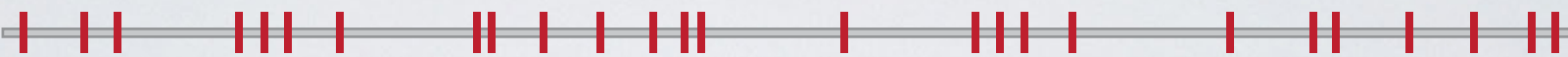


What is RAD-seq?

(Restriction-site Associated DNA)

Chr I 

22,830 *SbfI* sites in Stickleback

~ 45,000 RAD-Tags

HiSeq Illumina Lane:

150 million reads, 96 barcoded individuals

TGCAGG  TTCTGTTCACTGAAGCAGACGCGCGTGTATGGA

①  CCTGCAGG  CCTGCAGG

TGCAGG  TTCTGTTCACTGAAGCAGACGCGCGTGTATGGA

TGCAGG  CCTGCAGG

Stacks

1. Stacks is implemented in C++ with Perl wrappers
2. Core programs are shared-memory parallelized
3. Data visualization provided by MySQL / PHP web interface

30,000 Lines of C++
6,500 Lines of PHP
8,000 Lines of Perl



Angel
Amores



Bill
Cresko

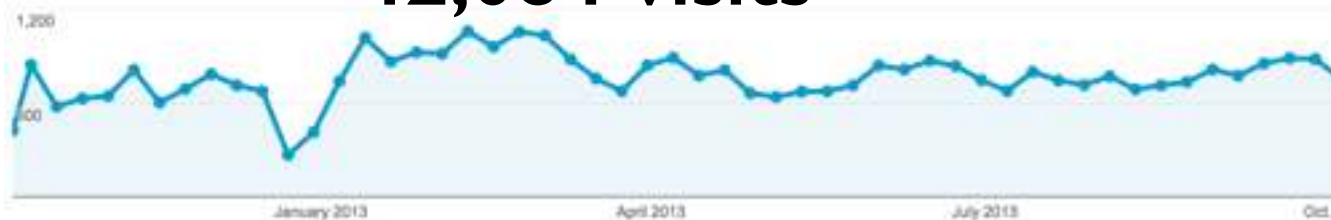


Paul
Hohenlohe

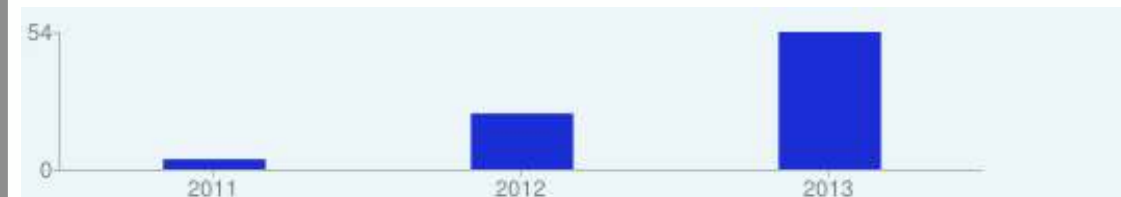
Dec. 2012 - Dec. 2013

13,254 visitors

42,084 visits



87 Citations

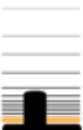


[Stacks: building and genotyping loci de novo from short-read sequences](#)
JM Catchen, A Amores, P Hohenlohe, W Cresko... - G3: Genes, Genomes, Genetics, 2011
[Cited by 80](#) - [Related articles](#) - [All 11 versions](#)

[Stacks: an analysis tool set for population genomics](#)
J Catchen, PA Hohenlohe, S Bassham, A Amores... - Molecular ecology, 2013
[Cited by 7](#) - [Related articles](#) - [All 4 versions](#)



Stacks



Stacks

Containers

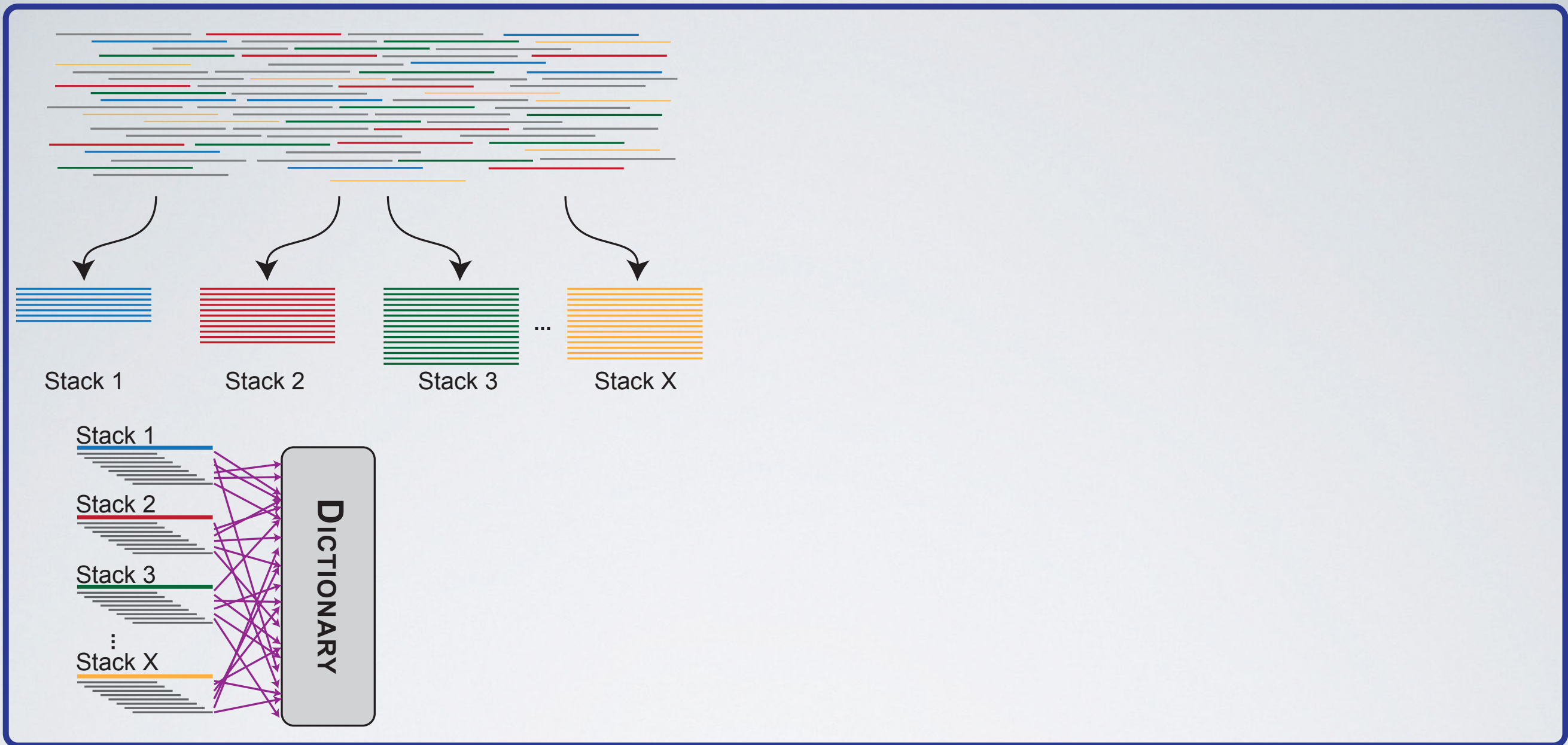




Stacks



Stacks



C

CT

Matching stacks using k-mers

A string can be represented as a stack of k-mers

TACCACATC
TACCACATC
ACCACATC
CCACATC
CACATC
ACATC
CATC
ATC
TC
C

AAATCCTTT
AAATCCTTT

1. Replace an n^2 algorithm with linear hash table lookup
2. Further reduce the search space:
 1. Given a number of allowed mismatches calculate the maximum useful k-mer size
 2. For each sequence: lookup matching k-mers and record associated reads
 3. For pairs of sequences with minimum number of k-mers, align reads.

A
AA
AAA
AAAT

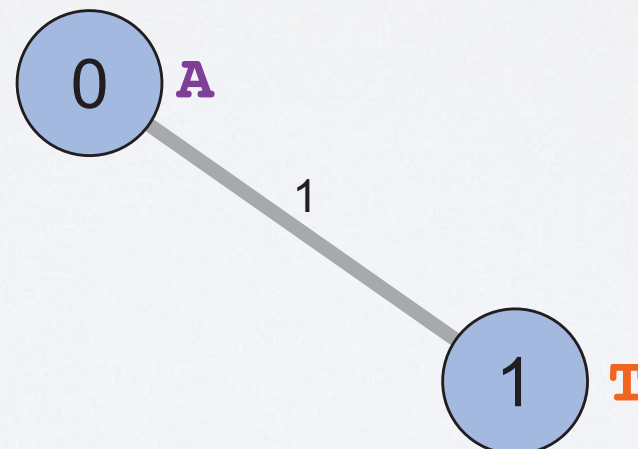
ACGTGTGAAAAATCC
CGTGTGAAAAATCCT
GTGTGAAAAATCCTT
TGTGAAAAATCCTTT

ACGTGTGAAAAATCC
CGTGTGAAAAATCCT
GTGTGAAAAATCCTT
TGTGAAAAATCCTTT

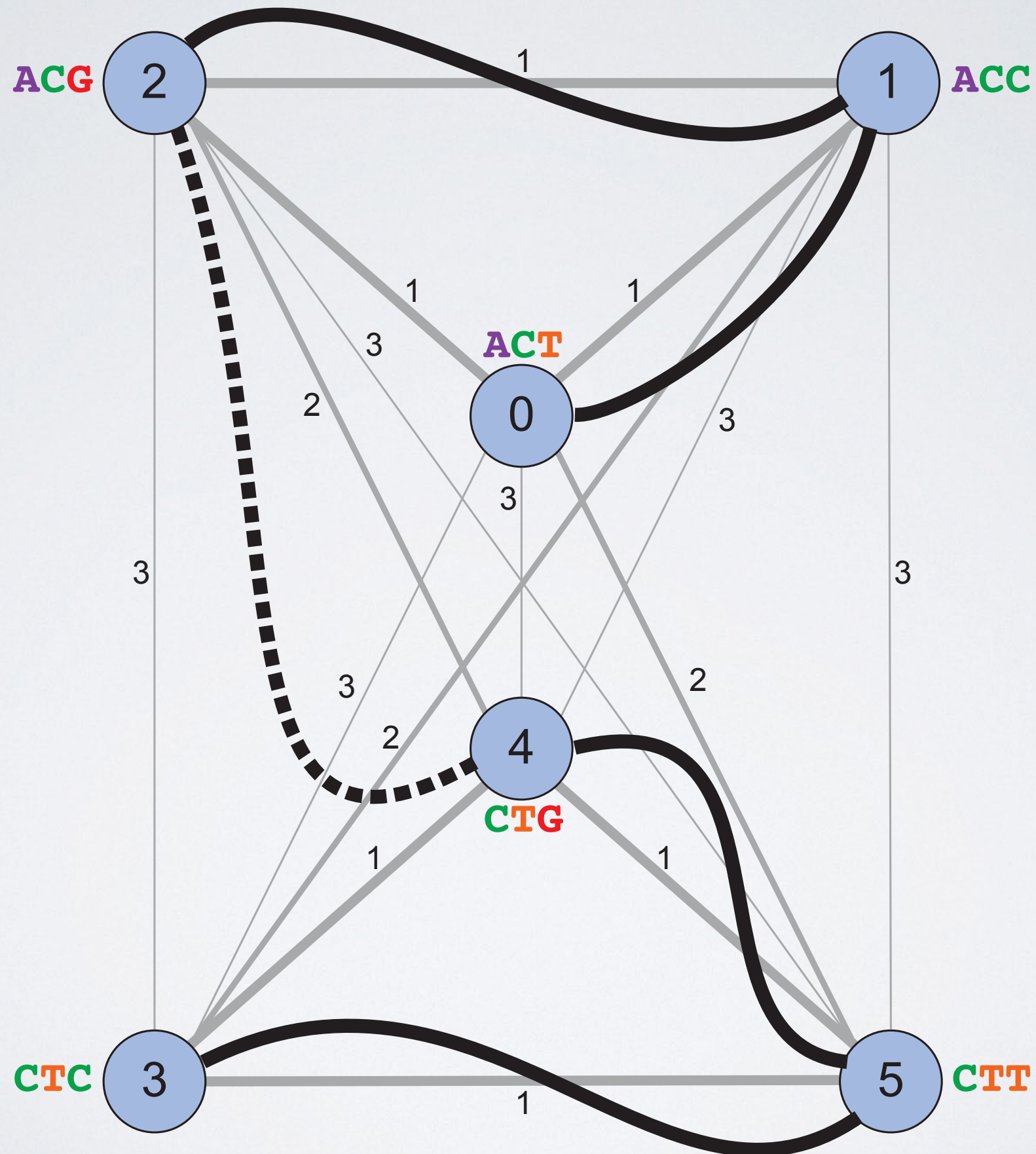
Building loci using minimum spanning trees

TGCAGGTACCCCGTCCCC**A**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**A**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**A**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**A**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**A**GCGGAGGACCTGTTACC

TGCAGGTACCCCGTCCCC**T**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**T**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**T**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**T**GCGGAGGACCTGTTACC
TGCAGGTACCCCGTCCCC**T**GCGGAGGACCTGTTACC



Building loci using minimum spanning trees



Differentiating SNPs from error

Restriction enzyme recognition site

Reference
genome
sequence

sequence
reads

```
CTTCAGGTTGGGTGAGTTGTCATCAGTCGGAATGCGCAGGTCACCTTACCTGCAGGCAGCTCTCTGAAGCGCAGGTACTCCATCGACCGGGTGGTGACTAG
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
ATCAGTCGGAATGCGCAGGTCACCTTACCTGCA
ATCAGTCGGAATGCGCAGGTCACCTTACCTGCA
ATCAGTCgAATGCTCAGGTCaCTTacctGcA
ATCAGTCGGAATGCGCAGGTCACCTTAcctGcA
ATCAtTCGGAATGCGCAGGTCACCTTAcctgcA
ATCAGTCGGAATGCGCAGGTCACCTTACCTGCA
ATCAGTCGGAATGCGCAGGTCACCTTAcCtGcA
ATCAGtCGGAATgCGCAGGTCACCTTACCTGCA
ATCAGTCGGAATGCGCAGGTCACCTTAcctgcA
cgcagtbggAAtgcGCAGgTcaCTTcctgcA
AtCAGTCGGAATGCGCAGGTCACCTTAcctgcA
ATCAGTCGGAATGCGCAGGTCACCTTAcctgcA
ATCAGTCgGAATGCGCAGGTCACCTTAcctgcA
ATCAGTCGGAATGCGCAGGTCACCTTAcctgcA
ATCAGTCGGAATGCGCAGGTCACCTTAcctgcA
ATCAGTCGGAATGCGCAGGTCACCTTAcctgcA
ATCAGTCGGAATGCGCAGGTCACCTTAcctgcA
TgcaggCAGcTCTCTGAAGCGCAGGtACTCCA
TgcaggCAGCTCTCTGAAGCGCAGGtACTCCA
TgCaggCAGCTCTCTGAAGCGCAGGgACTcCA
TGCAGGCAGCTCTCTGAAGCGCAGGtaCTCCA
TgcaggCAGCTCTCTGAAGCGCAGGtACTCcA
tgcaggcAGCtCTCTGAAGagaaGgtaCTCca
TgCaggCAGCTCTCTGAAGCGCAGGtACTcCA
TgcaggCAGCTCTCTGAAGCGCacGtaCtcCa
TgcaggcAGcTaTCTGAAGcgCAgGTActcca
TGCaggCAGCTCTCTGAAGCGCAGGtACTCCA
TgCaggCAGCTCTCTGAAGCGCAGGtAcTcca
TGCaGgCAGCTCTCTGAAGCGCAGGTACTCCA
tgcaggCAGCTcTaTGAAGcGCAGGtActcca
tgcaggCagCtCTCTgaAGcgAagGgACTcca
ggcaggCAGCtCTCtGaaGcGcAggtactcca
TgCaggCAGCTCTCTGAAGCGCAGGtAcTCca
```

Differentiating SNPs from error

T
T
G
T
T
T
T
T
T
T
T
T
G
T
T

```
CTTCAGGTTGGGTGAGTTGTCATCAGTCGGAATGCGCAGGTCACTTACCTGCAGGCAGCTCTCTGAAGCGCAGGTACTCCATCGACCGGGTGGTGACTAG
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
ATCAGTCGGAATGCGCAGGTCACTTACCTGCA
ATCAGTCGGAATGCGCAGGTCACTTACCTGCA
ATCAGTCgAATGCTCAGGTCaCTTacctGcA
ATCAGTCGGAATGCGCAGGTCACTTAcctGcA
ATCA+TCGGAATGCGCAGGTCACTTAcctgca
ATCAGTCGGAATGCGCAGGTCACTTACCTGCA
ATCAGTCGGAATGCGCAGGTCACTTAcCtGcA
ATCAGtCGGAATgCGCAGGTCACTTACCTGCA
ATCAGTCGGAATGCGCAGGTCACTTAcctgca
cgcagttggAAtgCGCAGgTcaCTTccctgca
AtCAGTCGGAATGCGCAGGTCACTTAcctgca
ATCAGTCGGAATGCGCAGGTCACTTAcctgca
ATCAGTCGGAATGCGCAGGTCACTTAcctgca
ATCAGTCgGAATGCGCAGGTCACTTAcctgca
ATCAGTCGGAATGCGCAGGTCACTTAcctgca
ATCAGTCGGAATGCGCAGGTCACTTAcctgca
ATCAGTCGGAATGCGCAGGTCACTTAcctgca
ATCAGTCGGAATGCGCAGGTCACTTAcctgca
TgcaggCAGcTCTCTGAAGCGCAGGtACTCCA
TgcaggCAGCTCTCTGAAGCGCAGGtACTCCA
TgCaggCAGCTCTCTGAAGCGCAGGgACTCCA
TGCAGGCAGCTCTCTGAAGCGCAGGtACTCCA
TgcaggCAGCTCTCTGAAGCGCAGGtACTCcA
tgcaggcAGCtCTCTGAAGagaaGtACTCca
TgCaggCAGCTCTCTGAAGCGCAGGtACTCcA
TgcaggCAGCTCTCTGAAGCGCAGGtACTCcA
TgcaggcAGcTATCTGAAGcgCAGGTActcca
TGCaggCAGCTCTCTGAAGCGCAGGtACTCCA
TgCaggCAGCTCTCTGAAGCGCAGGtACTCCA
TGCaGgCAGCTCTCTGAAGCGCAGGtACTCCA
tgcaggCAGCTcTATGAAGcGCAGGtActcca
tgcaggCagCtCTCTgaAGcgAgGgACTcca
ggcaggCAGCtCTCtGaaGcGcAggtactcca
TgCaggCAGCTCTCTGAAGCGCAGGtACTCCA
```

Differentiating SNPs from error

T
T
G
T
T
T
T
T
T
T
T
T
T
G
T
T

The reads are 14 T and 2 G:

GT heterozygote?

GG homozygote with error?

AA homozygote with lots of error?

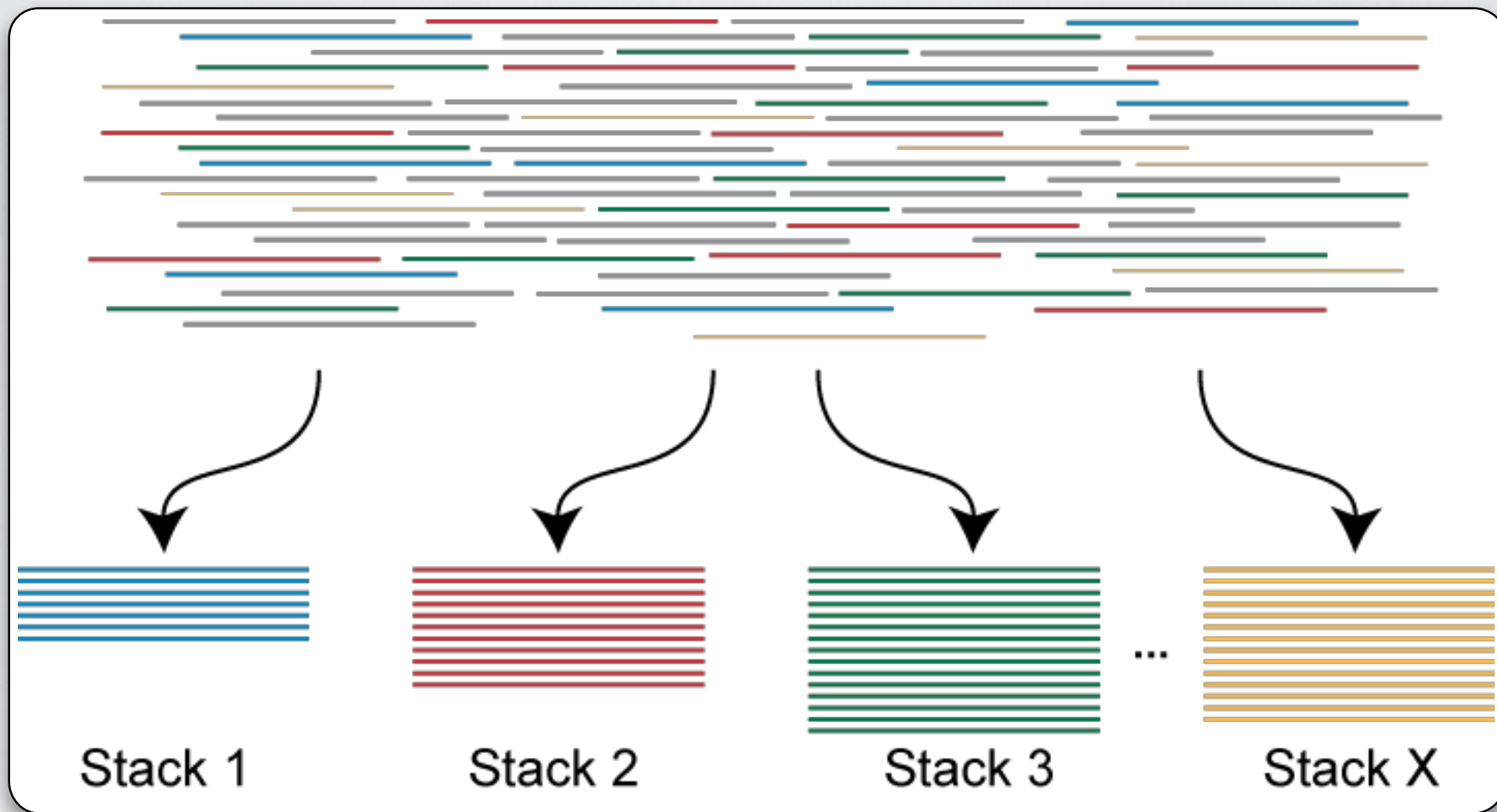
Needed a rigorous method to call genotypes

Parameters to control stack formation

Parameter Description	denovo_map.pl Parameter	Pipeline component	Component Parameter	Default Value
Minimum stack depth / minimum depth of coverage	<code>-m</code>	ustacks	<code>-m</code>	3
Distance allowed between stacks	<code>-M</code>	ustacks	<code>-M</code>	2
Distance allowed between catalog loci	<code>-n</code>	cstacks	<code>-n</code>	0

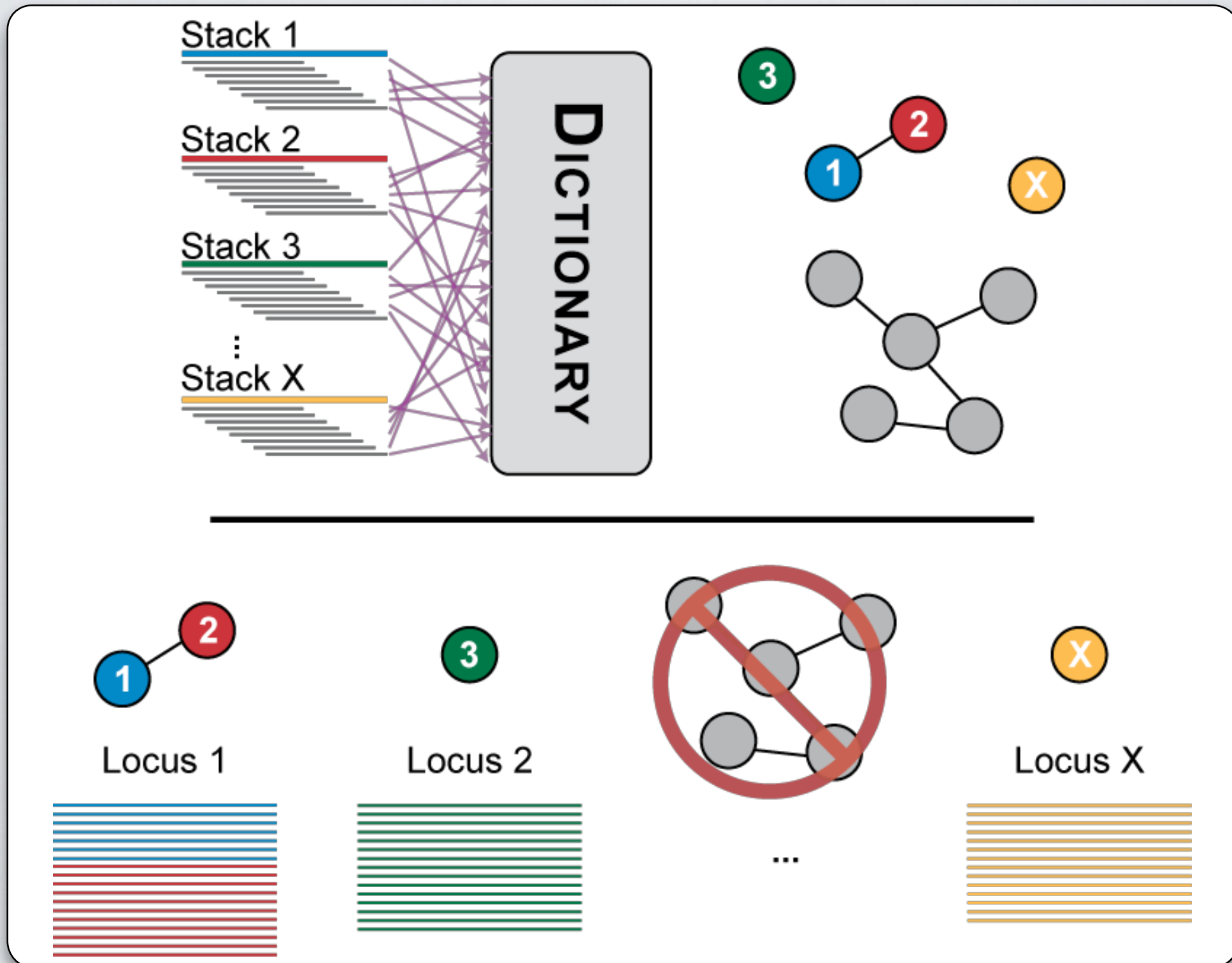
Parameters to control stack formation, ctd.

`-m <min stack depth>`



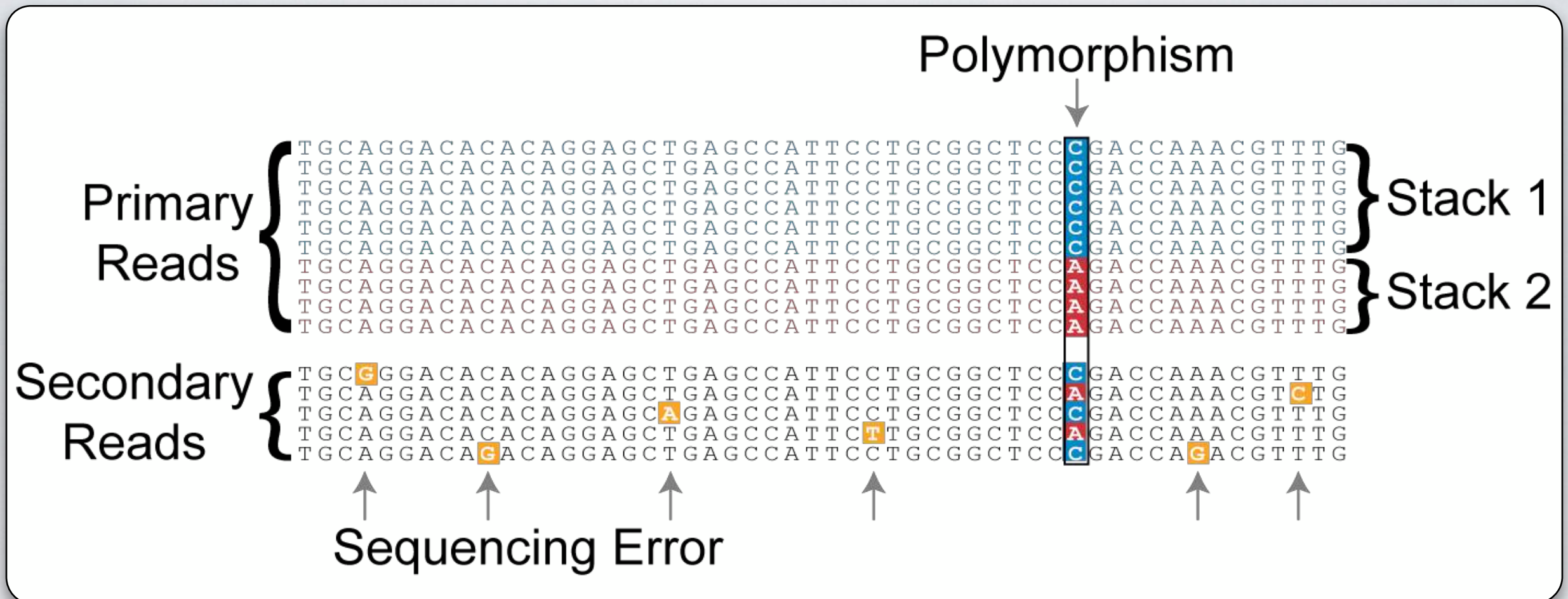
Parameters to control stack formation, ctd.

-M <distance allowed between stacks>



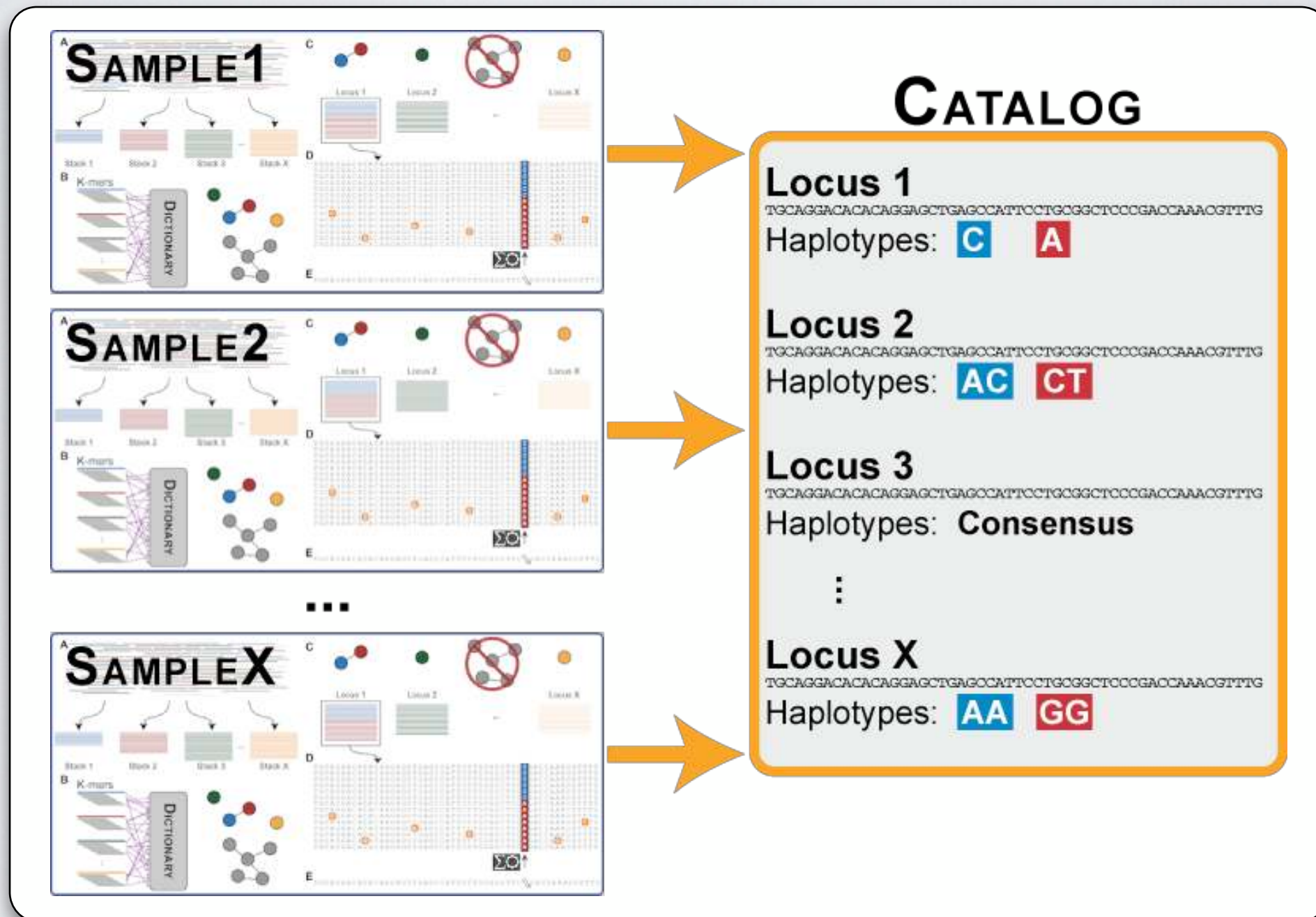
Parameters to control stack formation, ctd.

-N <secondary read distance>



Parameters to control stack formation, ctd.

-n <distance between catalog loci>



Stacks

Stacks Analysis Pipeline: RAD-Tag Catalog Viewer

http://genome.uoregon.edu/stacks/catalog.php?id=1&db=gartut_radtags&p=1&pp=10&filter_type[]=cata&filter_cata=103&filter_alle_l=1&fi Google

1 (1 tags) tags per page 10

Id	SNP	Consensus	Matching Parents	Progeny	Marker	Ratio	Genotypes
~ 103 annotate	Yes [2nuc]	TGCAGGAGCCCTCCCACTCGCTGATGGCCACTCCATTTCAGTGGACCGAGAGCCCAAAGCAACACTTCACATTC	2	92 / 91	ab/ac	aa: 25 (27.5%) ab: 24 (26.4%) ac: 18 (19.8%) bc: 24 (26.4%)	91

SNPs
 Column: 52; G/A
 Column: 70; T/G

Alleles
 a : GT
 b : GG
 c : AG

Matching Samples
 View: ☒ Haplotypes ☐ Allele Depths ☐ Genotypes

Male GT / GG	Female AG / GT	Progeny 1 GT	Progeny 2 AG / GG	Progeny 3 GG / AG	Progeny 4 GG / GT	Progeny 5 GG / AG	Progeny 6 AG	Progeny 7 GT / GG	Progeny 8 AG / GT
Progeny 9 GT	Progeny 10 GT	Progeny 11 GG / GT	Progeny 12 GT / AG	Progeny 13 GG / AG	Progeny 14 GT / AG	Progeny 15 GT / GG	Progeny 16 GG / GT	Progeny 17 GG / AG	Progeny 18 GT
Progeny 19 GT / AG	Progeny 20 AG / GG	Progeny 21 GT / AG	Progeny 22 AG / GT	Progeny 23 GG / AG	Progeny 24 GG / AG	Progeny 25 GT	Progeny 26 GG / GT	Progeny 27 GG / AG	Progeny 28 GG / GT
Progeny 29 GT / GG	Progeny 31 GT	Progeny 32 GT	Progeny 33 GT	Progeny 34 GT	Progeny 35 GT / GG	Progeny 36 GT	Progeny 37 GT / AG	Progeny 38 GT	Progeny 39 AG / GT
Progeny 40 GT	Progeny 41 GT	Progeny 42 GT	Progeny 43 GT / GG	Progeny 44 GG / GT	Progeny 45 GT	Progeny 46 GG / GT	Progeny 47 GG / AG	Progeny 48 GT	Progeny 49 GT / GG
Progeny 50 GT	Progeny 51 GT	Progeny 52 GT / AG	Progeny 53 GG / GT	Progeny 54 GT / GG	Progeny 55 AG / GG	Progeny 56 GT	Progeny 57 AG / GT	Progeny 58 GT / AG	Progeny 59 GG / GT
Progeny 60 GT / GG	Progeny 61 GT / GG	Progeny 62 GT / AG	Progeny 63 GG / AG	Progeny 64 GG / GT	Progeny 65 GT	Progeny 66 GT	Progeny 67 GG / GT	Progeny 68 GT	Progeny 70 GG / AG
Progeny 71 GG / AG	Progeny 72 AG / GG	Progeny 73 GT	Progeny 74 GG / AG	Progeny 75 GT / GG	Progeny 76 GT	Progeny 77 GG / AG	Progeny 78 GG / AG	Progeny 79 GT / GG	Progeny 80 GT
Progeny 81 GT / AG	Progeny 82 GT / AG	Progeny 83 GG / AG	Progeny 84 GT	Progeny 85 GT / GG	Progeny 86 GT / GG	Progeny 87 GT	Progeny 88 GG / AG	Progeny 89 GT	Progeny 90 GG / AG
Progeny 91 AG / GG	Progeny 92 GT / AG	Progeny 93 AG / GG	Progeny 94 GG / AG						

1 (1 tags)

tags per page 10

Id	SNP	Consensus	Matching Parents	Progeny	Marker	Ratio	Genotypes
~ 103 annotate	Yes [2nuc]	TGCAGGAGCCCTCCCACTCGCTGATGGCCACTCCATTCACTGGACCGAGAGCCAAAGCAACACTTCACATTC	2	92 / 91	ab/ac	aa: 25 (27.5%) ab: 24 (26.4%) ac: 18 (19.8%) bc: 24 (26.4%)	91

SNPs

Column: 52; G/A

Column: 70; T/G

Alleles

a : GT

b : GG

c : AG

Matching Samples

View: ☒ Haplotypes ☒ Allele Depths ☐ Genotypes

Male GT / GG 34 / 13	Female AG / GT 12 / 14	Progeny 1 GT 7	Progeny 2 AG / GG 8 / 16	Progeny 3 GG / AG 26 / 14	Progeny 4 GG / GT 15 / 11	Progeny 5 GG / AG 14 / 8	Progeny 6 AG 29	Progeny 7 GT / GG 22 / 11	Progeny 8 AG / GT 12 / 5
Progeny 9 GT 25	Progeny 10 GT 23	Progeny 11 GG / GT 32 / 14	Progeny 12 GT / AG 22 / 7	Progeny 13 GG / AG 7 / 8	Progeny 14 GT / AG 7 / 8	Progeny 15 GT / GG 2 / 3	Progeny 16 GG / GT 19 / 14	Progeny 17 GG / AG 9 / 4	Progeny 18 GT 15
Progeny 19 GT / AG 6 / 3	Progeny 20 AG / GG 6 / 9	Progeny 21 GT / AG 18 / 9	Progeny 22 AG / GT 4 / 5	Progeny 23 GG / AG 7 / 6	Progeny 24 GG / AG 8 / 10	Progeny 25 GT 7	Progeny 26 GG / GT 10 / 16	Progeny 27 GG / AG 3 / 3	Progeny 28 GG / GT 4 / 5
Progeny 29 GT / GG 8 / 5	Progeny 31 GT 11	Progeny 32 GT 10	Progeny 33 GT 17	Progeny 34 GT 20	Progeny 35 GT / GG 7 / 3	Progeny 36 GT 8	Progeny 37 GT / AG 12 / 4	Progeny 38 GT 9	Progeny 39 AG / GT 12 / 7
Progeny 40 GT 9	Progeny 41 GT 5	Progeny 42 GT 9	Progeny 43 GT / GG 9 / 12	Progeny 44 GG / GT 3 / 6	Progeny 45 GT 6	Progeny 46 GG / GT 4 / 11	Progeny 47 GG / AG 3 / 7	Progeny 48 GT 18	Progeny 49 GT / GG 5 / 6
Progeny 50 GT 18	Progeny 51 GT 9	Progeny 52 GT / AG 8 / 5	Progeny 53 GG / GT 10 / 8	Progeny 54 GT / GG 5 / 6	Progeny 55 AG / GG 8 / 10	Progeny 56 GT 22	Progeny 57 AG / GT 17 / 16	Progeny 58 GT / AG 23 / 24	Progeny 59 GG / GT 25 / 13
Progeny 60 GT / GG 12 / 18	Progeny 61 GT / GG 22 / 29	Progeny 62 GT / AG 7 / 23	Progeny 63 GG / AG 15 / 11	Progeny 64 GG / GT 13 / 20	Progeny 65 GT 44	Progeny 66 GT 27	Progeny 67 GG / GT 23 / 17	Progeny 68 GT 30	Progeny 70 GG / AG 14 / 13
Progeny 71 GG / AG 15 / 7	Progeny 72 AG / GG 9 / 6	Progeny 73 GT 42	Progeny 74 GG / AG 31 / 29	Progeny 75 GT / GG 15 / 22	Progeny 76 GT 41	Progeny 77 GG / AG 14 / 17	Progeny 78 GG / AG 25 / 17	Progeny 79 GT / GG 29 / 14	Progeny 80 GT 34
Progeny 81 GT / AG 17 / 29	Progeny 82 GT / AG 29 / 24	Progeny 83 GG / AG 16 / 25	Progeny 84 GT 41	Progeny 85 GT / GG 14 / 24	Progeny 86 GT / GG 6 / 4	Progeny 87 GT 15	Progeny 88 GG / AG 5 / 11	Progeny 89 GT 18	Progeny 90 GG / AG 5 / 17
Progeny 91 AG / GG 14 / 13	Progeny 92 GT / AG 12 / 6	Progeny 93 AG / GG 7 / 7	Progeny 94 GG / AG 3 / 2						

1 (1 tags)

tags per page 10

1 (1 tags)

tags per page 10

Id	SNP	Consensus	Matching Parents	Progeny	Marker	Ratio	Genotypes
~ 103 annotate	Yes [2nuc]	TGCAGGAGCCCTCCCACTCGCTGATGGCCACTCCATTTCAGTGGACCGAGAGC C AAAGCAACACTTCACA T TCCC	2	92 / 91	ab/ac	aa: 25 (27.5%) ab: 24 (26.4%) ac: 18 (19.8%) bc: 24 (26.4%)	91

SNPs

Alleles

Matching Samples

Column: 52; G/A
Column: 70; T/G

a : GT
b : GG
c : AG

View: ☒ Haplotypes ☒ Allele Depths ☒ Genotypes

Male GT / GG 34 / 13 aa	Female AG / GT 12 / 14 aa	Progeny 1 GT 7 aa	Progeny 2 AG / GG 8 / 16 bc	Progeny 3 GG / AG 26 / 14 bc	Progeny 4 GG / GT 15 / 11 ab	Progeny 5 GG / AG 14 / 8 bc	Progeny 6 AG 29 AC	Progeny 7 GT / GG 22 / 11 ab	Progeny 8 AG / GT 12 / 5 ac
Progeny 9 GT 25 aa	Progeny 10 GT 23 aa	Progeny 11 GG / GT 32 / 14 ab	Progeny 12 GT / AG 22 / 7 ac	Progeny 13 GG / AG 7 / 8 bc	Progeny 14 GT / AG 7 / 8 ac	Progeny 15 GT / GG 2 / 3 ab	Progeny 16 GG / GT 19 / 14 ab	Progeny 17 GG / AG 9 / 4 bc	Progeny 18 GT 15 aa
Progeny 19 GT / AG 6 / 3 ac	Progeny 20 AG / GG 6 / 9 bc	Progeny 21 GT / AG 18 / 9 ac	Progeny 22 AG / GT 4 / 5 ac	Progeny 23 GG / AG 7 / 6 bc	Progeny 24 GG / AG 8 / 10 bc	Progeny 25 GT 7 AC	Progeny 26 GG / GT 10 / 16 ab	Progeny 27 GG / AG 3 / 3 bc	Progeny 28 GG / GT 4 / 5 ab
Progeny 29 GT / GG 8 / 5 ab	Progeny 31 GT 11 aa	Progeny 32 GT 10 aa	Progeny 33 GT 17 aa	Progeny 34 GT 20 aa	Progeny 35 GT / GG 7 / 3 ab	Progeny 36 GT 8 aa	Progeny 37 GT / AG 12 / 4 ac	Progeny 38 GT 9 aa	Progeny 39 AG / GT 12 / 7 ac
Progeny 40 GT 9 aa	Progeny 41 GT 5 aa	Progeny 42 GT 9 aa	Progeny 43 GT / GG 9 / 12 ab	Progeny 44 GG / GT 3 / 6 ab	Progeny 45 GT 6 AC	Progeny 46 GG / GT 4 / 11 ab	Progeny 47 GG / AG 3 / 7 bc	Progeny 48 GT 18 aa	Progeny 49 GT / GG 5 / 6 ab
Progeny 50 GT 18 aa	Progeny 51 GT 9 aa	Progeny 52 GT / AG 8 / 5 ac	Progeny 53 GG / GT 10 / 8 ab	Progeny 54 GT / GG 5 / 6 ab	Progeny 55 AG / GG 8 / 10 bc	Progeny 56 GT 22 aa	Progeny 57 AG / GT 17 / 16 ac	Progeny 58 GT / AG 23 / 24 ac	Progeny 59 GG / GT 25 / 13 ab
Progeny 60 GT / GG 12 / 18 ab	Progeny 61 GT / GG 22 / 29 ab	Progeny 62 GT / AG 7 / 23 ac	Progeny 63 GG / AG 15 / 11 bc	Progeny 64 GG / GT 13 / 20 ab	Progeny 65 GT 44 aa	Progeny 66 GT 27 aa	Progeny 67 GG / GT 23 / 17 ab	Progeny 68 GT 30 aa	Progeny 70 GG / AG 14 / 13 bc
Progeny 71 GG / AG 15 / 7 bc	Progeny 72 AG / GG 9 / 6 bc	Progeny 73 GT 42 aa	Progeny 74 GG / AG 31 / 29 bc	Progeny 75 GT / GG 15 / 22 ab	Progeny 76 GT 41 aa	Progeny 77 GG / AG 14 / 17 bc	Progeny 78 GG / AG 25 / 17 bc	Progeny 79 GT / GG 29 / 14 ab	Progeny 80 GT 34 aa
Progeny 81 GT / AG 17 / 20 aa	Progeny 82 GT / AG 28 / 24 aa	Progeny 83 GG / AG 16 / 25 aa	Progeny 84 GT 41 aa	Progeny 85 GT / GG 14 / 24 aa	Progeny 86 GT / GG 6 / 4 aa	Progeny 87 GT 15 aa	Progeny 88 GG / AG 5 / 11 aa	Progeny 89 GT 18 aa	Progeny 90 GG / AG 5 / 17 aa

version 0.998

RAD-Tag Sample #2 [female]

Catalog ID	Depth	SNPs		Alleles		Deleveraged?	Lumberjackstack?	Blacklisted?
#103	26x	Column: 52	G/A	AG	46.15%	False	False	False
		Column: 70	T/G	GT	53.85%			

last updated: Sun Jan 8 08:59:35 PST 2012

Screen

- When you disconnect from a remote session all jobs are killed
- Screen allows you to connect/reconnect from any machine
 - type `screen` to start a new screen
 - `screen -ls` lists all screens you have running
 - `ctrl-a` sends commands to screen instead of the terminal
 - `ctrl-a ctrl-d` disconnects
 - type `exit` to destroy a screen