



Building Excellence in Genomics and Computational Bioscience

TILLING @ TGAC

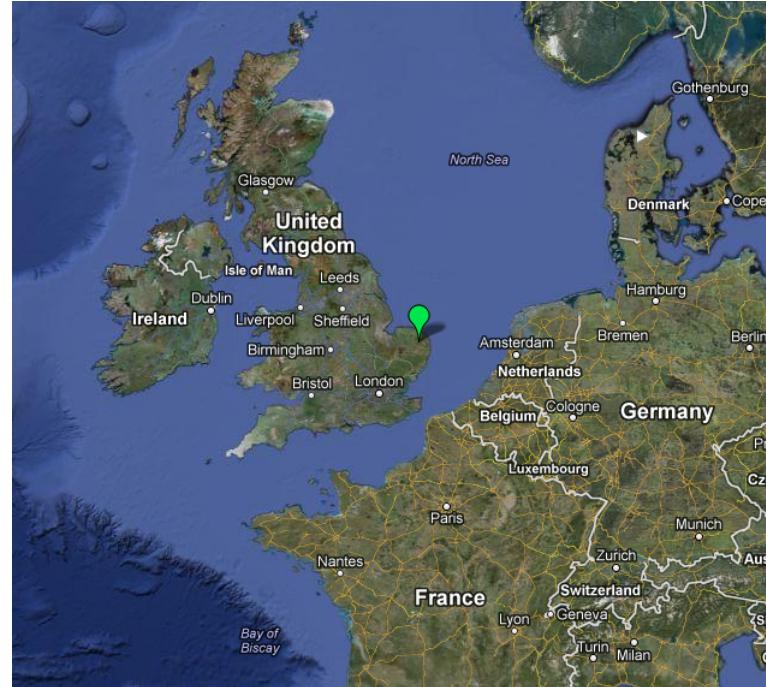
Sarah Ayling
Crop Genomics and Diversity
sarah.ayling@tgac.ac.uk



The Genome Analysis Centre



*Building Excellence in Genomics
and Computational Bioscience*



Forward genetics is the approach of determining the **genetic** basis responsible for a phenotype – classical genetics

Reverse genetics is an approach to discover the function of a gene by analyzing the phenotypic effects of specific engineered gene sequences.

TILLING: Targeting Induced Local Lesions in Genomes

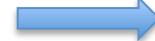
Source: Wikipedia

TILLING by sequencing



Induce
genetic
variation

Base changes:
 $C > T$; $G > A$



GCGATGAGCTCGAAGAGCGCA
↓
GCGATAAGCTCGAAGAGTGCAT
↓

Detect base
changes in
genes



Phenotype
plants with
changes in
your gene of
interest



Sequence once, database for everyone...

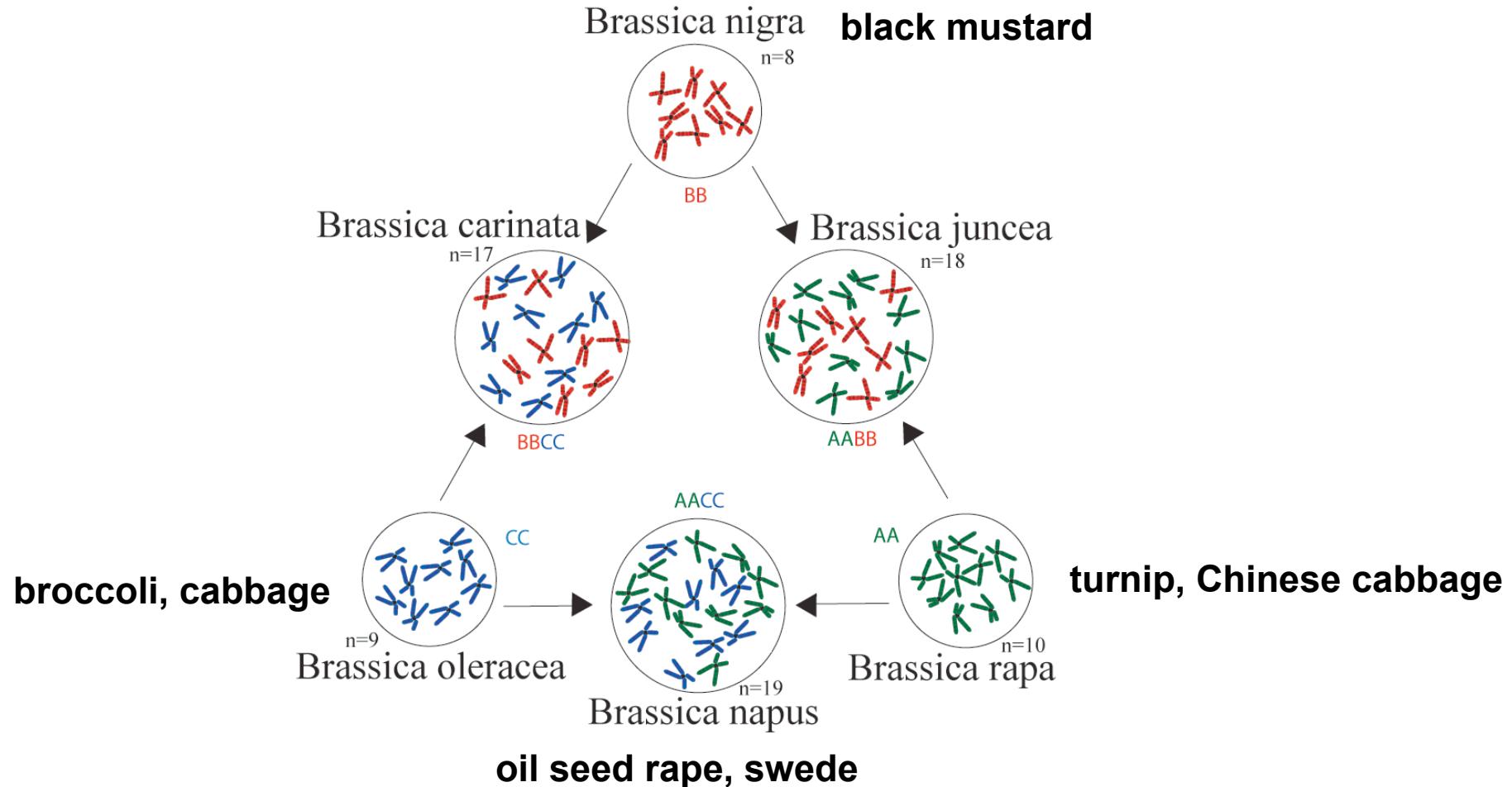
TILLING projects:



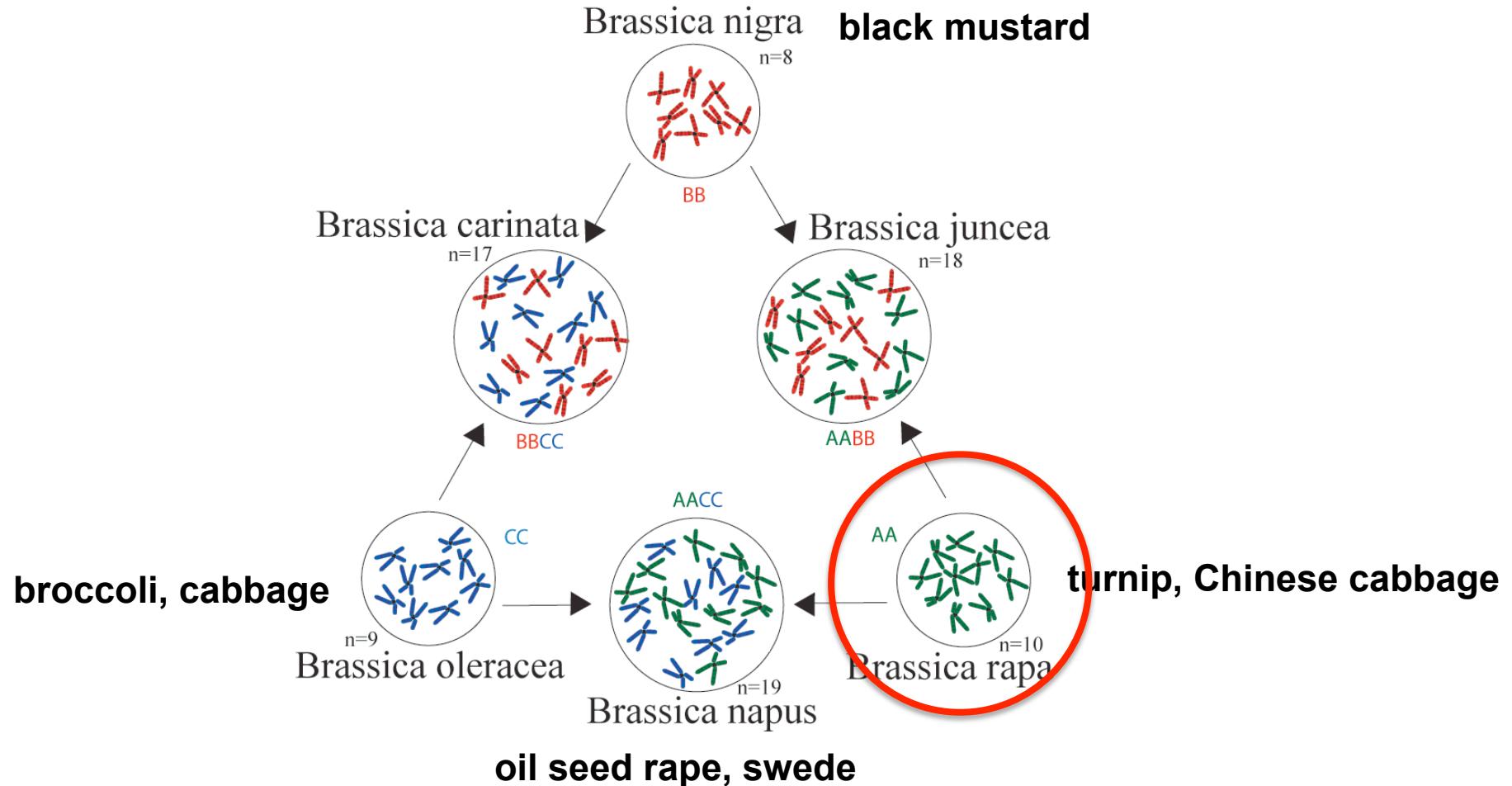
Brassica rapa



Triticum aestivum



Wikipedia (U, Jpn J Bot
7:389452, 1935)

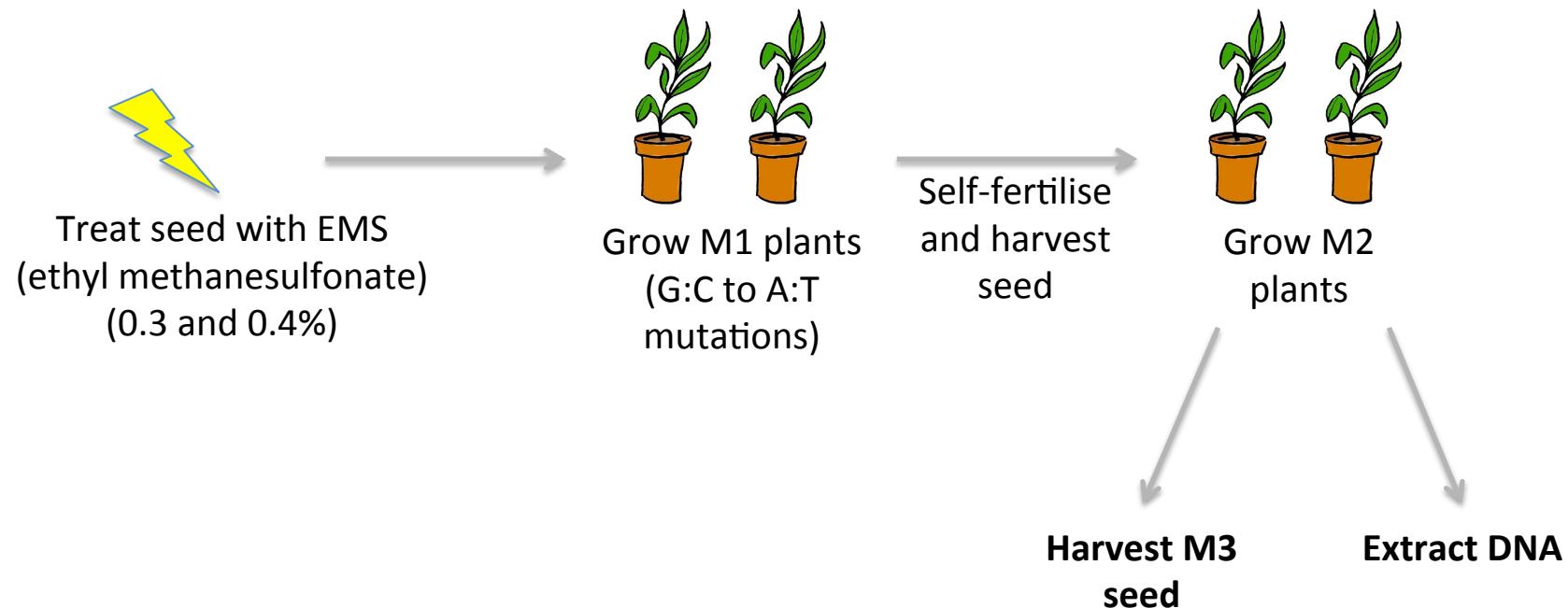


Wikipedia (U, Jpn J Bot
7:389452, 1935)

Genotype: R-o-18

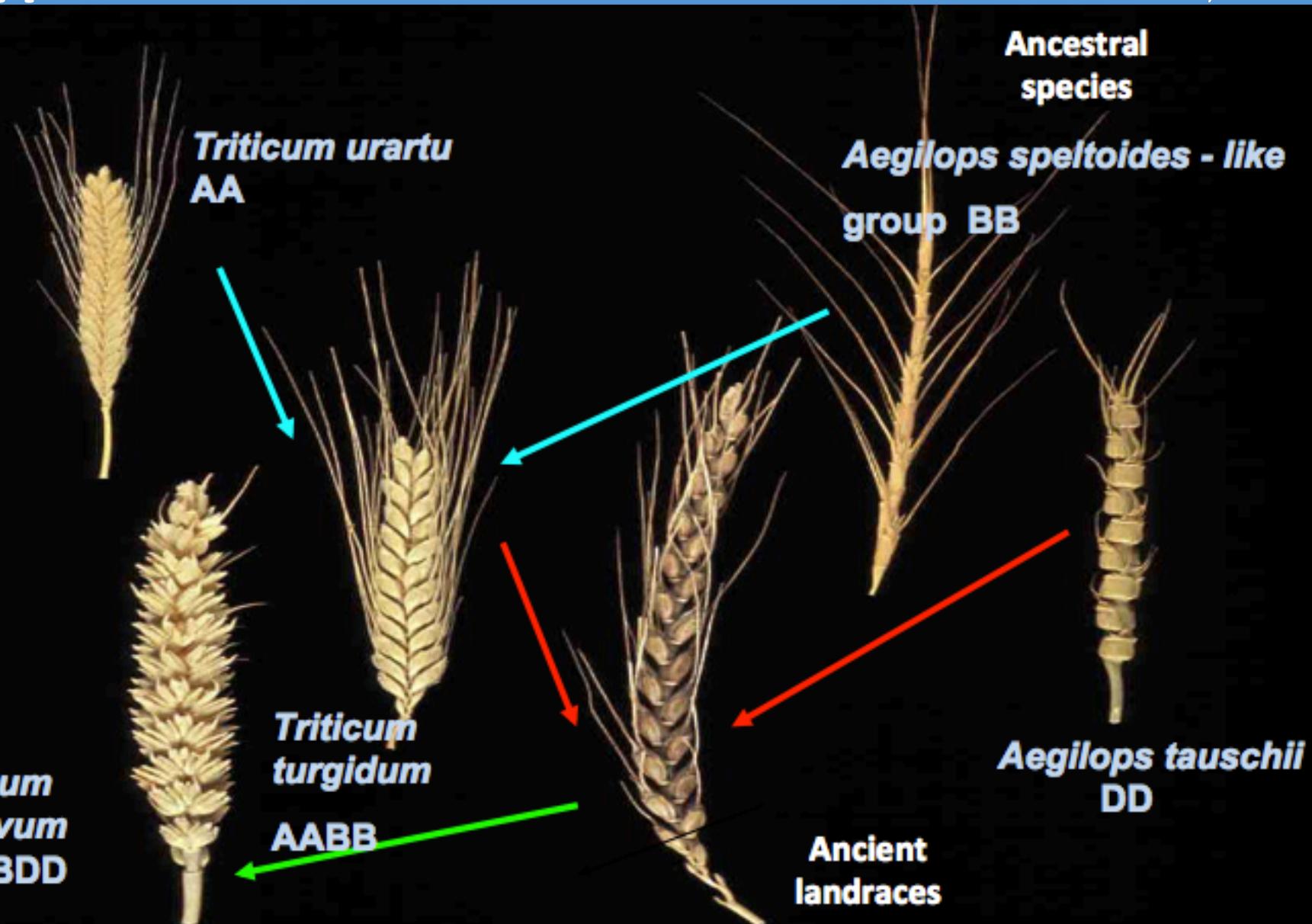
- Inbred line of *B. rapa* subsp. *Trilocularis* (Yellow Sarson)
- Closely related to *B. rapa* oilseed crops grown in Pakistan

EMS population developed in Lars Østergaard group, >1000 lines

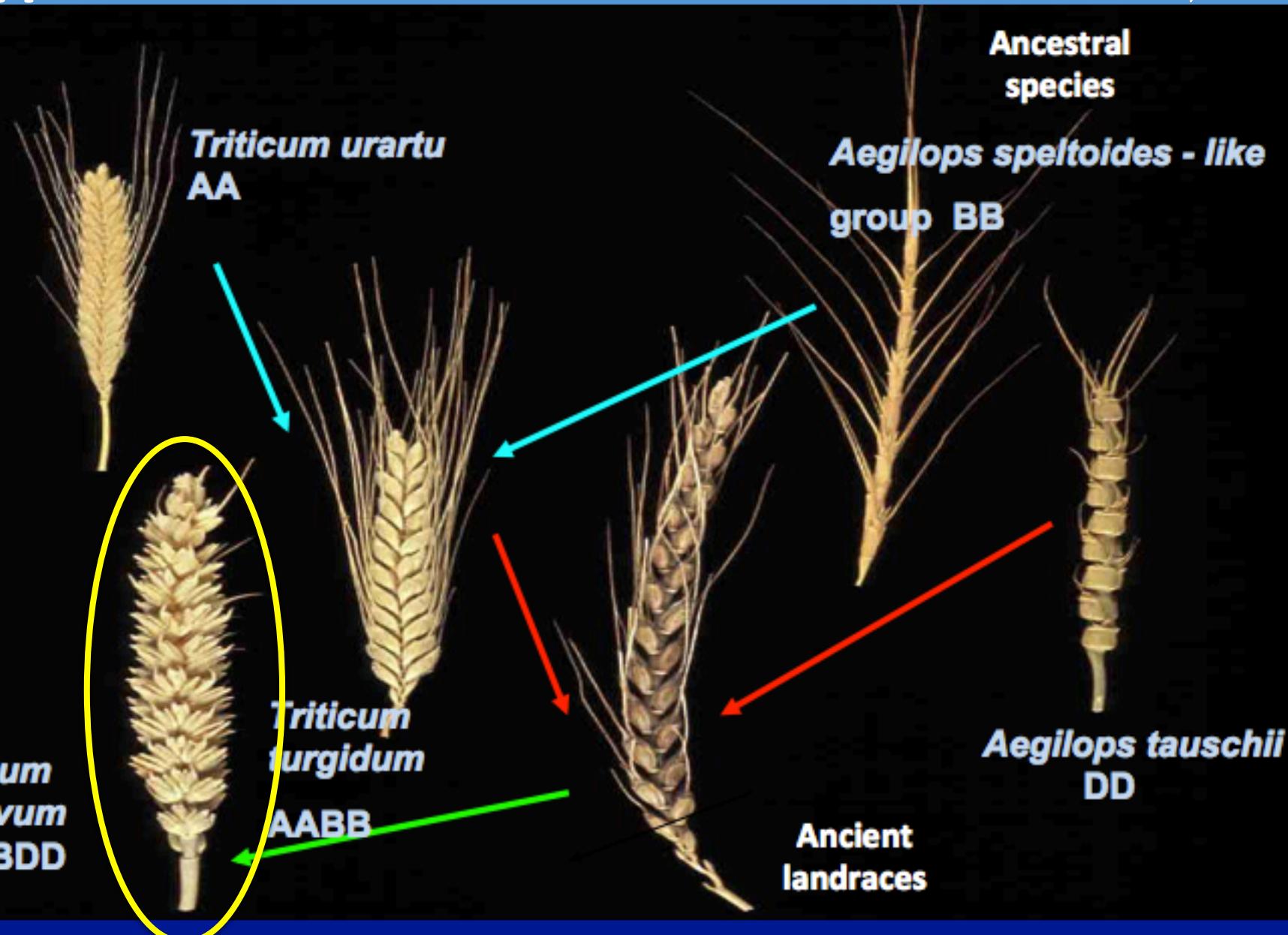


Stephenson *et al.*, BMC Plant Biology 2010

Polyplloid wheat



Polyplloid wheat

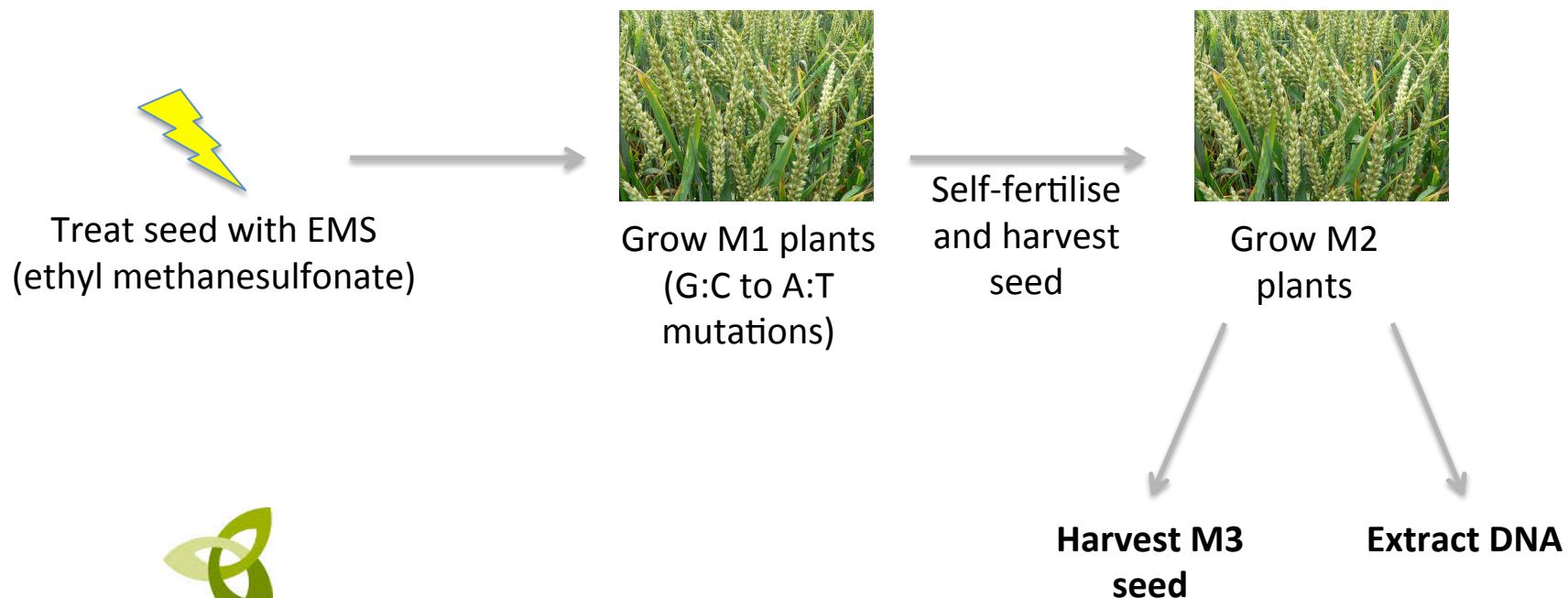


Mutant population - wheat

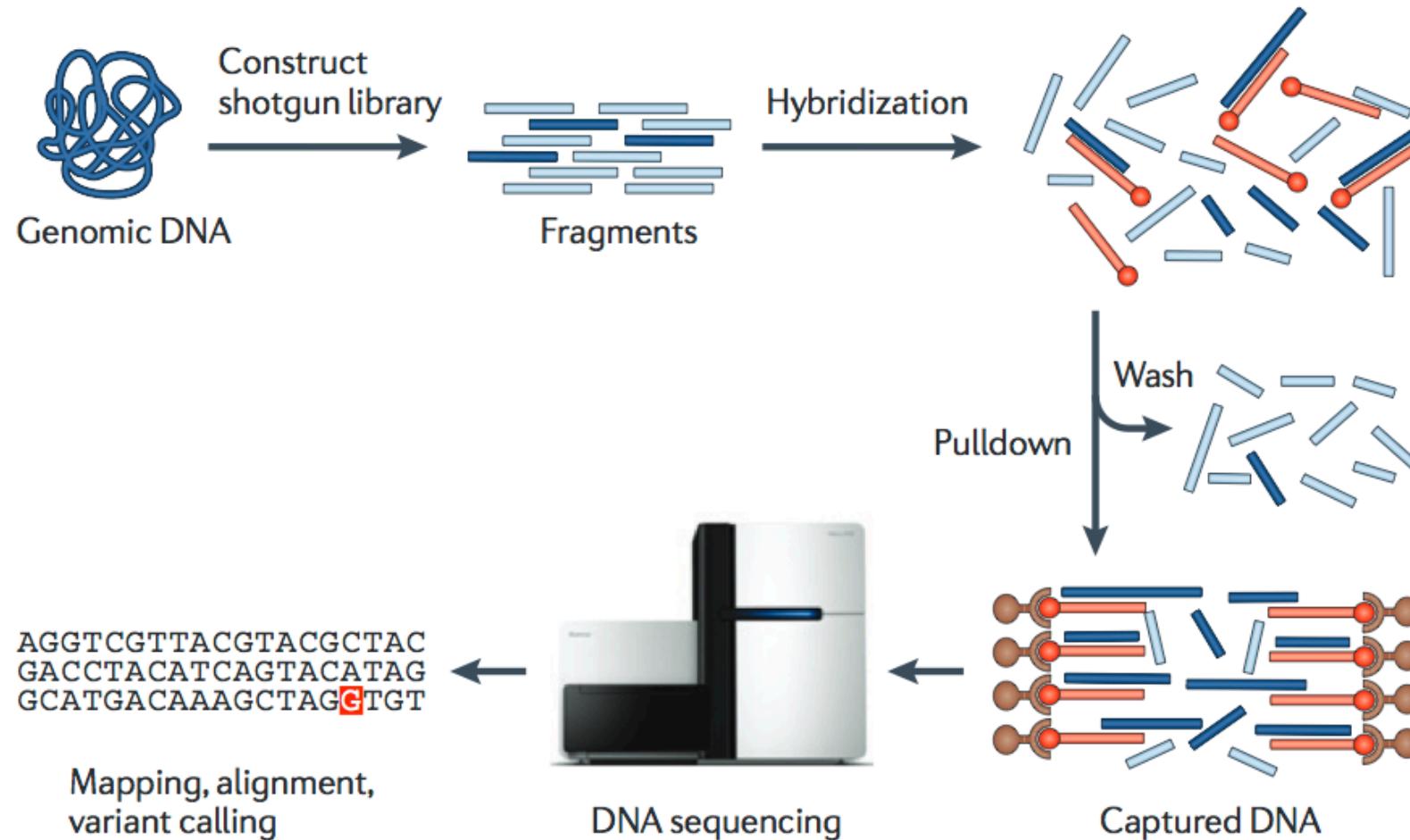
Genotype:

Triticum aestivum cv Cadenza (Bread wheat)

EMS population developed in Andy Philips group,
>1500 lines as part of WGIN1



Exome capture



Bamshad *et al.*, Nature Reviews Genetics, 2011

B. rapa genome & annotation

 BLAST | BioMart |
Brassica rapa ▾

 **Brassica rapa**
Brassica rapa subsp. *pekinensis*
Data Source [IVFCAAS](#) | Taxonomy ID [51351](#)

Go

Gene counts

Coding genes:	41,018
Non coding genes:	1,011
Small non coding genes:	1,009
Long non coding genes:	1
Misc non coding genes:	1
Pseudogenes:	817
Gene transcripts:	42,853

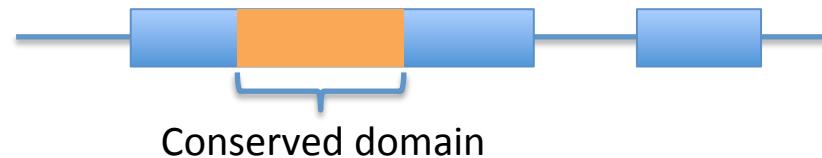
- Ian Bancroft and Martin Trick at the John Innes Centre for providing the gene annotation.

B. rapa capture design

B. rapa genome

<http://brassicadb.org>, v1.2

41,019 protein-coding genes, 48 Mb total sequence



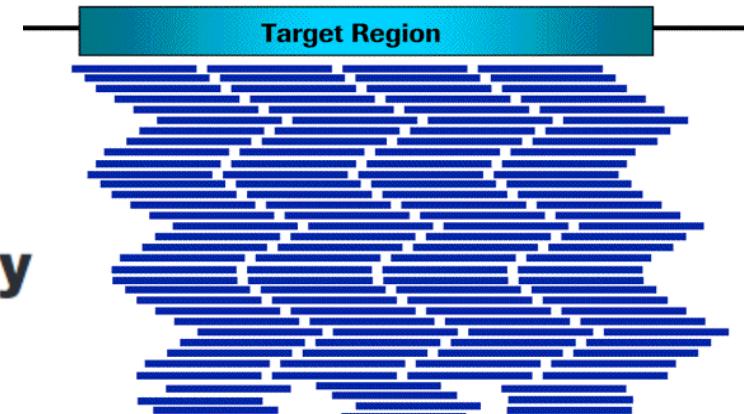
48 Mb



30 Mb

CODDLE-like: Comai lab
(Codons Optimized to Detect Deleterious Lesions)


NimbleGen
SeqCap EZ Library



Sequencing strategy

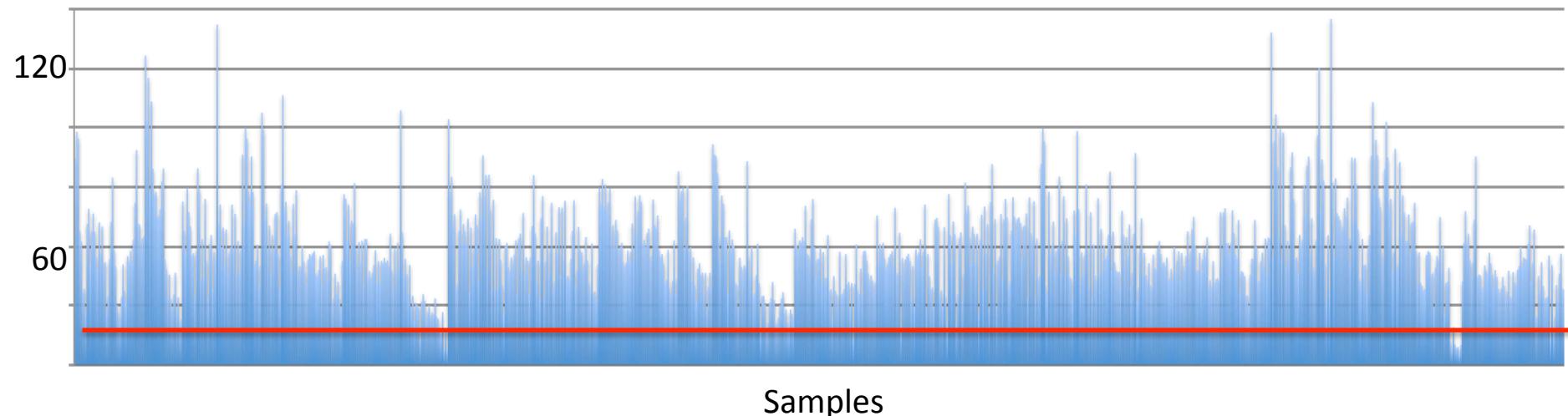
- Paired-end Illumina
- Fragment size: 300bp
- Read lengths: 2x150bp
- 24-plex run over 2 lanes
 - = 12-plex



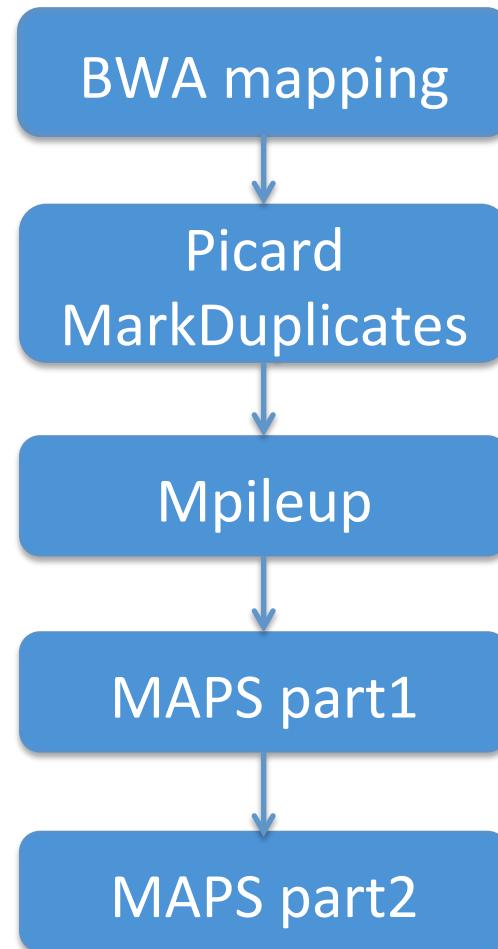
Sequencing results

Number of samples processed	998
Mean number of reads per sample	25,892,695
Minimum number of reads for a sample	3,755,166
Mean target coverage	38.5
Mean percentage of targets with <2x coverage	5.1
Mean percentage of target bases with $\geq 10x$ coverage	86.3

Target coverage variability between samples

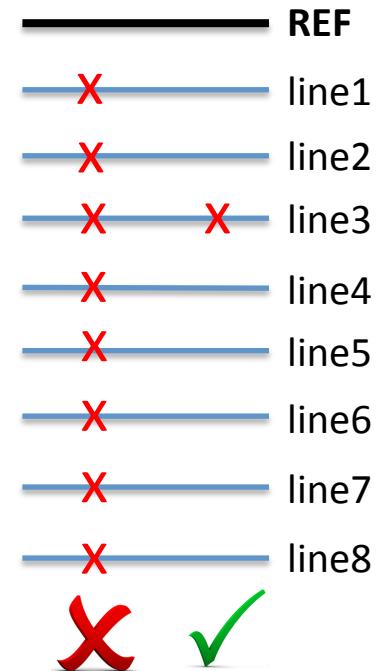


MAPS pipeline



MAPS part1:
(whole batch)
 $\text{minLib}=n-2$
 $\text{minCov}=10$
 $\text{maxCov}=1000$
 $\text{hetOneMinPer}=20$

MAPS part2:
(per sample)
 $\text{minLib}=n-2$
 $\text{minCov}=10$
 $\text{hetMinPer}=30$
 $\text{homMinCov}=3$
 $\text{hetMinCov}=6$



Henry *et al.* 2014. Plant Cell

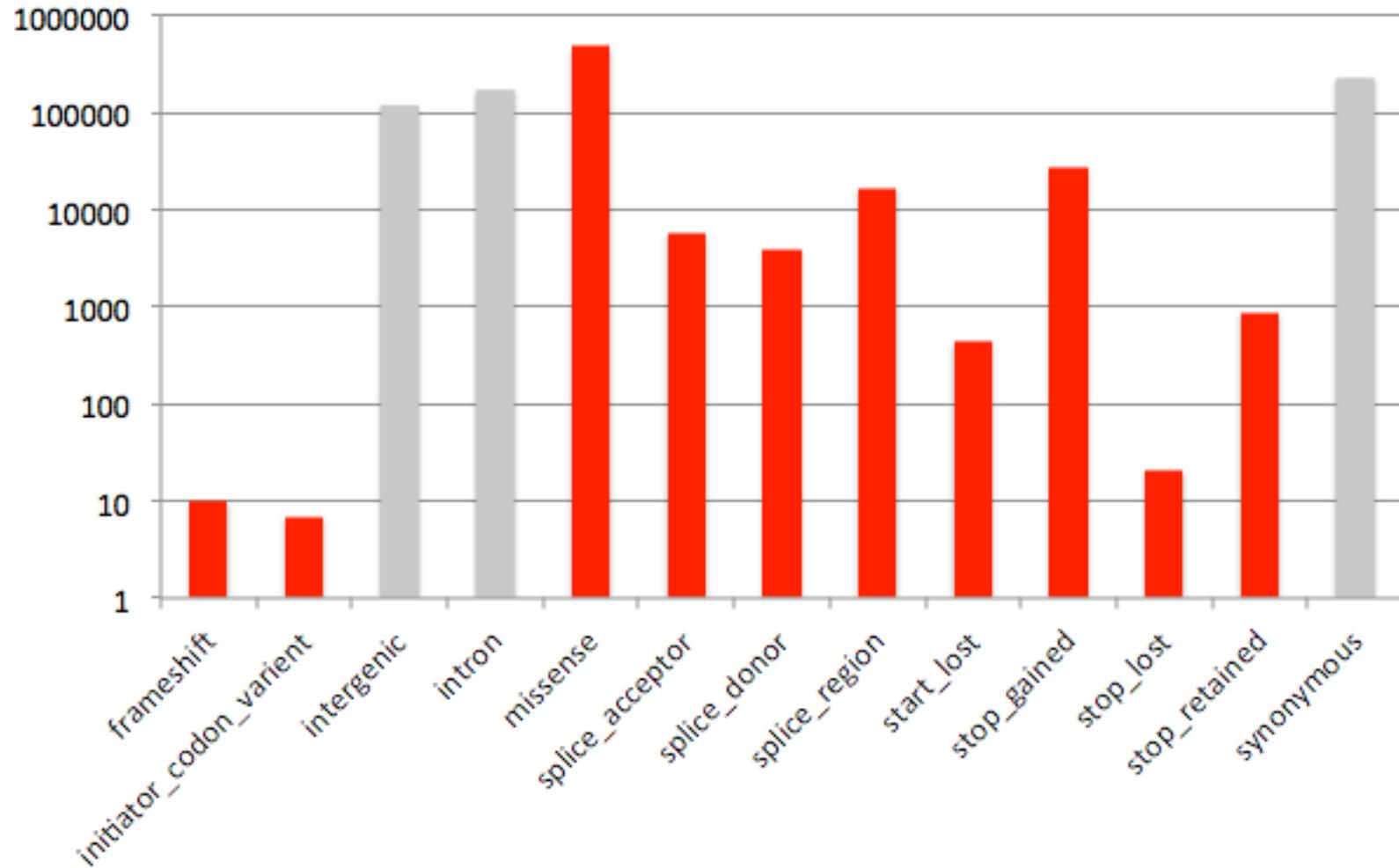
Mean number of mutations per sample	1188.7
Mean mutation density	1 per 49.9 kb
% Homozygous	10
% Heterozygous	90
Mean % EMS mutations (C:G->A:T)	99.1



SnpEff

Genetic variant annotation and effect prediction toolbox.

92.2% of *B. rapa* genes contain non-synonymous mutations



Accessing the data

The screenshot shows the TGAC Browser interface for the brassica_revgen dataset. The main window displays genomic tracks for Order line LIB12115, including Genes, EMS_mutation, Bait_Interval, and Target_Interval. A message window is open in the foreground, allowing the user to compose an email. The message details are as follows:

From: Tgac (Jon Wright)

To: orders@revgenuk.ac.uk

Cc: [empty]

Subject: Order line LIB12115

The message window also features a rich text editor toolbar with various formatting options like bold, italic, underline, and alignment.

Slide: Jon Wright

Accessing the data

The screenshot shows a web browser window for the TGAC Browser. The URL in the address bar is `browser.tgac.ac.uk/brassica_revgen/index.jsp?query=A01&coord=1`. The search term `brassica_revgen : bra033779` is entered in the search bar. The results are displayed under the heading "Results for searched query". A note indicates "(Limited to first 100 match)". The search result table has columns: Track, Gene, Reference Name, Reference Coord Sys, Position, and Link. The data row shown is: Gene_genes, Bra033779, A01, 1, 14312347-14314113, [♂](#).

Track	Gene	Reference Name	Reference Coord Sys	Position	Link
Gene_genes	Bra033779	A01	1	14312347-14314113	♂

© 2011 - 2014 [The Genome Analysis Centre](#) | Version: 0.2.1 | [Feedback](#)

Accessing the data

Screenshot of the TGAC Browser interface showing a BLAST search results page.

The browser window title is "TGAC Browser - brassica_revgen". The URL is "browser.tgac.ac.uk/brassica_revgen/blast.jsp". The search term is "brassica_revgen : A01".

Parameter settings:

- Gap Costs: Existence: 5 Extension: 2
- Word Size: 16
- Short Queries: Automatically adjust parameters for short input sequences.

BLAST History 1:

- BLAST job v9GyXPJf

Query #	Subject #	% identity	alignment length	mismatches	gap openings	q.start	q.end	s.start	s.end	e-value	bit score	Subject db	Download Sequence
my_gene	A01	99.60	1767	0	3	1	1760	14312347	14314113	0.0	3328	Brapa_sequence_v1	Download
my_gene	A01	88.64	449	30	3	573	1020	2514568	2514140	7e-152	540	Brapa_sequence_v1	Download
my_gene	A01	87.80	451	39	5	573	1022	11459833	11460268	3e-140	502	Brapa_sequence_v1	Download
my_gene	A01	87.90	463	39	7	572	1020	9234522	9234063	1e-139	500	Brapa_sequence_v1	Download
my_gene	A01	94.80	269	13	1	756	1024	15447658	15447925	1e-117	427	Brapa_sequence_v1	Download
my_gene	A01	85.24	454	41	7	572	1023	9198217	9198646	2e-115	419	Brapa_sequence_v1	Download
my_gene	A01	94.01	267	14	2	754	1019	8599362	8599097	3e-110	402	Brapa_sequence_v1	Download
my_gene	A01	92.62	271	18	2	754	1023	13990327	13990058	1e-105	387	Brapa_sequence_v1	Download
my_gene	A01	93.23	266	15	3	757	1020	16040003	16039739	3e-103	379	Brapa_sequence_v1	Download
my_gene	A01	92.54	268	17	3	754	1020	15428525	15428790	6e-101	371	Brapa_sequence_v1	Download
my_gene	A01	92.54	268	16	3	754	1019	21296996	21297261	6e-101	371	Brapa_sequence_v1	Download

Page footer: © 2011 - 2014 The Genome Analysis Centre | Version: 0.2.1 | [Feedback](#)



[FAQs](#) [How to Order](#) [Contact Us](#) [Home](#)

Latest News

About RevGenUK ▶

Models

TILLING

Links

Resources

Publications

Client Portal

TILLING is sweet: see Lin et al. *Nature*



Welcome to RevGenUK

The John Innes Centre (JIC) is at the forefront of reverse genetics through its establishment of TILLING and de-TILLING technologies in model legumes. Reverse genetics involves making a change to a specific gene, seeing what effect this has, and so working out the gene's normal function. We aim to develop these into a biological and bioinformatics resource for the wider plant scientific community via the RevGenUK project, providing an integrated reverse genetics platform for model plants.

RevGenUK aims to be "A 'single-stop' shop for use in functional genomics research consisting of an efficient, self-sustaining, extendable service for reverse genetics."

Poor	<input type="checkbox"/>
Satisfactory	<input type="checkbox"/>
Good	<input type="checkbox"/>
Excellent	<input checked="" type="checkbox"/>

We should be very grateful if you would take a few minutes to fill in our [questionnaire](#). Your answers and opinions will help us to shape and improve the RevGenUK service in the future.



Next steps?

- Order your seeds (multiple mutants per gene)
- Grow your plants
 - Check phenotypes
- Cross/self to obtain homozygous alleles
 - Check with KASP markers
- Grow your plants
 - Check phenotypes
- Publish ☺

Acknowledgements



Brassica:

Jon Wright

Wheat:

Paul Bailey

Ksenia Krasileva

Ricardo Ramirez-

Gonzales

Platforms & Pipelines:

Leah Clissold

and team



Trevor Wang

Fran Robson

Lars Østergaard

Saleha Bakht



John Innes Centre

Cristobal Uauy



UNIVERSITY OF CALIFORNIA

Jorge Dubcovsky

Hans Vasquez-

Gross

Luca Comai

Meric Lieberman

Isabelle Henry



ROTHAMSTED
RESEARCH

Andy
Phillips

