



The Zebrafish Mutation Project

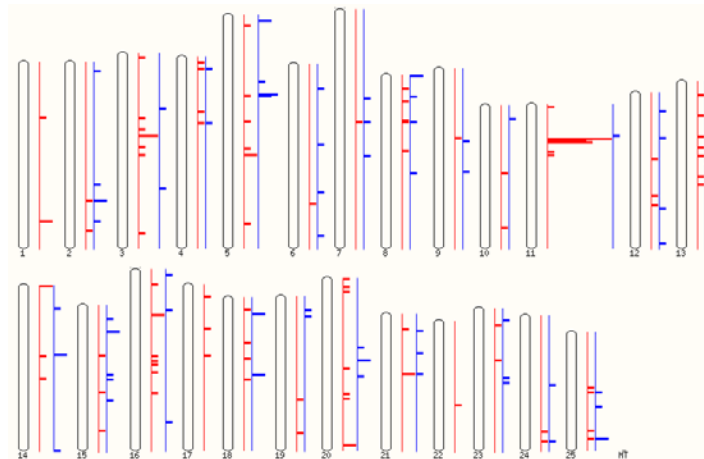
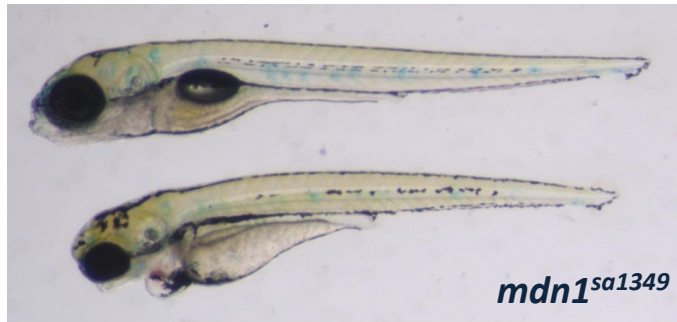
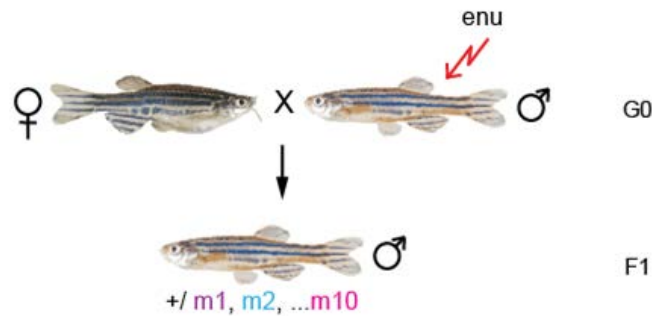
Derek L Stemple

Wellcome Trust Sanger Institute

Linking Disease Model Phenotypes to Human Conditions

10 September 2015

Functional annotation of a vertebrate genome

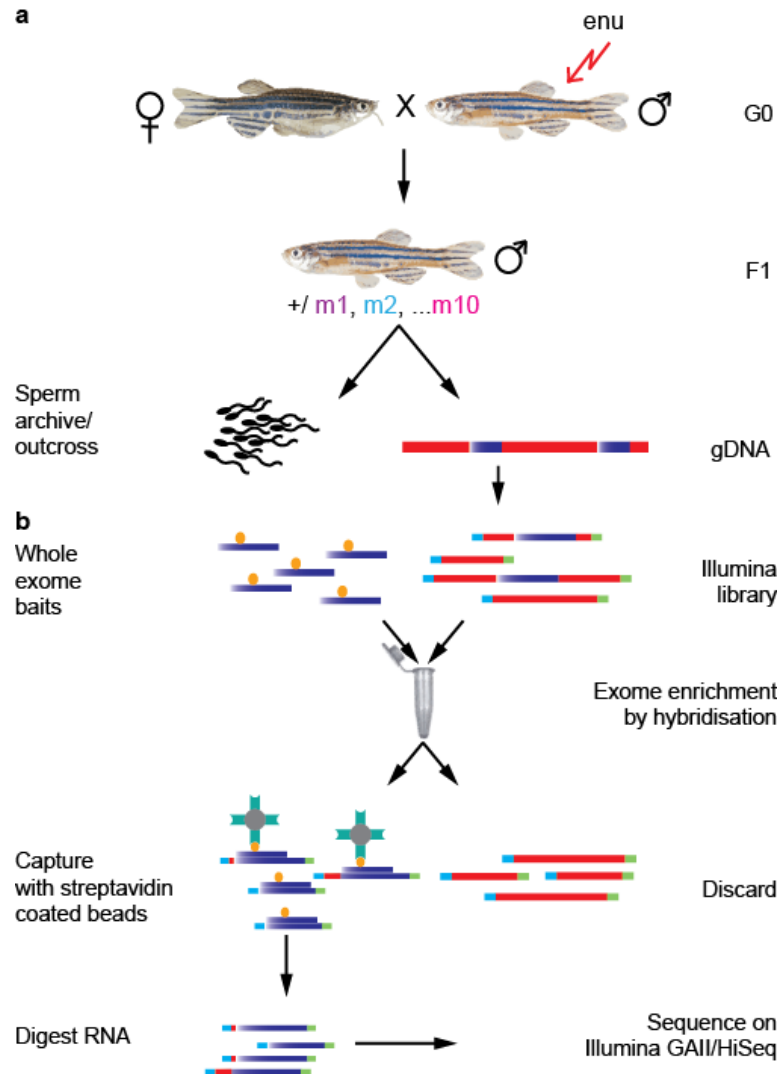


Identify a disruptive mutations in every protein-coding gene in zebrafish

Provide morphological phenotype description of 8,000 genes

mRNA expression profiles of alleles producing abnormal phenotypes

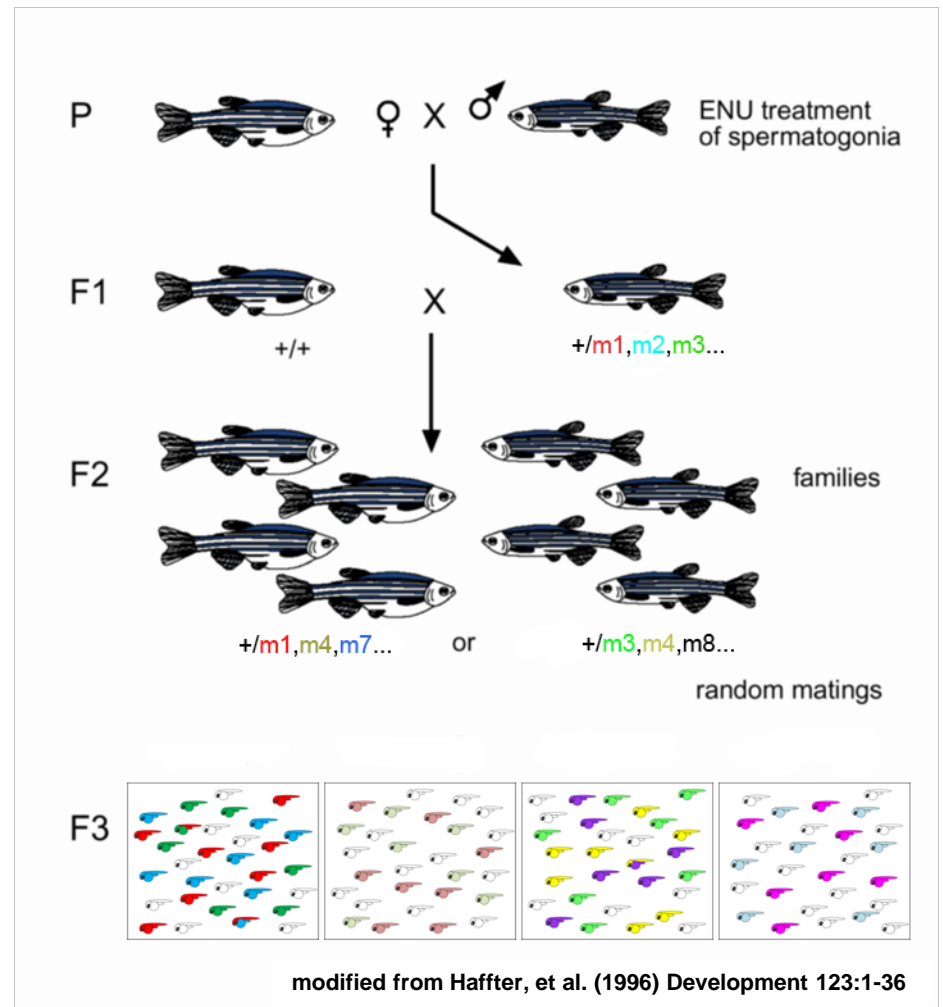
Chemical Mutagenesis and Mutation Detection



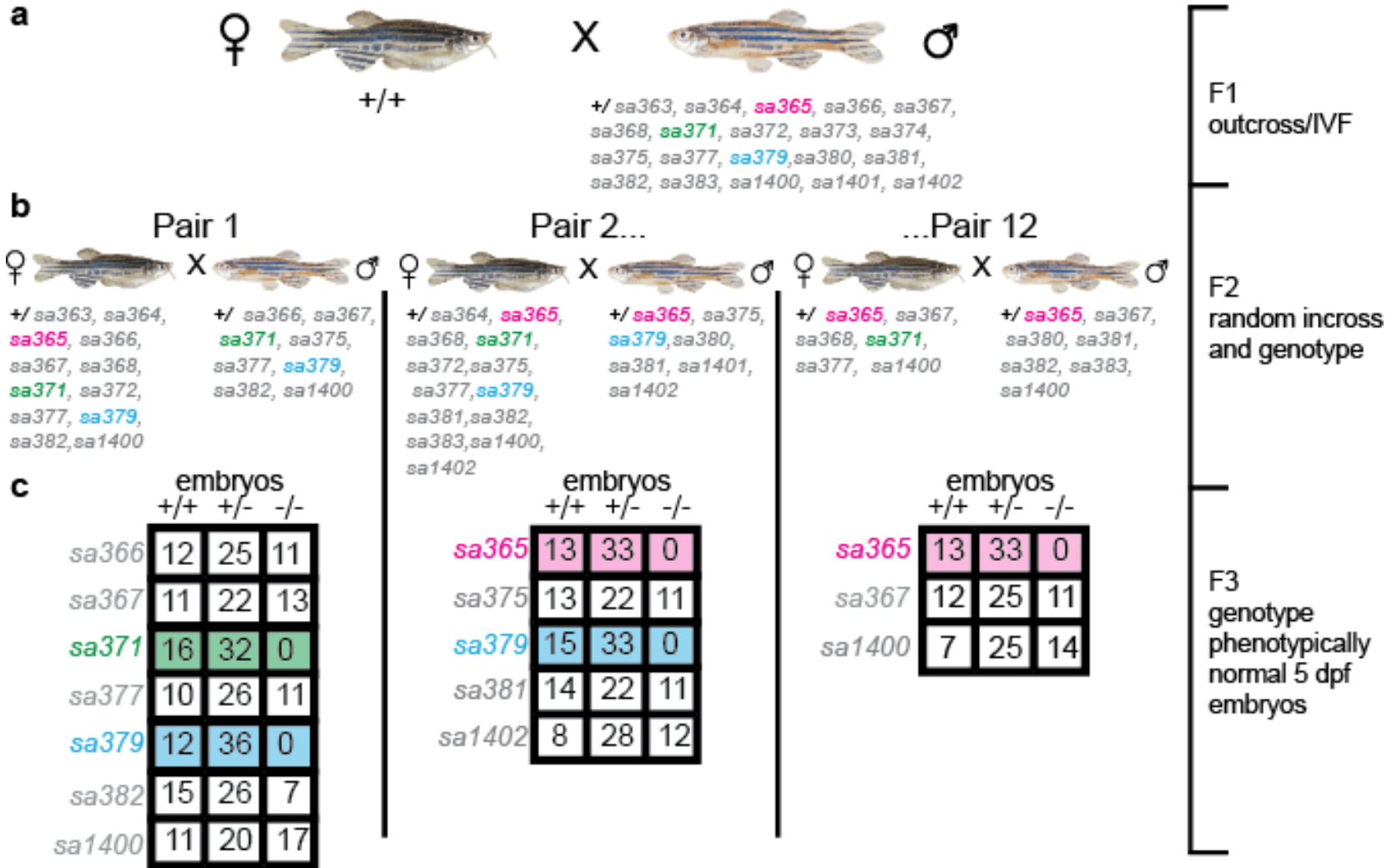
Kettleborough RN et al., Nature, 2013

Phenotyping: reverse genetics

- Gene driven
- Identify mutations carried by families
- Associate phenotype after mutation discovery



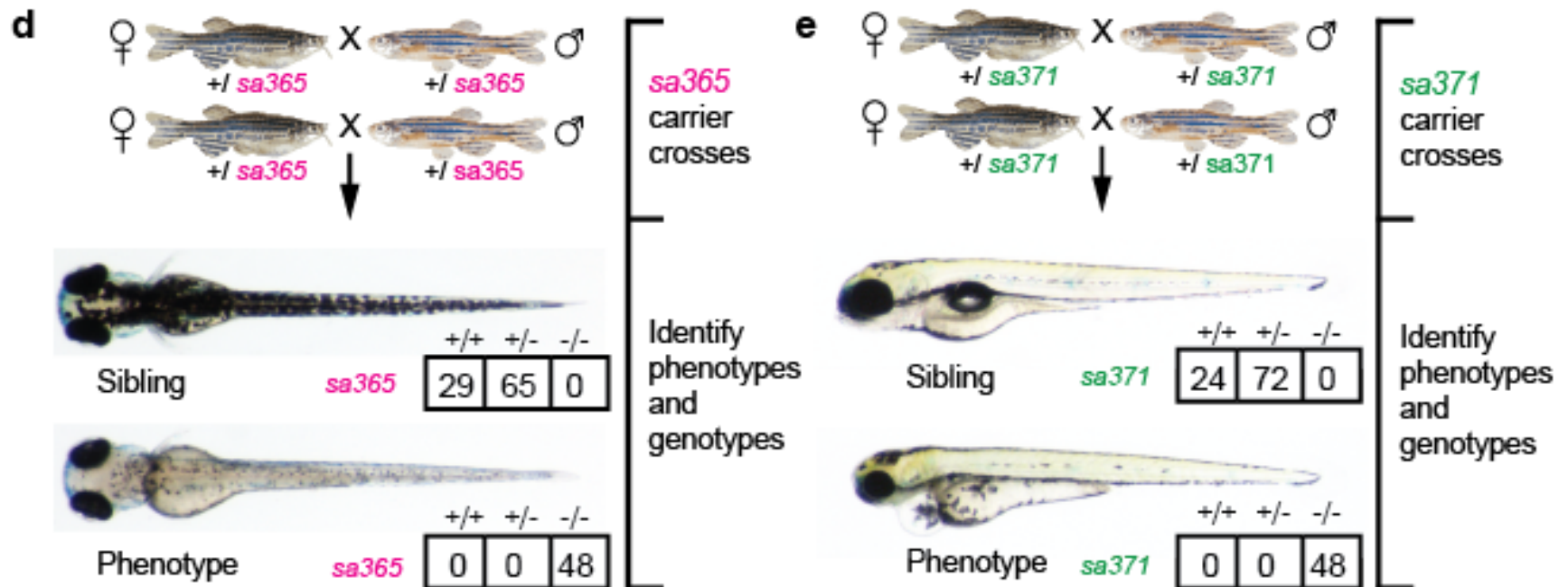
Morphological phenotyping



Kettleborough RN et al., Nature, 2013

Dooley CM et al., Methods, 2013

Morphological phenotyping



Kettleborough RN et al., Nature, 2013

Dooley CM et al., Methods, 2013

Zebrafish Mutation Project: *lamc1*

http://www.sanger.ac.uk/sanger/Zebrafish_Zmpgene/ENSDARG00000036279

www.sanger.ac.uk/sanger/Zebrafish_Zmpgene/ENSDARG00000036279


OS X Reference Google Maps YouTube Wikipedia iGoogle eSNAP Menu AGRESSO Business W Home page - Wellcome Form C Editor Zebrafish Mutation P

wellcome trust
sanger
institute

Search for Enter search here...

Home Research Scientific resources Work & study About us

Mouse Zebrafish Data Software Databases Technologies Talks & training



Search Zebrafish Mutation Project

Search for:

e.g. [ENSDARG00000093395](#) or [slc2a11b](#) or [ENSG00000135862](#) or [LAMC1](#) or [sa457](#)

You can look for mutant lines by [browsing a complete list](#) or by searching for a particular gene

lamc1

Ensembl ID: [ENSDARG00000036279](#)

ZFIN ID: [ZDB-GENE-021226-3](#)

Description: Laminin subunit gamma-1 [Source:UniProtKB/Swiss-Prot;Acc:Q1LVF0]

Human Orthologue: LAMC1

Human Description: laminin, gamma 1 (formerly LAMB2) [Source:HGNC Symbol;Acc:6492]

Mouse Orthologue: Lamc1

Mouse Description: laminin, gamma 1 Gene [Source:MGI Symbol;Acc:MGI:99914]

Zebrafish Mutation Project: *lamc1*

http://www.sanger.ac.uk/sanger/Zebrafish_Zmpgene/ENSDARG00000036279

Mutation Details

Allele Name: sa379
Current Status: **Available for shipment**
 For more information about the meaning of this status and other statuses, please see our [FAQs](#).
Availability: [Order Allele From ZIRC](#)
Mutation: G > A
Consequence: Essential Splice Site

Transcript ID	Consequence	Amino Acid Affected	Amino Acid Total	Exon Affected	Exon Total
ENSDART0000004227	Essential Splice Site	325	1593	5	28

Genomic Location: Chromosome 2 (position 35675287)
KASP Assay ID: 554-0359.1 (used for ordering genotyping assays from [LGC Genomics](#))
KASP Sequence: AATATTTTTATAGGCACATTTTCTGACTGGTGTGGTGTGTGTTTCAG/ACTTGTAAGTGAATGGAAAGAGCGCTGAATGTTACTTCGACCCGTGAGCTG
Associated Phenotype: This allele has been associated with this phenotype by genetic linkage analysis and may not be causal. See [FAQs](#) for more info.



Stage	Entity	Quality	Tag
Pharyngula:Prim-5 ZFS:0000029	brain ZFA:0000008	quality PATO:0000001	abnormal PATO:0000460
Pharyngula:Prim-5 ZFS:0000029	extension ZFA:0000106	quality PATO:0000001	abnormal PATO:0000460
Pharyngula:Prim-5 ZFS:0000029	eye ZFA:0000107	quality PATO:0000001	abnormal PATO:0000460
Pharyngula:Prim-5 ZFS:0000029	notochord development GO:0030903	disrupted PATO:0001507	abnormal PATO:0000460
Pharyngula:Prim-5 ZFS:0000029	somite ZFA:0000155	quality PATO:0000001	abnormal PATO:0000460

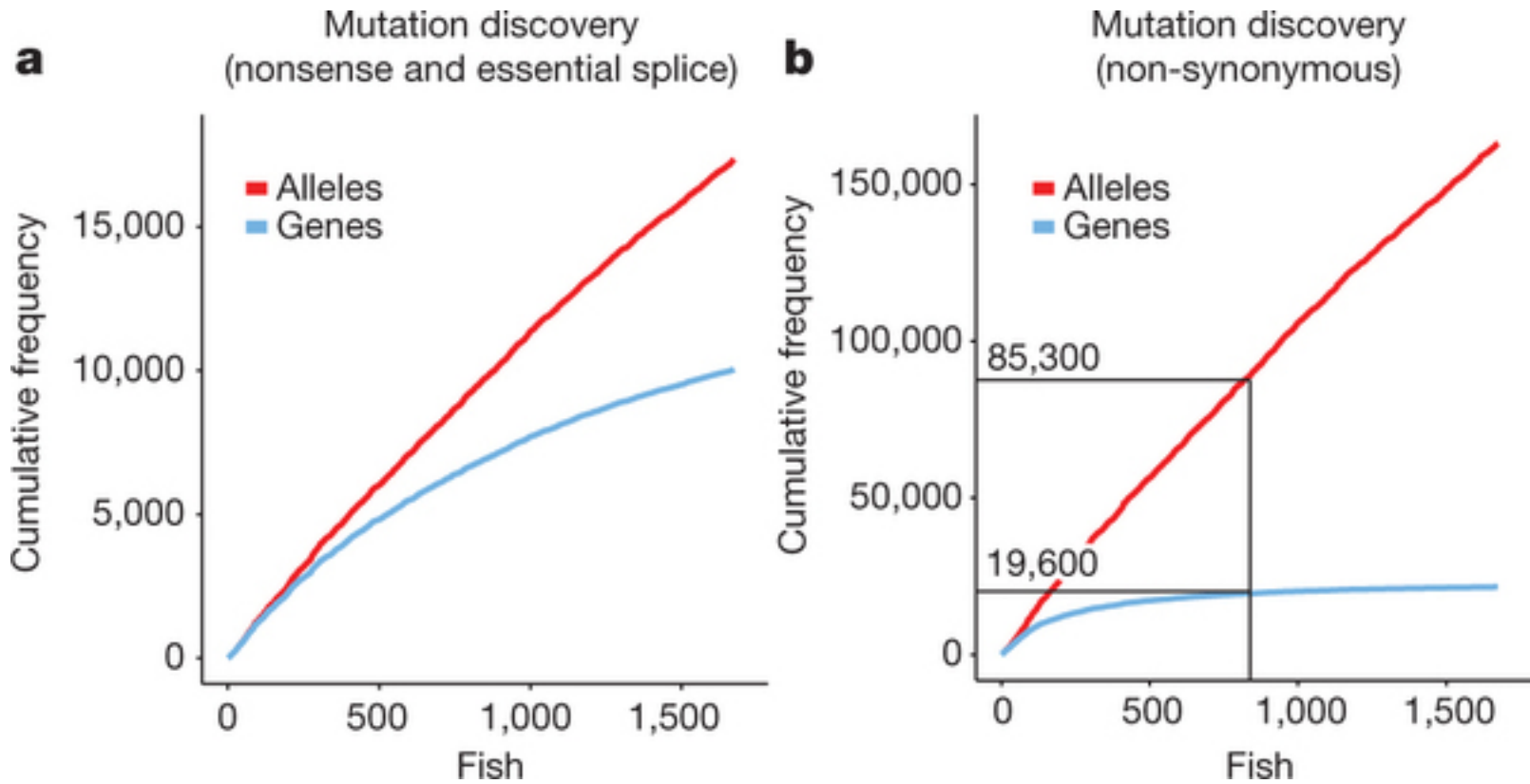
Transcriptome Profiling Preview: [View complete transcriptome profile](#)

Region	3' end position	3' end strand	Adjusted p-value	Log2 fold change (mutant/sibling)	Closest Ensembl gene 3' end	Gene name	e70 Ensembl Gene ID
1:4393283-4394800	4393283	-1	< 1 × 10 ⁻¹⁶	3.3	0	fn1b	ENSDARG00000006526
12:37046025-37046500	37046025	-1	< 1 × 10 ⁻¹⁶	3.2	-2	socs3b	ENSDARG00000026611
13:1656801-1657243	1657243	1	< 1 × 10 ⁻¹⁶	-2.0	0	CABZ01044297.1	ENSDARG00000074745
13:22812801-22813207	22813207	1	< 1 × 10 ⁻¹⁶	-2.0	-17	sncga	ENSDARG00000034423
13:39679457-39680100	39679457	-1	< 1 × 10 ⁻¹⁶	-1.7	12	col9a1	ENSDARG00000031483

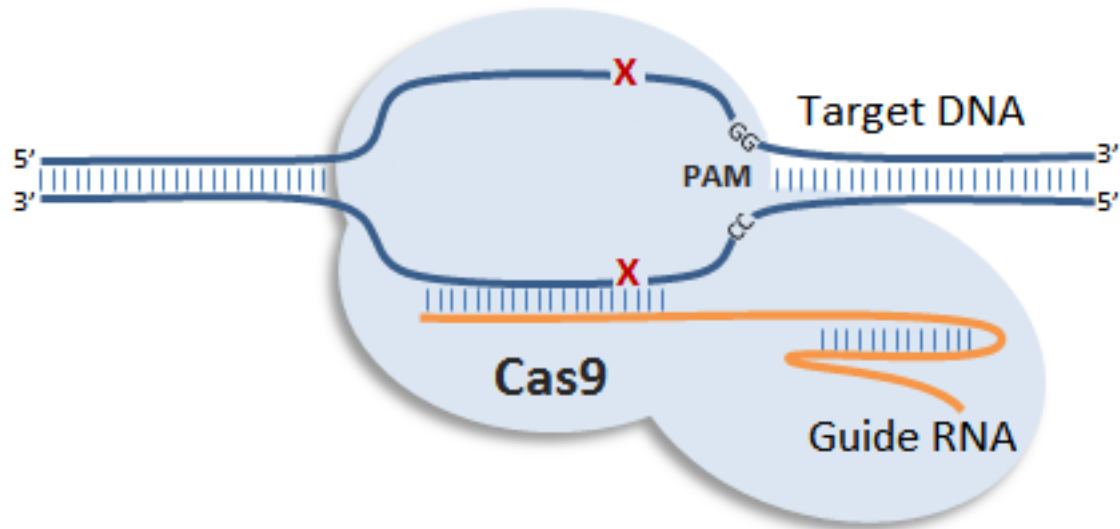
Mutation Detection

- All detected mutations are available for perusal at the European Nucleic Acid Archive (ENA)
<http://www.ebi.ac.uk/ena/data/view/PRJEB1830>
- So far we have sequenced 2,839 F1 males
- We have identified 32,383 nonsense and disruptive splice-site alleles in 14,059 genes (53% of total)
- BUT, we are reaching a point of diminishing returns

Saturation Mutation Detection

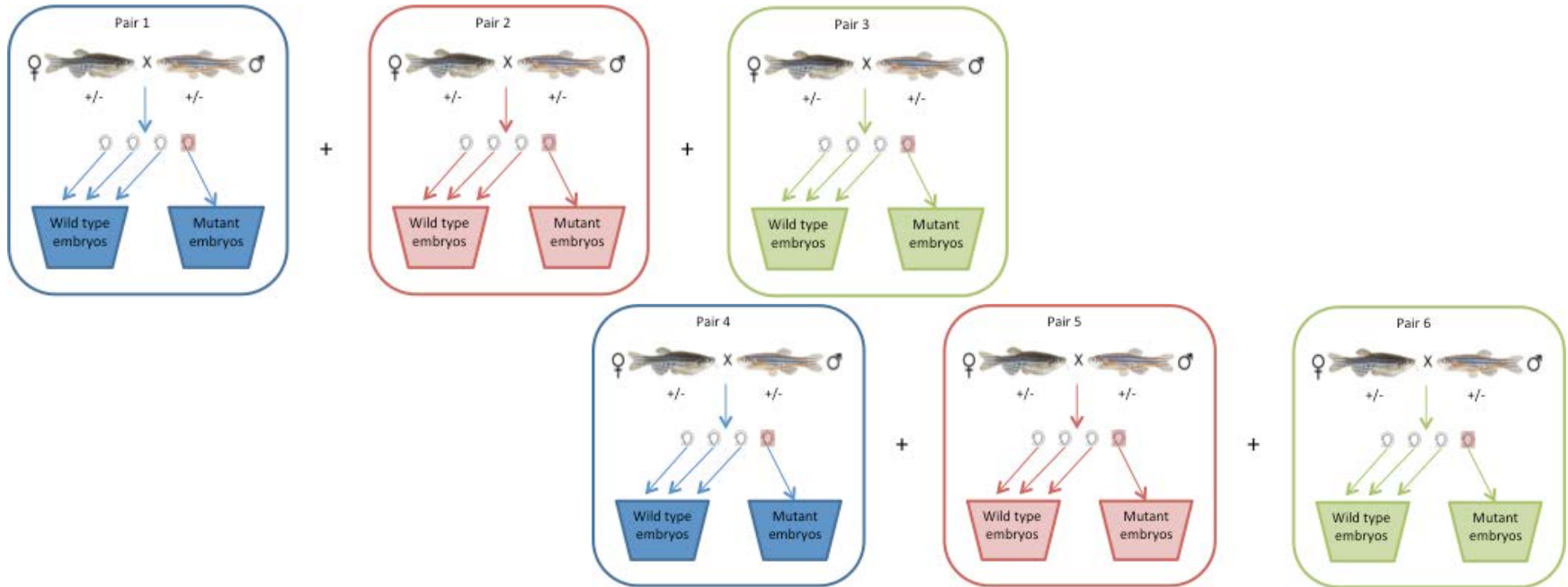


CRISPR/Cas9



- CRISPRs (Clustered Regularly Interspaced Short Palindromic Repeats)
- Bacterial immunity
- Cas9 Double-stranded endonuclease
- Hybrid 'guide RNA' with Cas9 binding sequence and targeting sequence (20bp)
- With codon optimisation and NLS very efficient activity in eukaryotic cells

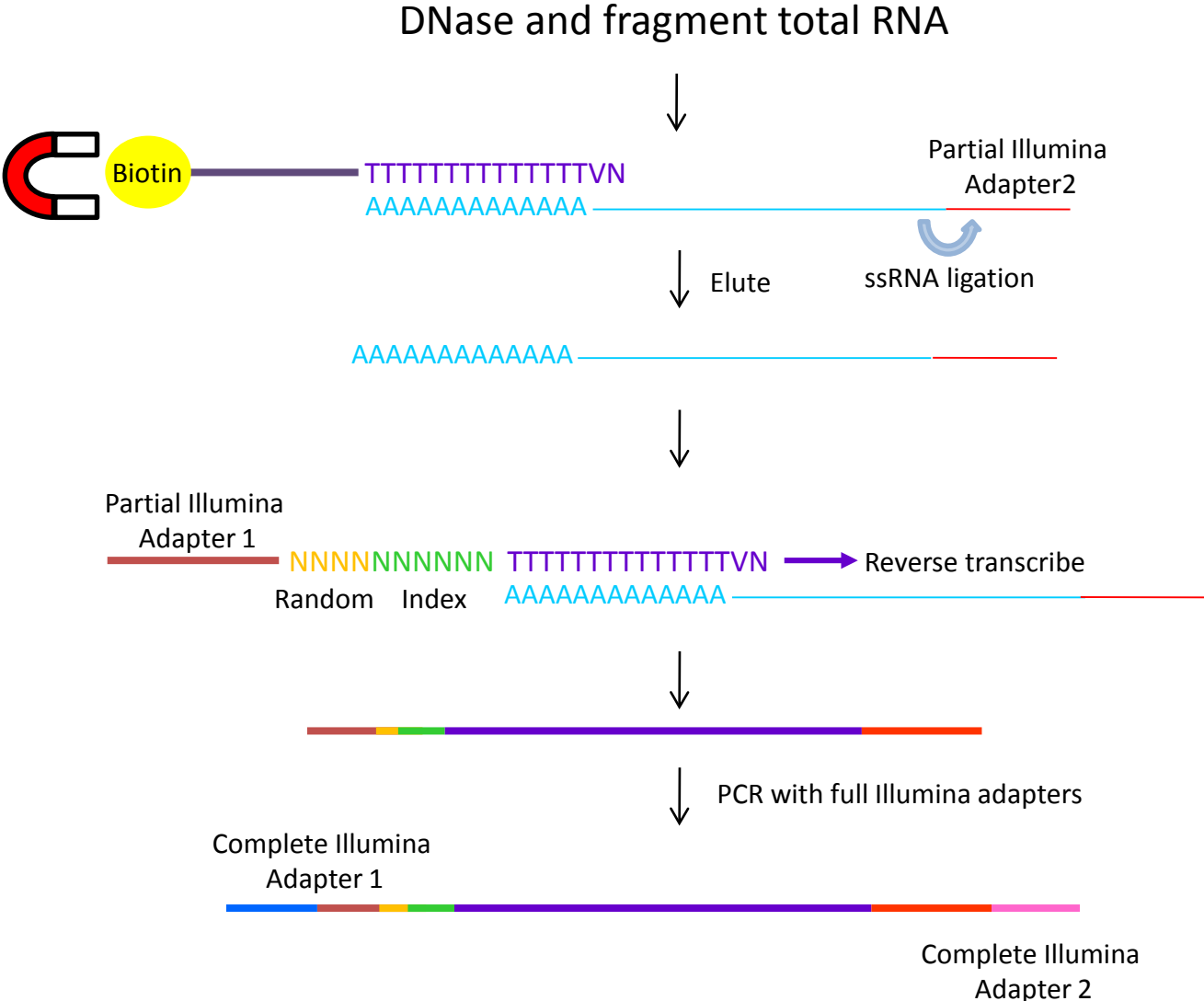
DeTCT (Differential expression transcript counting technique)



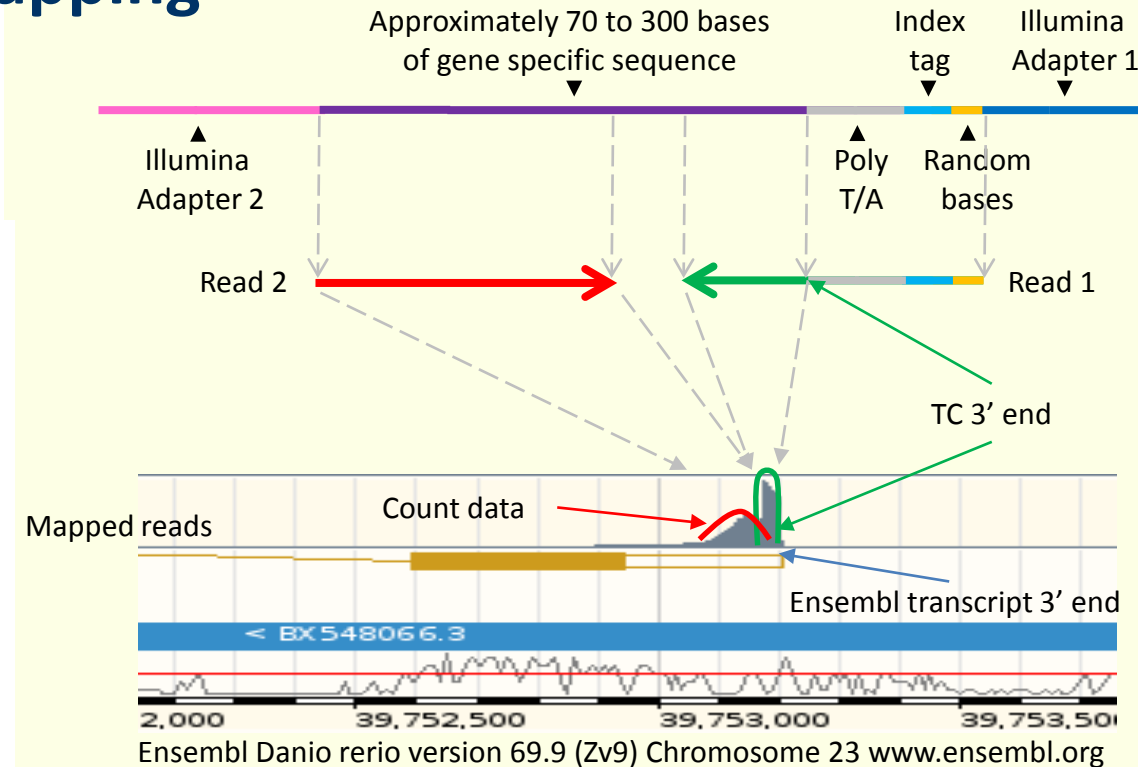
Quantitative analysis of changes in transcript levels of mutant vs. wild-type siblings

- Restricted to 3' ends – more sequencing space than RNASeq
- Identifies differentially abundant polyA transcripts
- Identifies functional non-coding RNAs
- Highly sensitive even for rare cell types (e.g. pigment cells)

Molecular Phenotyping: DeTCT

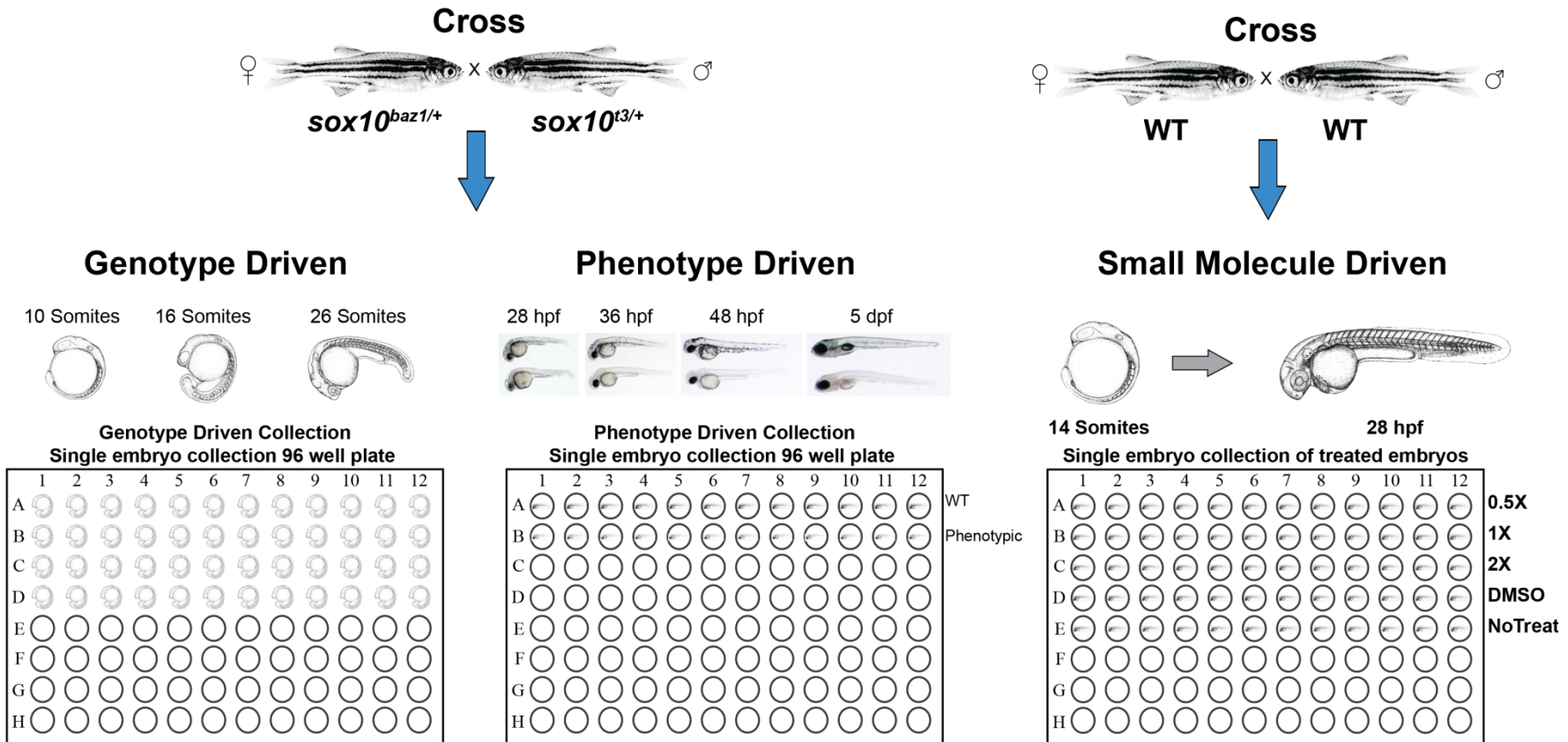


Read Pair Mapping



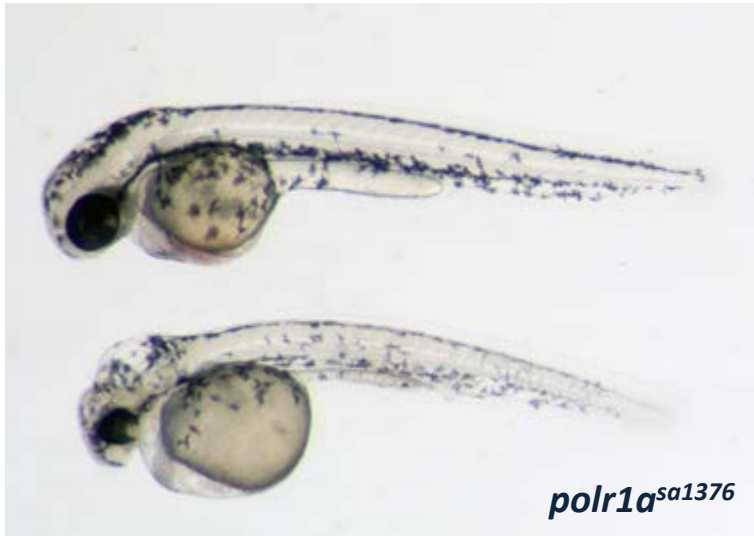
1. Illumina libraries comprising 3' ends of each transcript are generated
2. Each sample is indexed with a different tag
3. Reads are mapped to the reference genome
4. Genomic regions are detected as a peak of mapped read pairs above background
5. Differentially expressed regions are identified by DEseq2
Anders S, Huber W. (2010) Differential expression analysis for sequence count data, *Genome Biology* **11**:R106
6. Regions are associated with annotated genes

DeTCT input



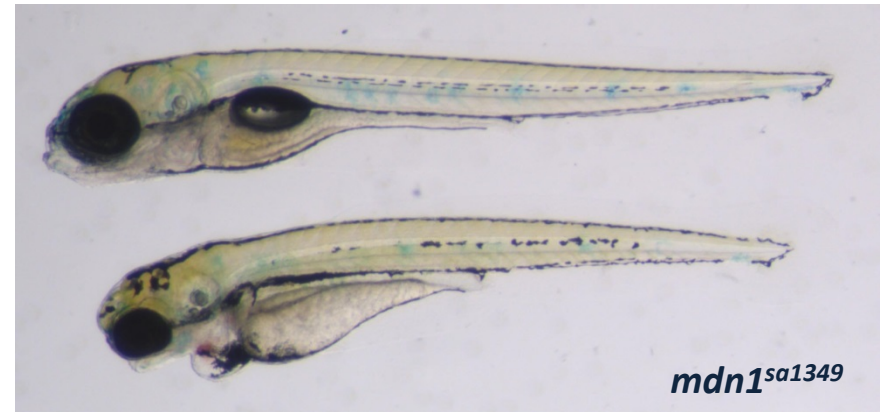
Many biological replicates: 12 embryos per condition
 Output: Table of differentially abundant transcript ends
 ~ 100 experiments analysed

Ribosome mutants



polymerase (RNA) I polypeptide A

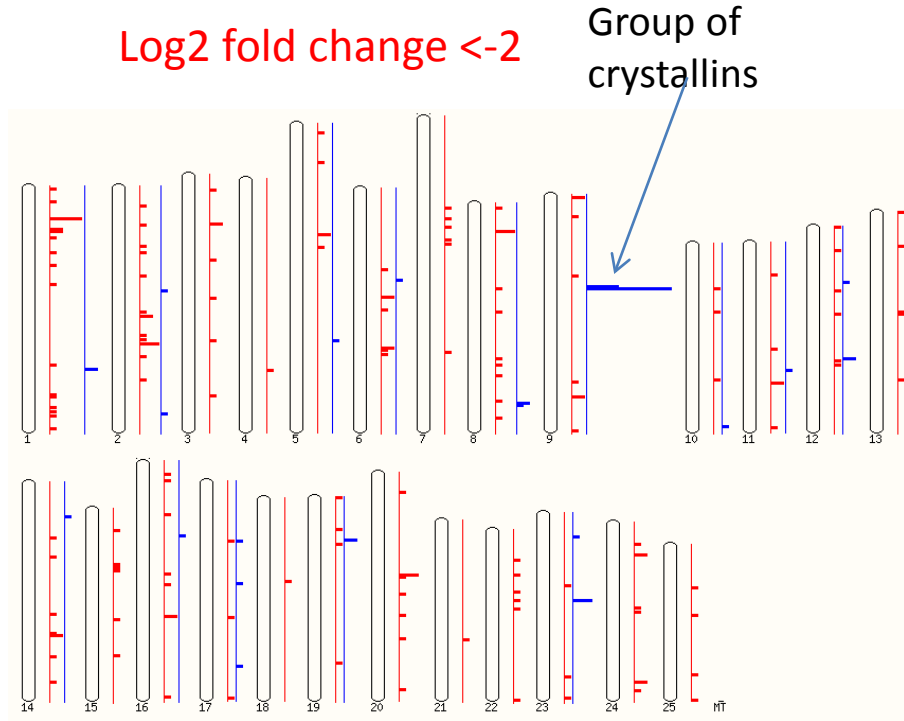
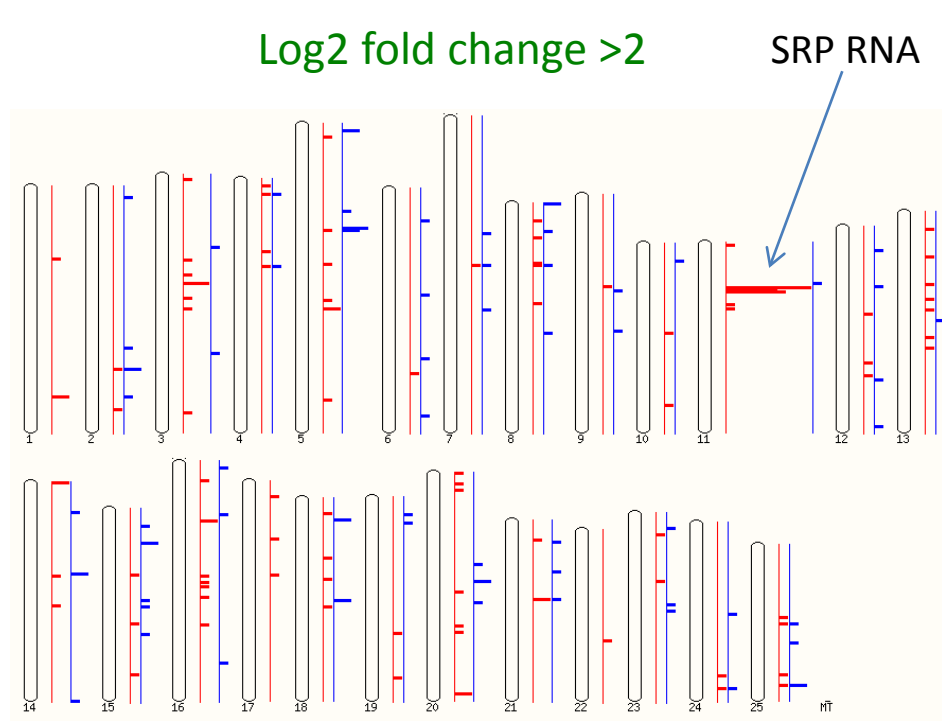
- Largest subunit of RNA polymerase I
- Transcribes large ribosomal RNA, cleaved into three rRNAs
- Treacher Collins
- Phenotype first apparent early day 2



midasin homolog (yeast)

- Nuclear chaperone required for maturation and nuclear export of pre-60S ribosome subunits
- Mouse KO early embryonic lethal
- Phenotype first apparent early day 5

Ribosome mutants: Global comparison



Blue: *polr1a*^{sa1376} 243 significant regions
Red: *mdn1*^{sa1349} 960 significant regions

Within 1000 bases
P-val <0.05

ZFIN: Comprehensive gene expression patterns

polr1a <http://zfin.org/ZDB-GENE-030131-5286>



Home Genes / Markers / Clones BLAST GBrowse Expression Antibodies Mutants / Knockdowns / Tg Constructs Anatomy / GO Publications Maps

Research

General Information

ZIRC

Site Search:

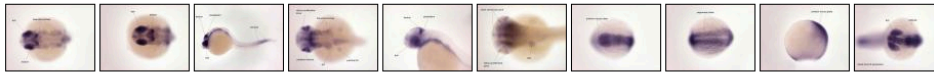


Modify Search

Your Input Welcome

Figure Gallery (12 images)

◀ 1 / 2 ▶



Search matched on current symbol: **polr1a**

Expression Pattern Search Results for *polr1a*

(7 figures with expression from 2 publications)

[[Show only figures with images](#)]

Publication (current status)	Data	Genotype or Background	Stage Range	Anatomy
Stuckenholtz et al., 2009	Fig. 3	<i>s854Tg/s854Tg</i> <input type="checkbox"/>	Long-pec to Day 6	gut <input type="checkbox"/>
Thisse et al., 2001 - Present [cb935]	Fig. 1	AB/Tuebingen	50%-epiboly to Bud	neural plate <input type="checkbox"/>
	Fig. 2	AB/Tuebingen	1-4 somites to 10-13 somites	neural plate <input type="checkbox"/> , segmental plate <input type="checkbox"/>
	Fig. 3	AB/Tuebingen	14-19 somites	immature eye <input type="checkbox"/> , midbrain <input type="checkbox"/> , somite <input type="checkbox"/>
	Fig. 4	AB/Tuebingen	20-25 somites to Prim-5	alar plate midbrain region <input type="checkbox"/> , cerebellum <input type="checkbox"/> , immature eye <input type="checkbox"/> , optic tectum <input type="checkbox"/> , pharyngeal arch 3-7 skeleton <input type="checkbox"/> , somite <input type="checkbox"/>
	Fig. 5	AB/Tuebingen	Prim-15 to Prim-25	cerebellum <input type="checkbox"/> , eye <input type="checkbox"/> , gut <input type="checkbox"/> , optic tectum <input type="checkbox"/> ... (all 7) ▶
	Fig. 6	AB/Tuebingen	High-pec to Long-pec	liver <input type="checkbox"/> , retina <input type="checkbox"/> , ventricular zone <input type="checkbox"/>

polr1a^{sa1376}: anatomical analysis

Enrichment analysis of ZFA terms

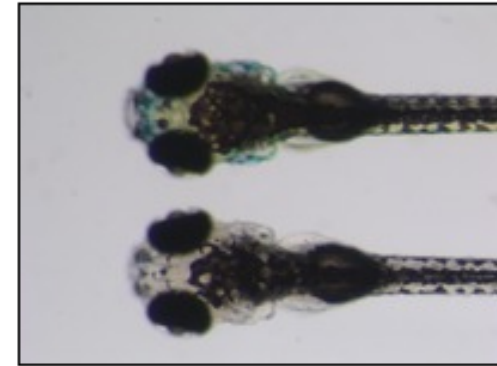
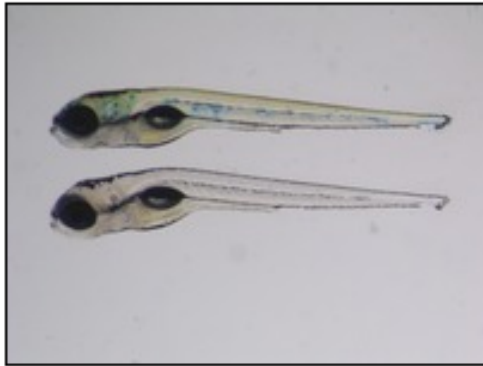
ZFA:0001477	portion of tissue	4.03E-18
ZFA:0001105	embryonic structure	6.12E-18
ZFA:0001512	anatomical group	1.03E-17
ZFA:0001439	anatomical system	2.24E-17
ZFA:0001308	organism subdivision	3.02E-14
ZFA:0000396	nervous system	3.85E-13
ZFA:0001490	cavitated compound organ	4.82E-13
ZFA:0001488	multi-tissue structure	5.15E-13
ZFA:0000496	compound organ	5.88E-12
ZFA:0000282	sensory system	7.81E-11
ZFA:0001127	visual system	1.36E-10
ZFA:0000107	eye	4.38E-10
ZFA:0001478	anatomical cluster	1.06E-09
ZFA:0009000	cell	2.74E-09
ZFA:0001114	head	7.62E-09
ZFA:0000093	blastomere	2.30E-08
ZFA:0001175	blastodisc	3.59E-08
ZFA:0000035	lens	4.16E-08
ZFA:0005566	anterior segment eye	7.89E-08
ZFA:0000445	optic tectum	1.57E-07

Visual system

Nervous system

ZFA:0001328	neuromere	8.35E-06
ZFA:0000029	hindbrain	1.06E-05
ZFA:0001094	whole organism	1.42E-05
ZFA:0000037	anatomical structure	2.08E-05
ZFA:0000079	telencephalon	7.24E-05
ZFA:0000128	midbrain	8.75E-05
ZFA:0000101	diencephalon	5.28E-04
ZFA:0000109	forebrain	0.001206
ZFA:0000633	caudal tuberculum	0.001387
ZFA:0000516	periventricular grey zone	0.002074
ZFA:0000100	cerebellum	0.002128
ZFA:0000075	spinal cord	0.00255
ZFA:0000008	brain	0.003093
ZFA:0000599	torus semicircularis	0.003164
ZFA:0000012	central nervous system	0.004276
ZFA:0000098	proliferative region	0.006039
ZFA:0001261	ventricular system	0.01604
ZFA:0001679	solid lens vesicle	0.032223
ZFA:0000159	tectal ventricle	0.0374

Transcriptional profile of *slc2a11b*^{sa1577}



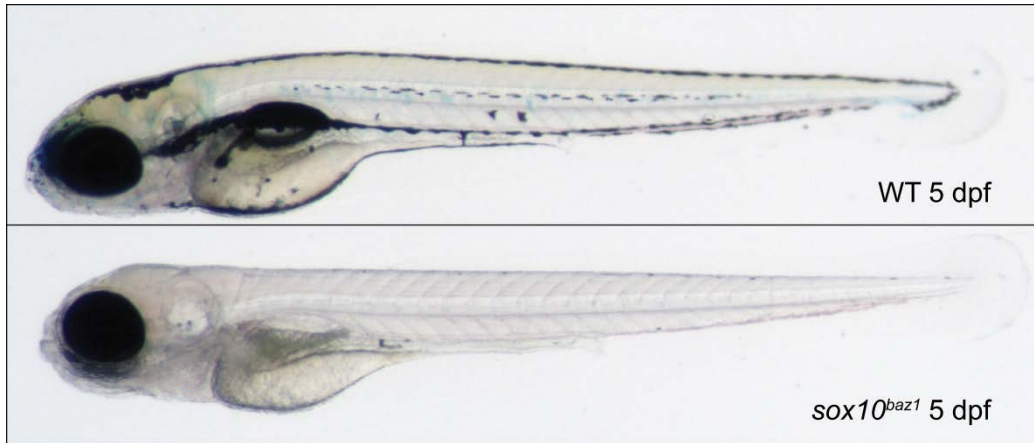
Stage	Entity	Quality	Tag
Larval:Day 5 (ZFS:0000037)	iridophore (ZFA:0009199)	quality (PATO:0000001)	normal (PATO:0000461)
Larval:Day 5 (ZFS:0000037)	melanocyte (ZFA:0009091)	quality (PATO:0000001)	normal (PATO:0000461)
Larval:Day 5 (ZFS:0000037)	xanthophore (ZFA:0009198)	absent (PATO:0000462)	abnormal (PATO:0000461)

Chr	3' end position	3' end strand	p value	Adjusted p value
5	10953779	1	0	0
12	26618916	1	0	0
21	14306023	-1	0	0
5	10963944	1	6.11E-08	0.000367
11	31618426	1	8.97E-06	0.024088



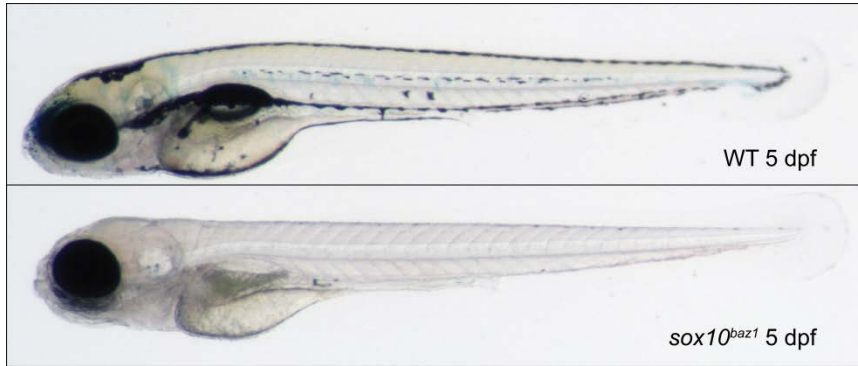
1	ENSDART00000137543	zgc:66350	-1.71583
6	ENSDART00000103270	slc22a7a	-0.99554

Transcriptional profile of *sox10^{baz1}*



- Specification of non-ectomesenchymal neural crest derivatives:
 - Pigment cells
 - Neurons: sensory, enteric, para-, sympathetic
 - Schwann cells
- Heterozygous mutations in humans cause Waardenburg syndrome through haploinsufficiency

Transcriptional profile of *sox10^{baz1}*



- 33 affected regions <100bp



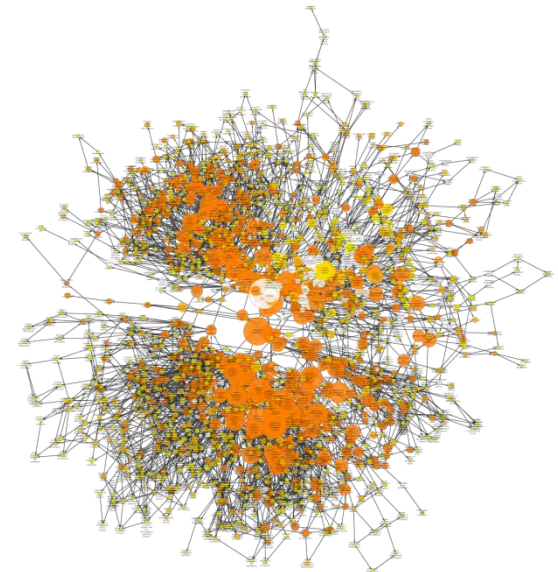
Expression enrichment

ZFA:0009090	pigment cell	
ZFA:0009076	oligodendrocyte	
ZFA:0000045	neural crest	
ZFA:0009074	macroglial cell	
ZFA:0007009	nerve	
ZFA:0009147	glial cell	
ZFA:0009169	glioblast	0.042073
ZFA:0009198	xanthophore	0.04609

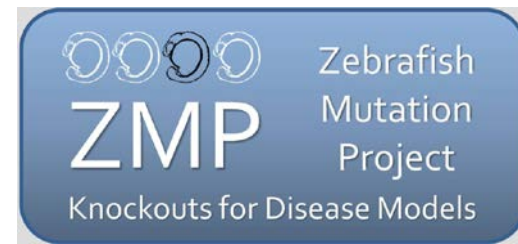
Ensembl Gene ID	Gene	Chr.
ENSDARG00000071475	<i>aox3</i>	22
ENSDARG00000013340	<i>gch2</i>	12
ENSDARG00000038609	<i>mpz</i>	2
ENSDARG00000078680	<i>vcana</i>	5
ENSDARG00000036186	<i>mbpa</i>	19
ENSDARG00000016470	<i>anxa5b</i>	1
ENSDARG00000089413	<i>mbpb</i>	16
ENSDARG00000011929	<i>plp1b</i>	5
ENSDARG00000022303	<i>hig1</i>	2
ENSDARG00000009505	<i>slmo2</i>	6
ENSDARG00000052734	<i>hmgcra</i>	5
ENSDARG00000052738	<i>hmgcs1</i>	10
ENSDARG00000034572	<i>gpr143</i>	9
ENSDARG00000058476	<i>stc1l</i>	10
	<i>cyp2k6</i>	3
	<i>zgc:73324</i>	3
	<i>stx11b.1</i>	20
	<i>MREG</i>	19
	<i>otomp</i>	2
	<i>tyrp1b</i>	1
	<i>cldnk</i>	3
	<i>nmrk2</i>	2
	<i>mlpha</i>	6
	<i>slc45a2</i>	21
	<i>ponzr4</i>	7
	<i>coq10b</i>	9
	<i>zgc:110339</i>	25
	<i>ponzr3</i>	7
	<i>slc22a7a</i>	11
ENSDARG00000060316	<i>cish</i>	6
ENSDARG00000094929	<i>si:dkey-7f3.15</i>	16
ENSDARG00000040200	<i>si:dkeyp-86f5.1</i>	3
ENSDARG00000058336	<i>SLC25A22 (1 of 2)</i>	7

Functional annotation: Current focus

- Unbiased vs. targeted analysis
 - Random mutagenesis for **unbiased** gene analysis
 - KO alleles in 26,000 genes
 - Phenotypes for 8,000 genes
 - CRISPR/Cas9 for **targeted** groups of genes
- Transcriptional profiling using DeTCT
 - Thus far 39 experiments complete, with 78 sets of samples collected
- Expand profile network analysis
- Single embryo transcriptional profiling
 - Independent from morphological phenotype
- Single cell transcript counting



Acknowledgements



ZMP:

Elisabeth Busch-Nentwich

Jeffrey Barrett

Freek van Eeden

Ross Kettleborough

Mari Niemi

Steve Harvey

James Morris

ZIRC, Eugene:

Amanda Hall

Zoltan Varga

Catherine Scahill

Animal Support:

April Freeman

Chris Dooley

James Bussell

Ian Packham

Diane Hazlehurst

ZFIN, Eugene:

Ian Sealy

Nick Harman

Leyla Ruzicka

John Collins

Nicola Goodwin

Jorge Zamora

EZRC, KIT Karlsruhe:

Neha Wali

DNA pipelines

Robert Geisler

Peter Clarke

Jana Maier

Richard Clark

Richard Gibbons

Richard White

Roiseann Cain

Samantha Carruthers

Zsofia Pusztai

