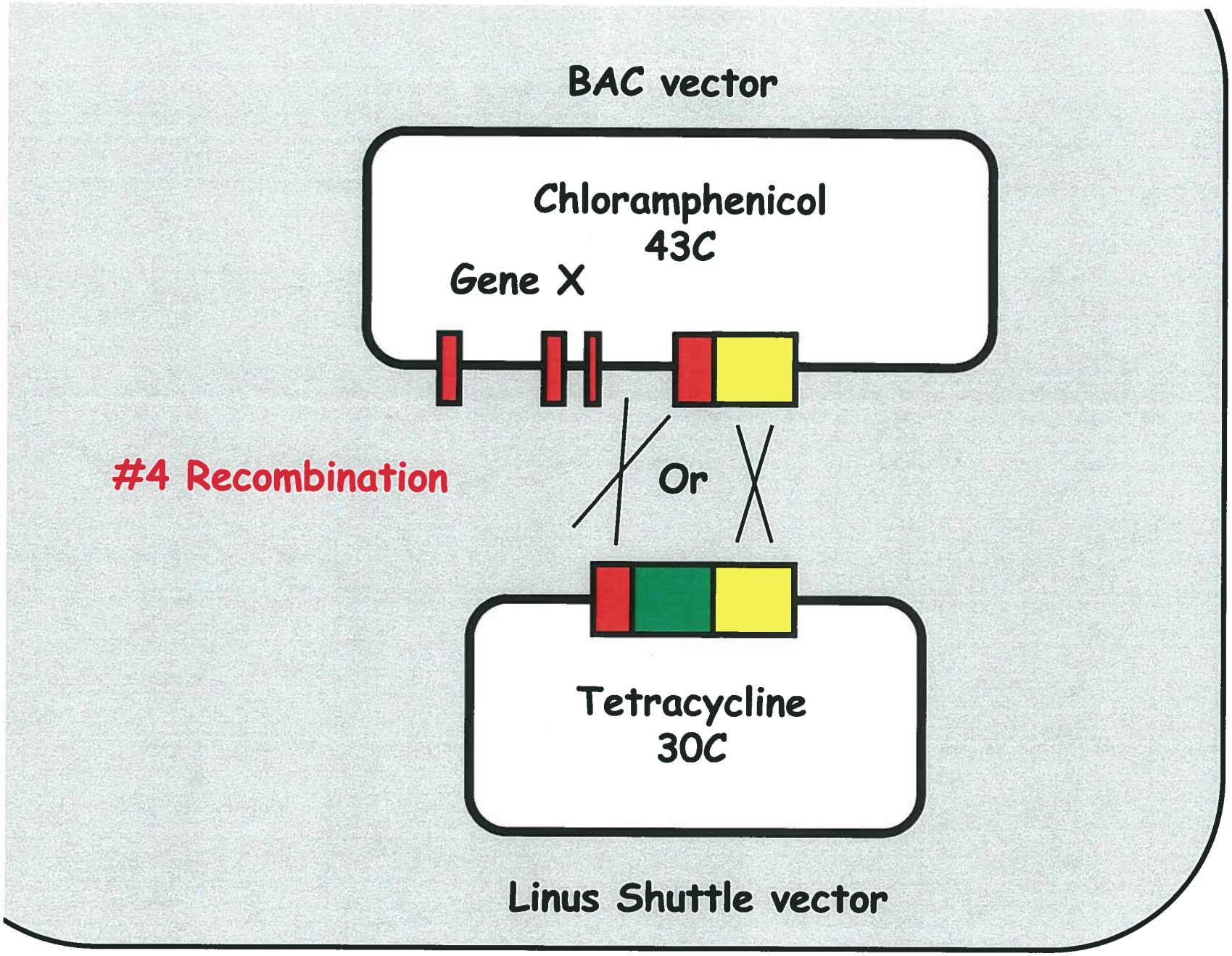
# #1 Build the Construct
## Triple PCR



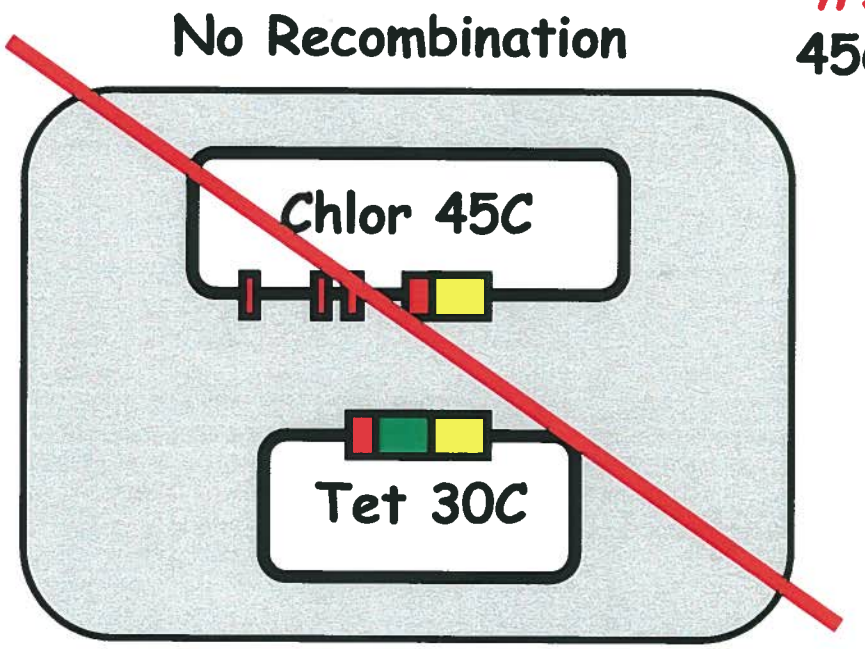

*attP sites

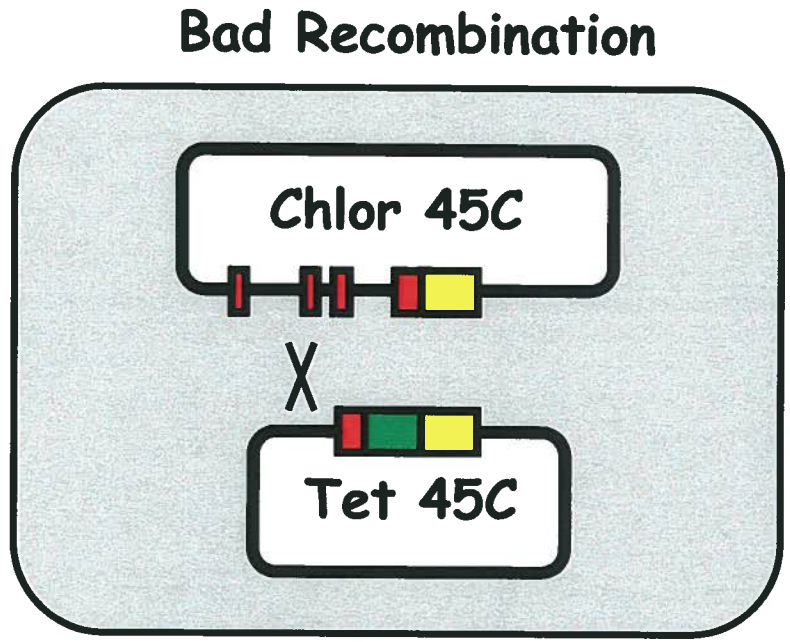

Tetracycline
Tet repressor
116
attP1
ccdB
Chloramphenicol
attP2
117
10000
Linus map
13.992 kb
5000
repA pSC101p4

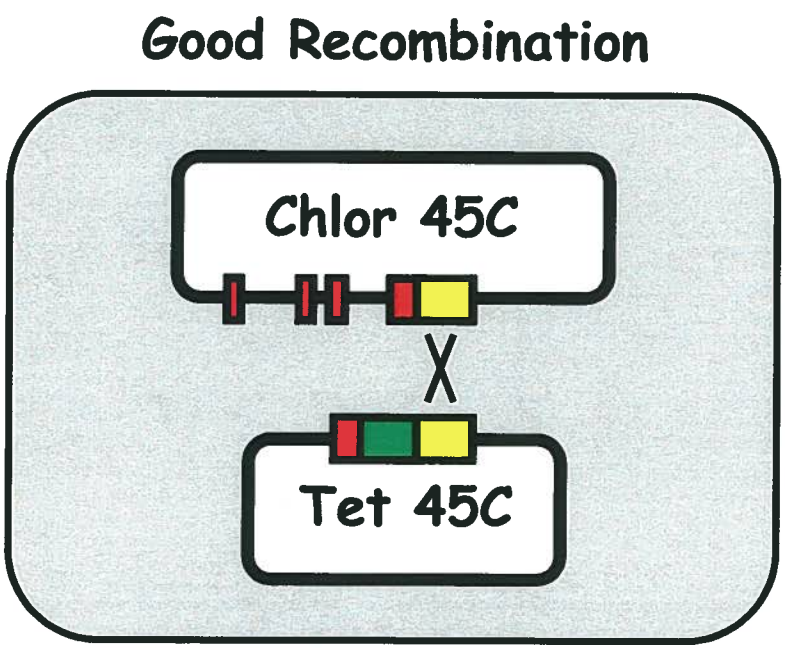

#2 Competent cells
BAC
Chloramphenicol
43C
Gene X
#3 Transform
Tetracycline
30C
Linus Shuttle vector


BAC vector
Chloramphenicol
43C
Gene X
#4 Recombination
Or
Tetracycline
30C
Linus Shuttle vector

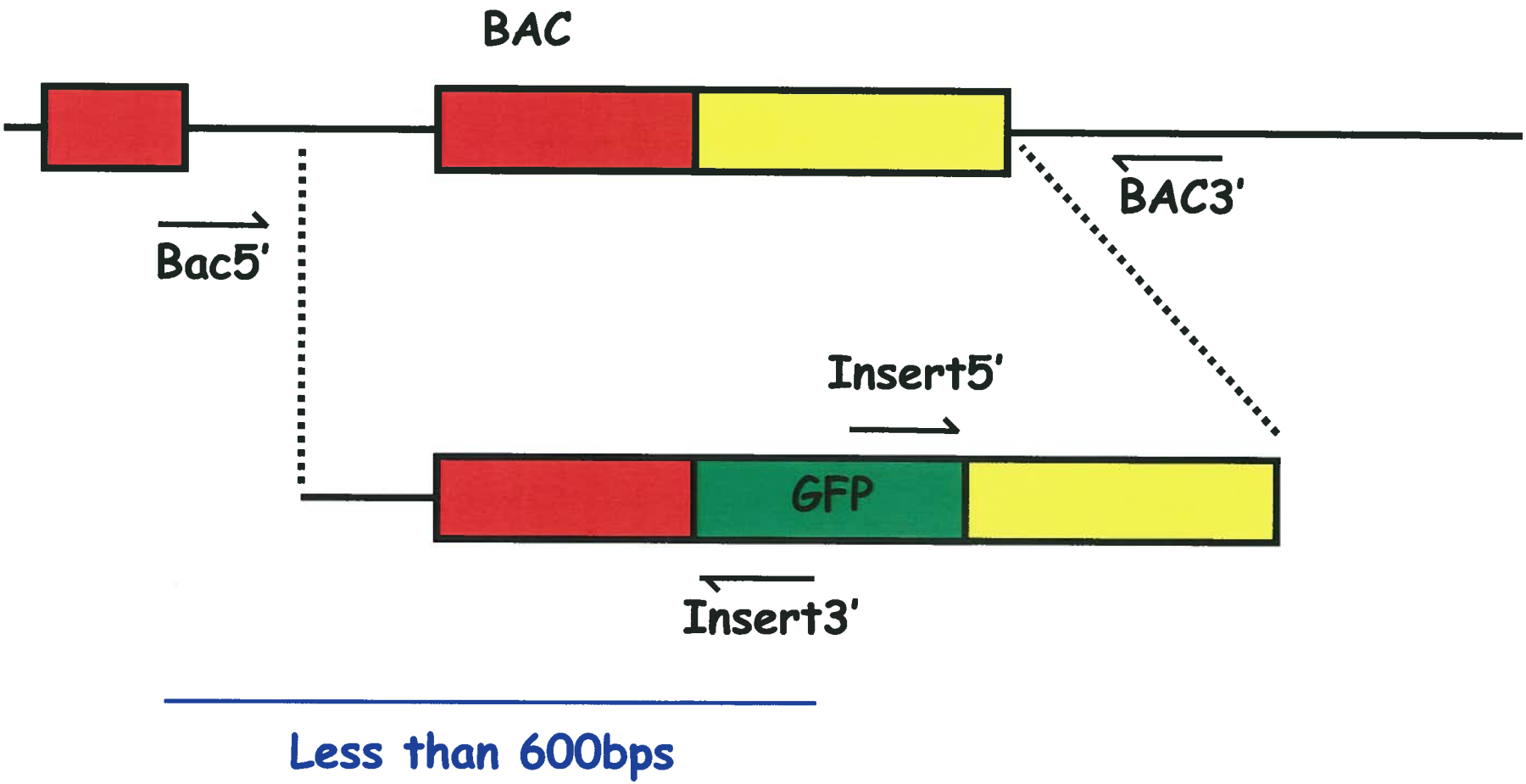
## #5 Selection
45C Tet/Chlor



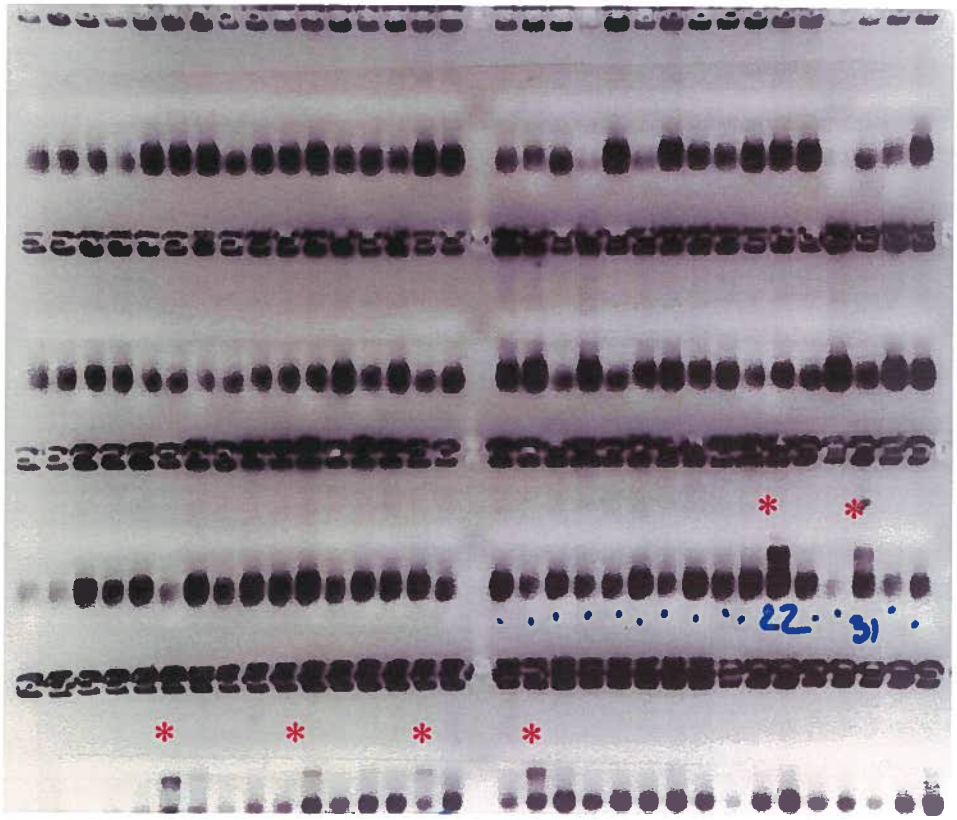



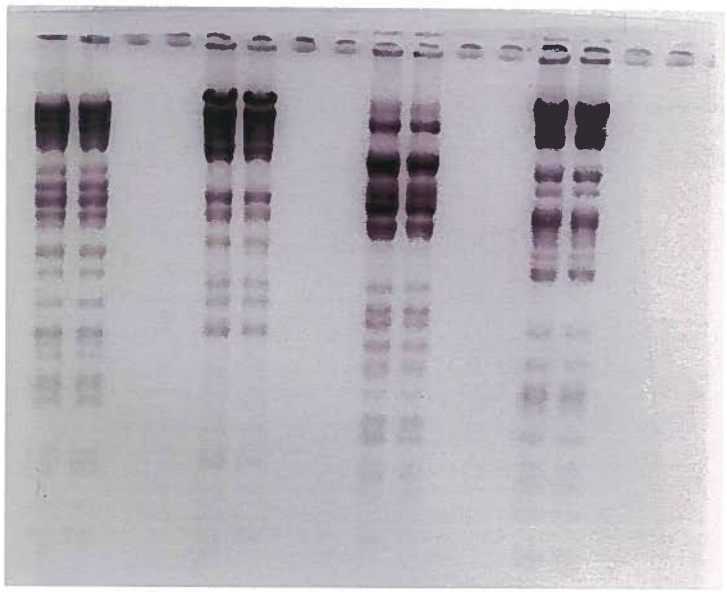
# #6 Screening

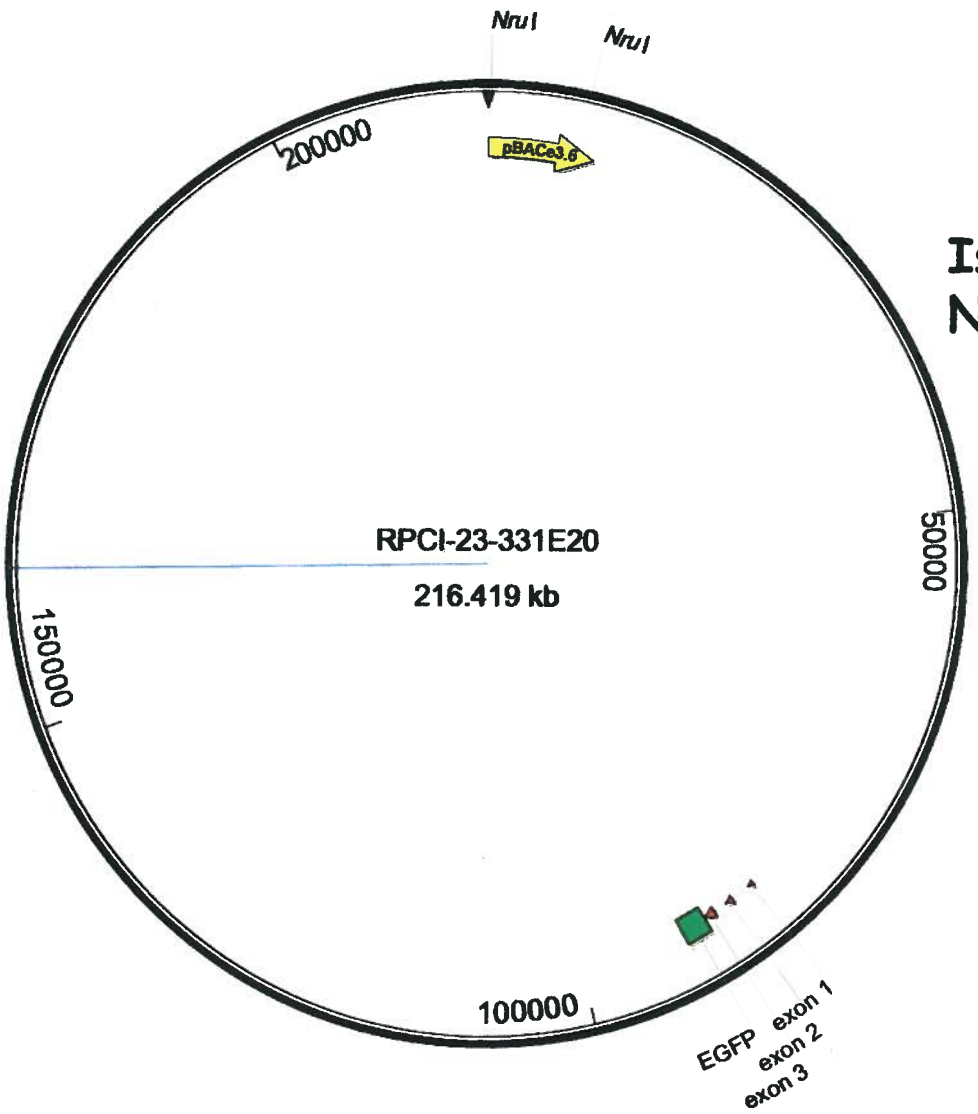


## PCR screening




## Final constructs






Isolate BAC DNA using NucleoBond BAC100 kit

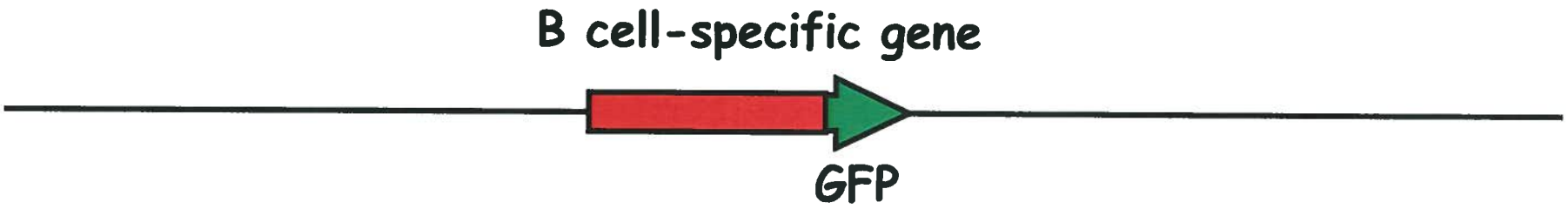
Restrict DNA with NruI or NotI to drop pBACe3.6

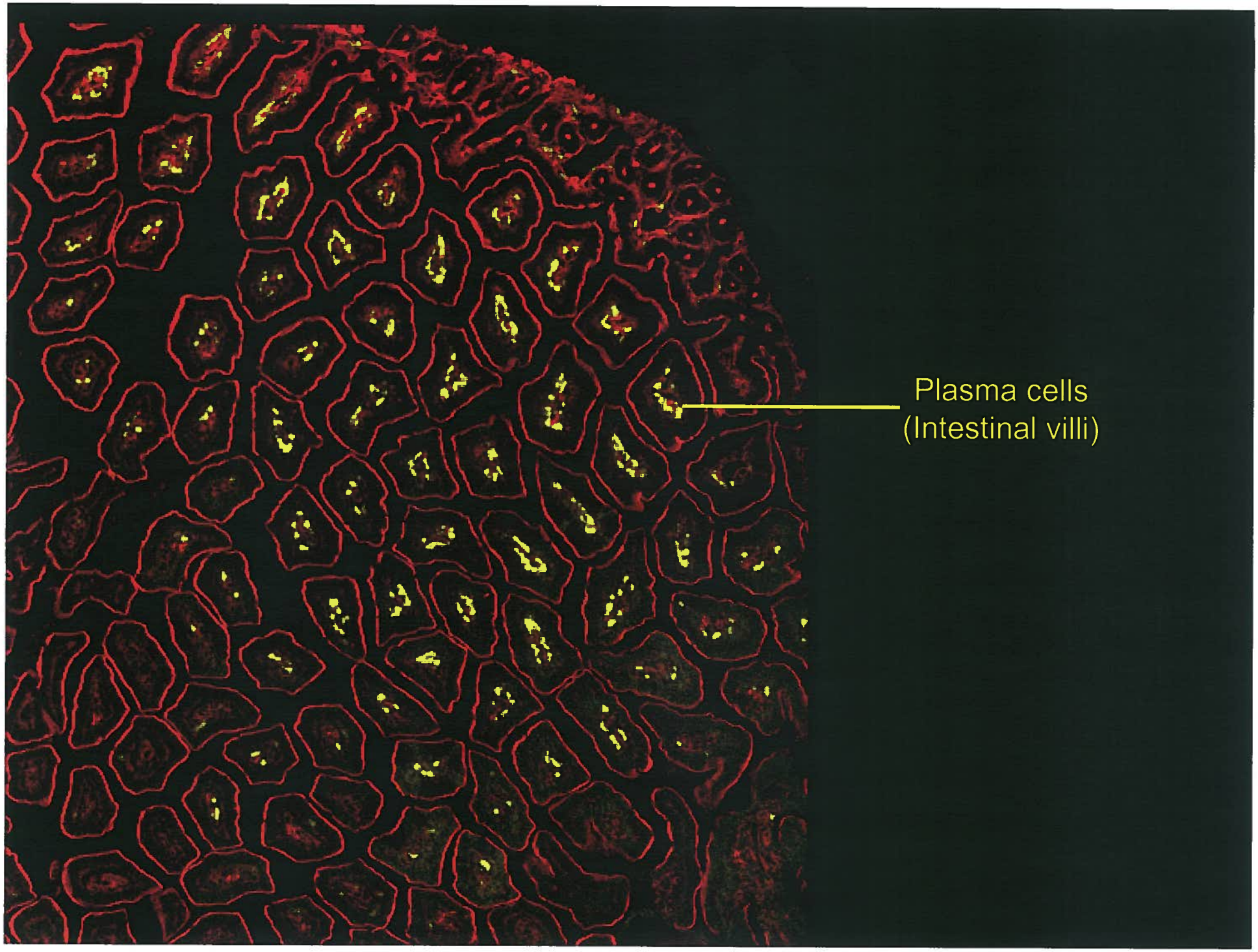
Dialyze DNA

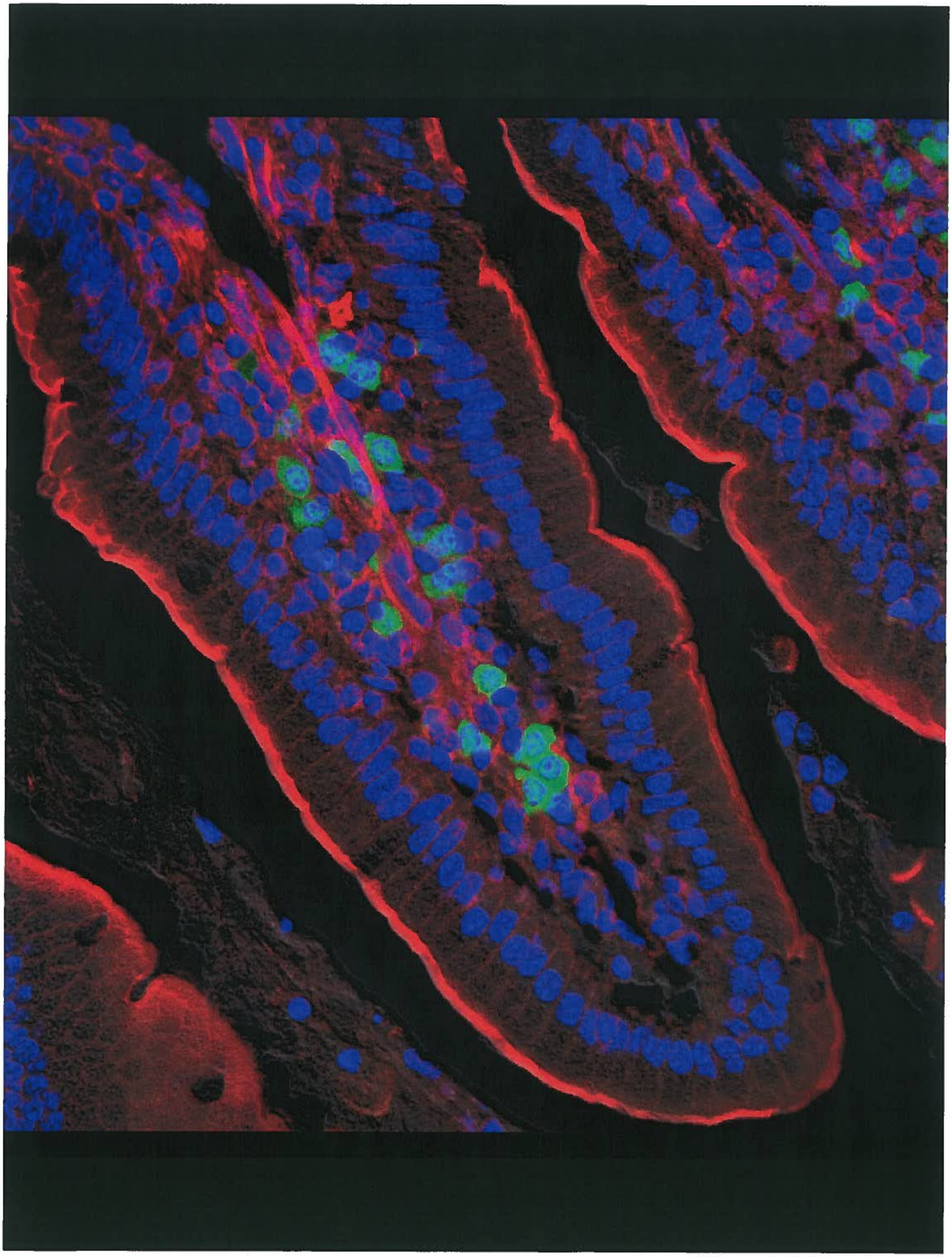
Microinject

# Applications

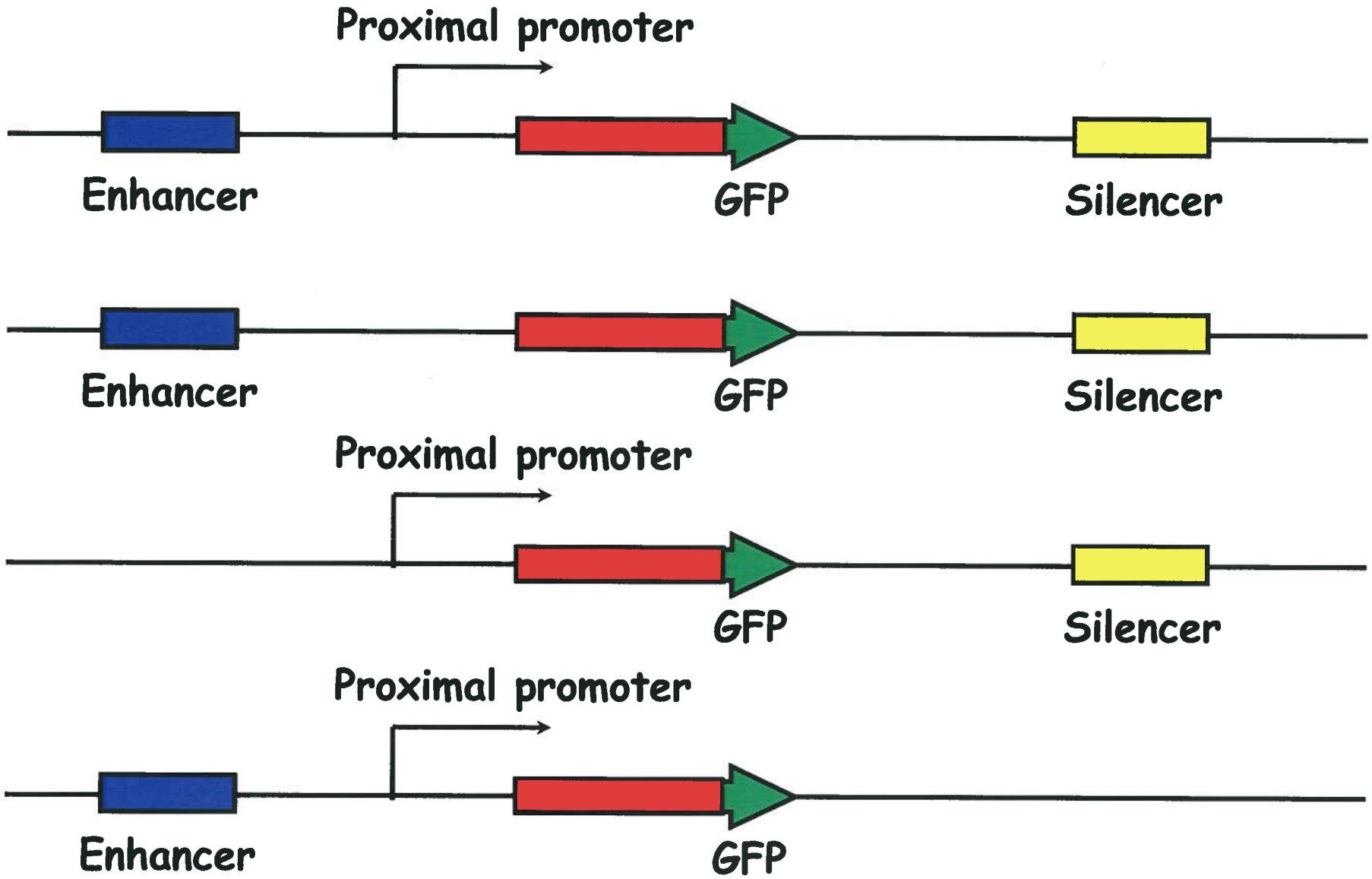
## Applications: 1 in-vivo labeling



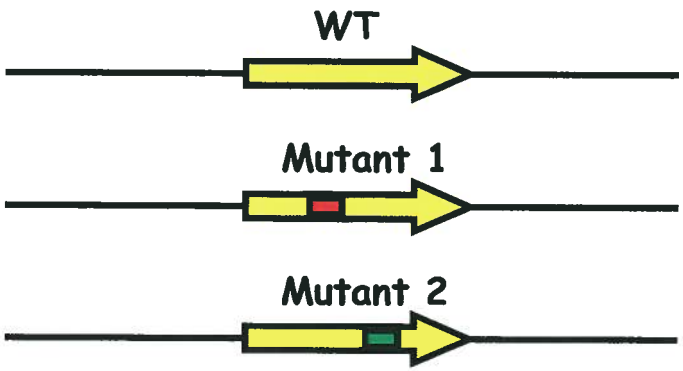

Plasma cells
(Intestinal villi)



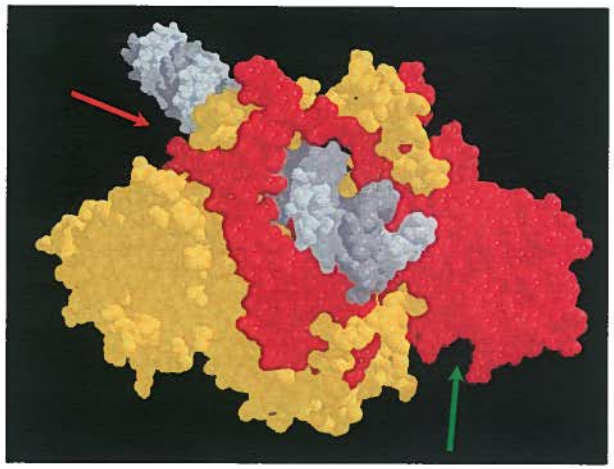
# Applications: 2 Transctriptional studies



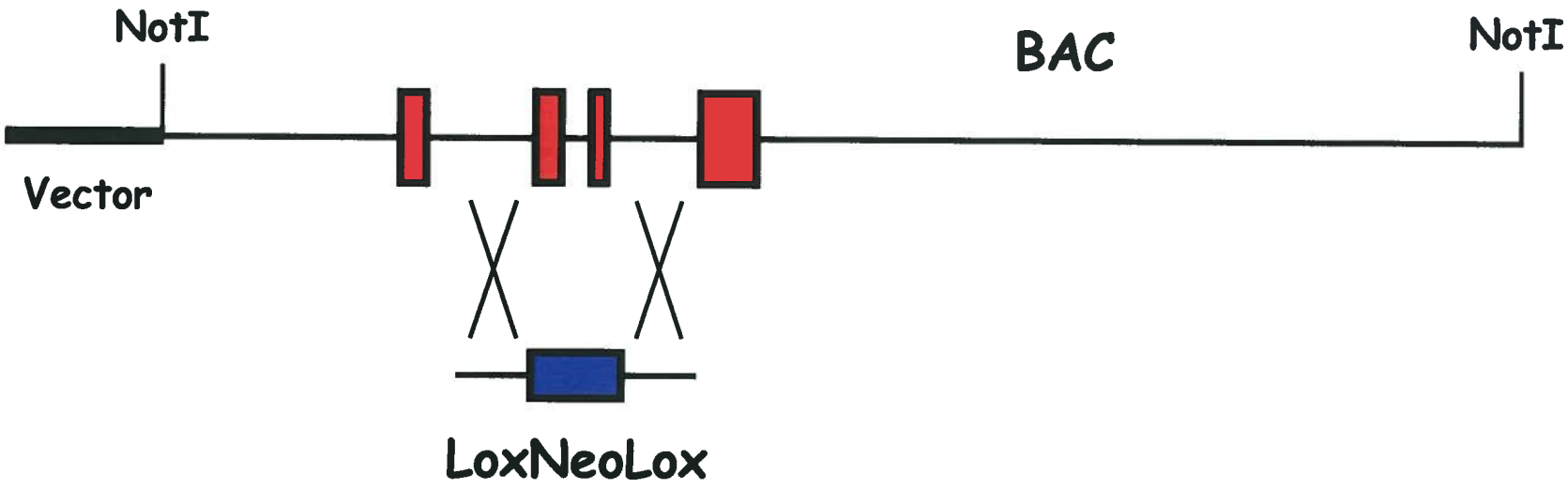
# Applications: 3 knock-out reconstitution



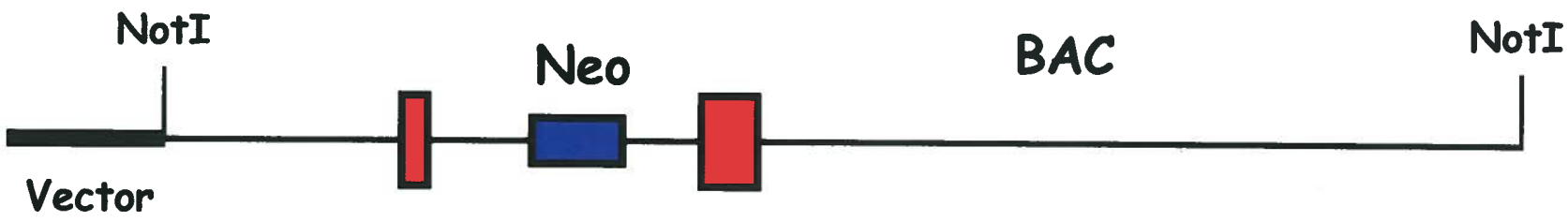
-/-



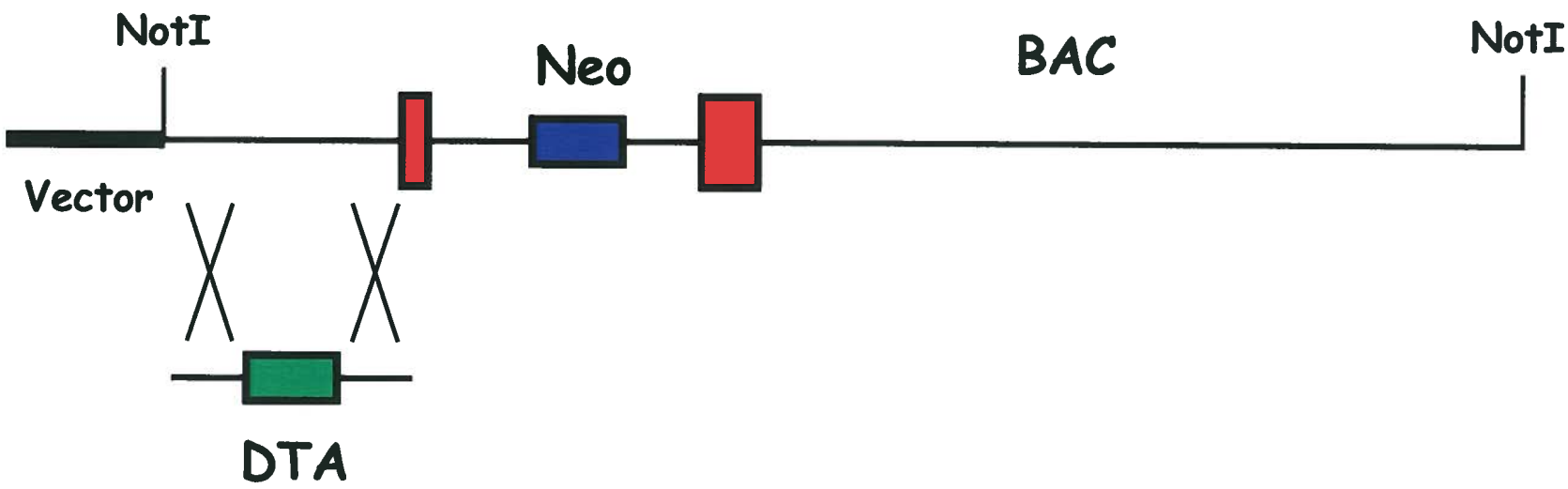
# Applications: 4 knock-out constructs



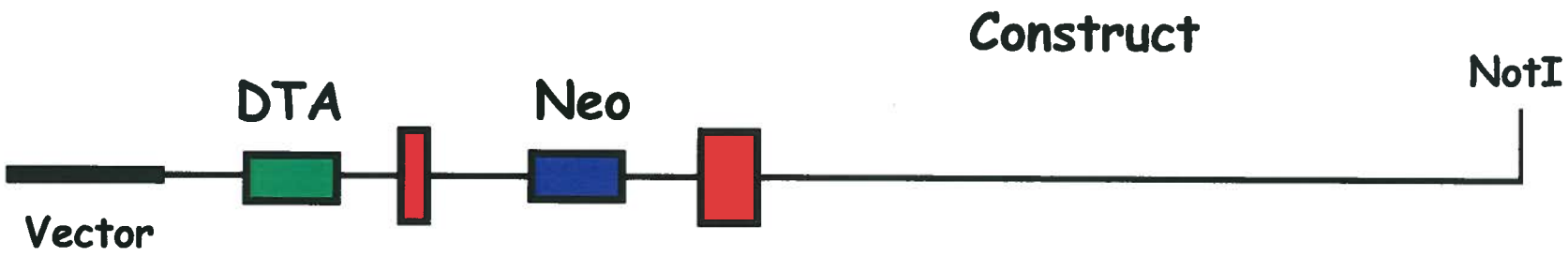
# Applications: 4 knock-out constructs



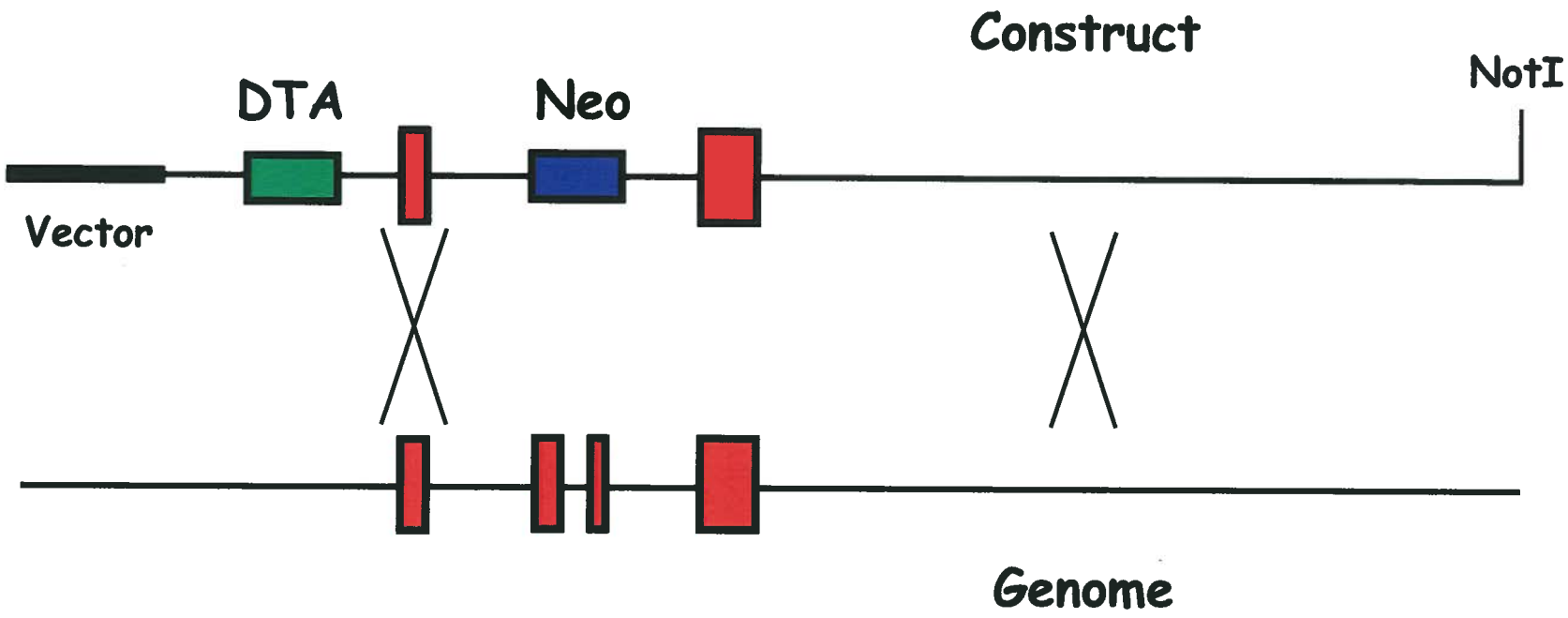
# Applications: 4 knock-out constructs

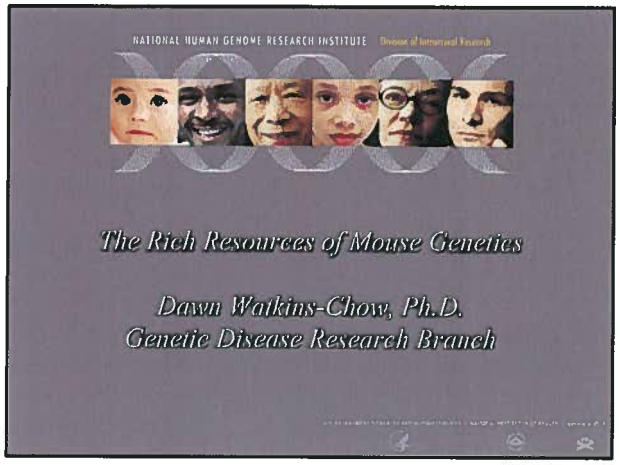


# Applications: 4 knock-out constructs



# Applications: 4 knock-out constructs





## Outline

- History of Mouse Genetics
- The Mouse Genome
- Reverse Genetics: Engineered Mice
- Forward Genetics: Screens for Induced Mutations
- Strategic Breeding for Diversity
- Public Database Resources

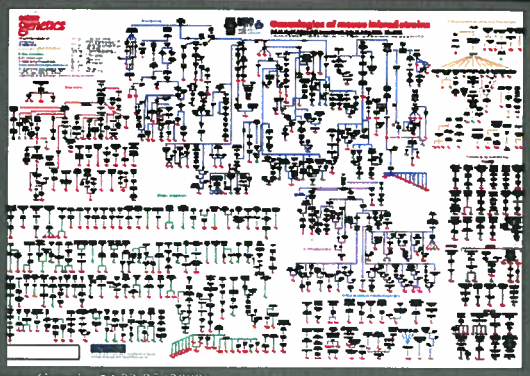
## History of the Laboratory Mouse

- 1100 BC "Fancy" mouse breeding of coat color variants in Asia & Europe
- 1900 Retired schoolteacher Abbie Lathrop breeds fancy mice in Granby, MA
- 1909 first inbred strains established
- 1929 The Jackson Laboratory founded
- 1962 nude mouse
- 1980 first transgenic mouse
- 1989 first knockout mouse
- 1990 explosion of engineered mice & beginning of mouse genome project

## Mouse Stocks

- Outbred Stocks
  - Genetically undefined, bred to maintain heterozygosity
  - Cheap & healthy for when genotype does not matter
  - Should not be used for breeding experiments
- Inbred Strains
  - Bred for homozygosity (~100 strains available)
- Several commercial vendors available
  - JAX (jaxmice.jax.org)
  - Taconic (www.taconic.com)
  - Charles River (www.criver.com)

## Inbred Strain Geneologies

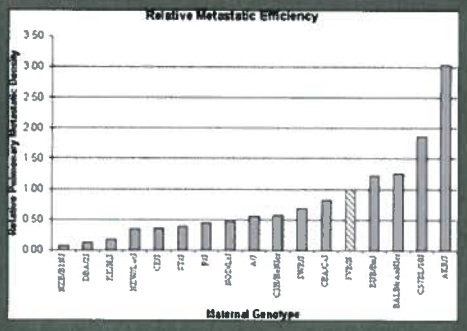


*Nature Genetics* 24:23-25 (2000)

## Mouse Phenome Database

http://phenome.jax.org/pub-cgi/phenome/mpdcgi

## Affect of Background Strain




T. Lilsted et al. *Int. J. Cancer* (1998) 77:640-644
K. Hunter. *Cime Lamer* (2003) 200:97-105

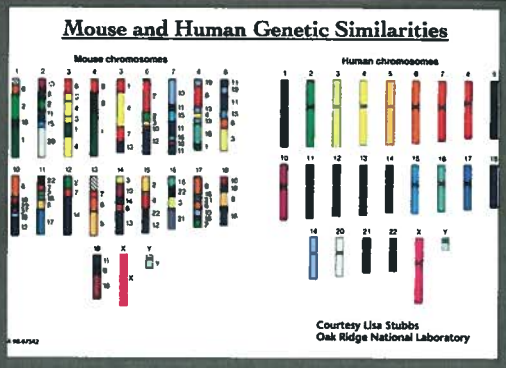
## *Live Linkage Map of the Mouse*

10th International Congress of Genetics, Montreal
Created by Staff of The Jackson Laboratory, 1958

## *The Mouse Genome*

*Nature* 420:520-562 (2002)

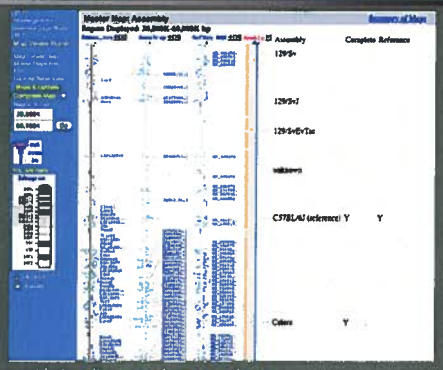
## *Synteny*




## *The Mouse Genome*

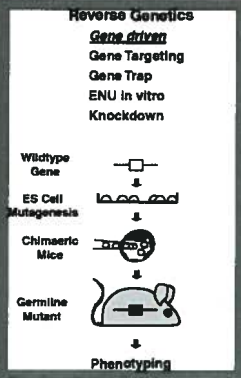
- 19 autosomes, XY
- NCBI Build 37
  - $2.66 \times 10^9$ bp
    - 124 contigs mostly > 5000 kb
  - 28,278 transcripts
  - C57BL/6J is reference strain
- Web browsers
  - http://genome.ucsc.edu
  - http://www.ensembl.org
  - http://www.ncbi.nlm.nih.gov
    - includes alternate partial assemblies

## The Mouse Genome




http://www.ncbi.nlm.nih.gov/projects/mapview/

## Strategies to identify gene function




From Gene Knockout Protocols

## Large Scale Efforts

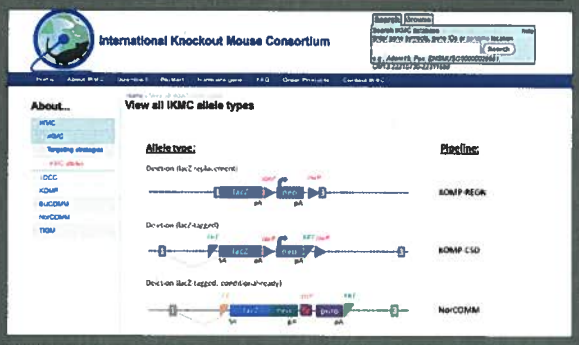
**International Knockout Mouse Consortium**

IKMC Gene Progress Summary

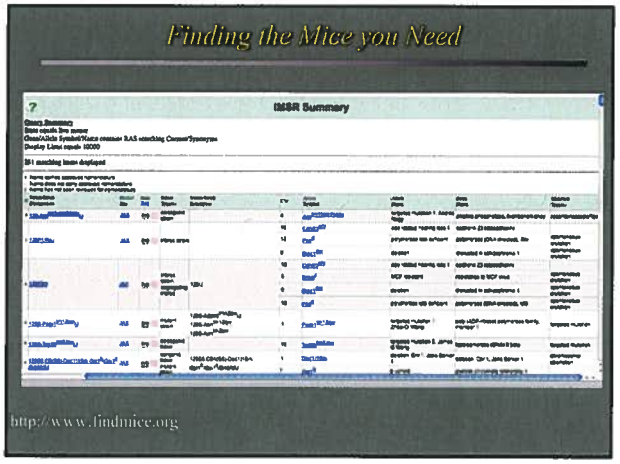
| Total Genes | IKOMP | | EUCOMM | NorCOMM | TIGM |
|---|---|---|---|---|---|
| | CSD | Regeneron | | | |
| Project goal | 5000 | 3500 | 6000 | 500 | - |
| Vectors generated | 5099 | 3295 | 4477 | 312 | - |
| Vectors available | 4791 | 2296 | 4477 | 312 | - |
| ES cells generated | 2550 | 1808 | 2386 | 45 | - |
| ES cells available | 1940 | 1136 | 2386 | 45 | 10699 |
| Mutant mice generated | 168 | 160 | 279 | 0 | - |
| Mutant mice available | 168 | 71 | 279 | 0 | - |

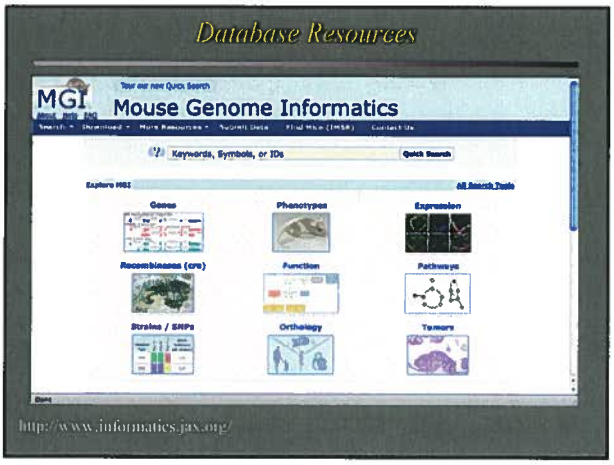
http://www.knockoutmouse.org

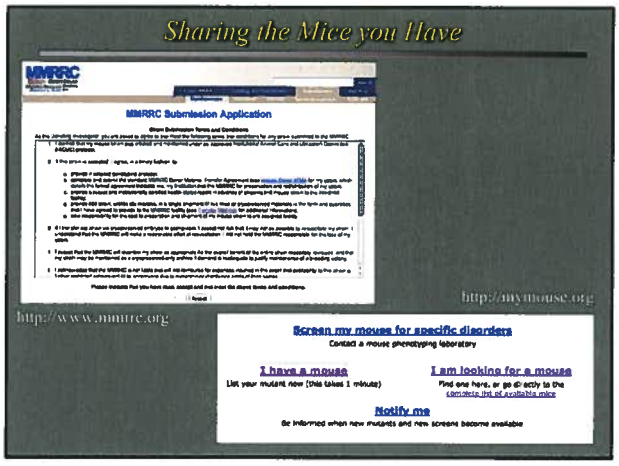
## Large Scale Efforts

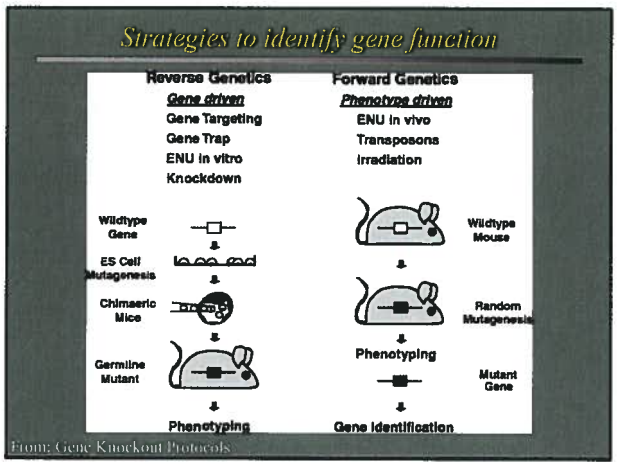



http://www.knockoutmouse.org


Finding the Mice you Need
IMSR Summary
http://www.findmice.org


Database Resources
MGI
Mouse Genome Informatics
Keywords, Symbols, or IDs
Quick Search
Genes
Phenotypes
Expression
Recombinases (cre)
Function
Pathways
Strains / SNPs
Orthology
Tumors
http://www.informatics.jax.org/

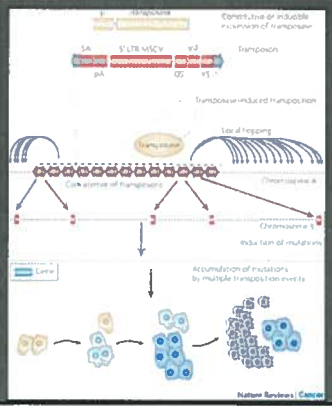

Sharing the Mice you Have
MMRRC
MMRRC Submission Application
http://www.mmrrc.org
http://mymouse.org
Screen my mouse for specific disorders
Contact a mouse phenotyping laboratory
I have a mouse
List your mutant now (This takes 1 minute)
I am looking for a mouse
Find one here, or go directly to the
complete list of available mice
Notify me
Be informed when new mutants and new screens become available


Strategies to identify gene function
Reverse Genetics
Gene driven
Gene Targeting
Gene Trap
ENU in vitro
Knockdown
Forward Genetics
Phenotype driven
ENU in vivo
Transposons
Irradiation
Wildtype Gene
ES Cell Mutagenesis
Chimaeric Mice
Germline Mutant
Phenotyping
Wildtype Mouse
Random Mutagenesis
Phenotyping
Mutant Gene
Gene Identification
From: Gene Knockout Protocols

## *Forward Genetics*

- *Retroviral Insertional Mutagenesis*
  - *MMTV, MuLV*
- *Transposon Mediated Mutagenesis*
- *ENU Mutagenesis*

## *Transposon Mediated Mutagenesis*



Kool & Berns
*Nature Reviews Cancer*
*(2009)* 9:389-399.

## *Transposon Mediated Mutagenesis*

- Sleeping beauty

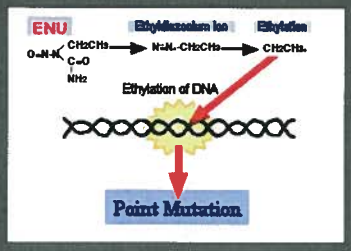
  Keng et al. (2005) *Nature Methods* **2**: 763-769

- piggyBac

  Dinget al. (2005) *Cell* **122**: 473-483

## *ENU Mutagenesis*

- ENU mutagenesis (ethylnitrosourea)




- Alkylating agent
- Mutations introduced into spermatogonial stem cells

## ENU Mutagenesis Large Scale Screens

Riken
- broad phenotyping

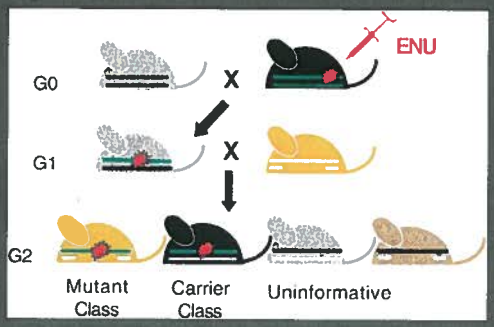
MRC Harwell
- broad phenotyping

The Jackson Laboratory
- Neurological, reproductive, heart, lung, blood

Baylor College of Medicine
- Chromosome 11

## ENU Mutagenesis

Region specific screens




Justice Hum Mol Gen (1999) 8:1955-63

## ENU Mutagenesis

Sensitized/Modifier screens

Matera et al. Hum Mol Gen (2008) 17:2118-2131

## Mapping Phenotypes

SNP or SSLP markers polymorphic between the parental mouse strains (C57BL/6J and BALB/cJ) are used to genotype affected mice throughout the genome

| SNP Position | BALB/cJ C57BL/6J affected mice |
|---|---|
| 13.001.770 | |
| 13.003.265 | |
| 13.008.983 | |
| 13.015.500 | |
| 13.019.050 | |
| 13.028.450 | |
| 13.034.725 | |
| 13.042.480 | |
| 13.051.415 | |
| 13.063.230 | |
| 13.066.460 | |
| 13.069.159 | |
| 13.087.830 | |
| 13.094.010 | |
| 13.102.770 | |
| 13.114.540 | |

Mutation

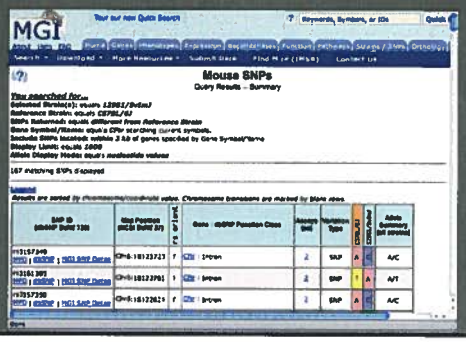
- Homozygous BALB/cJ
- Homozygous C57BL/6J
- Heterozygote
- No call

## Polymorphic Variation

- RFLP
- Simple sequence length polymorphism (SSLP) SSRP, STRP, Dinucleotide repeat, microsatellites
- Copy number variation (CNVs)
- Single Nucleotide Polymorphisms
  - Over 10,000,000 SNPs in NCBI dbSNP
  - From 86 strains

## Searching for Polymorphic Variation

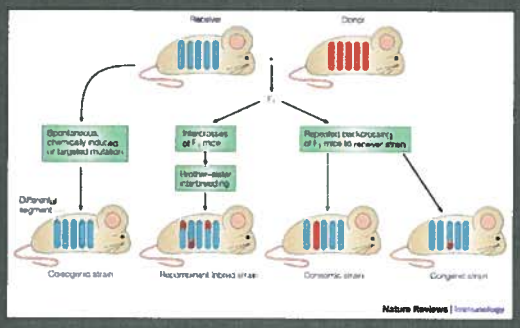
- http://phenome.jax.org & http://www.informatics.jax.org
- 167 SNPs polymorphic between B6 and 129 within 2 kb of C4b



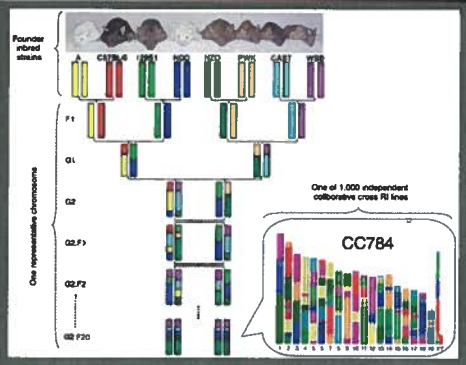

## Standard Crosses

- Incross: mating between individuals of an inbred strain. Used for inbred strain maintenance.
- Outcross: mating between 2 unrelated strains. To generate F1s.
- Backcross: mating between a heterozygous F1 and one of the parent strains. Used in linkage analysis.
- Intercross: mating between two identically heterozygous individuals. Used in linkage analysis.

## Designer Crosses



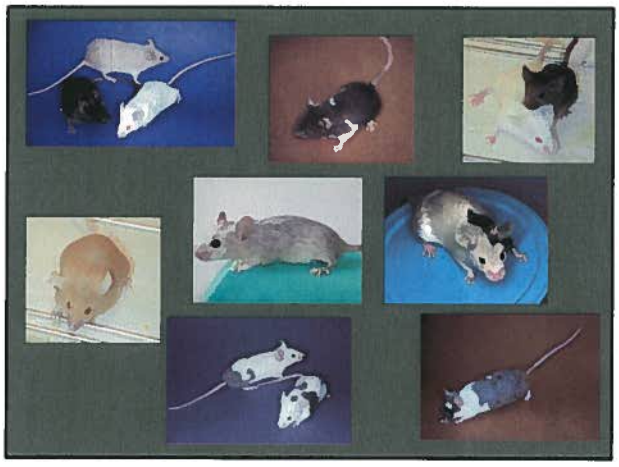
## Collaborative Cross




http://compgen.unc.edu/wp/?page_id=99

## Reference Material

- *Mouse Genetics, Concepts and Applications* (1995) Silver, L.M. Oxford University Press, New York, NY
- *Genetic Variants and Strains of the Laboratory Mouse*, 3rd Edition (1996) Edited by M. F. Lyon, S. Rastan and S.D.M. Brown. Oxford University Press,New York, NY
- Online Books:

  http://www.informatics.jax.org/resources.shtml#res_books
- Online Resources:

  http://www.informatics.jax.org/resources.shtml