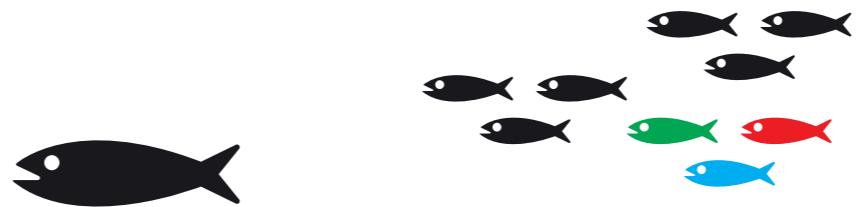
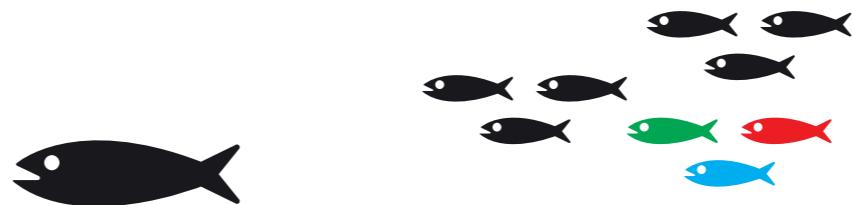


Introductory Bioinformatic Course to Sequencing Data Processing

Marker development

GBS/RAD-seq





↓
Chr. I ||| ||| ||| ||| ||| |||
Chr. II ||| ||| ||| ||| ||| |||

↓ + restriction enzyme (SbfI)

TGCGG TGATGACTGGTACACATGTTATGACGAT...

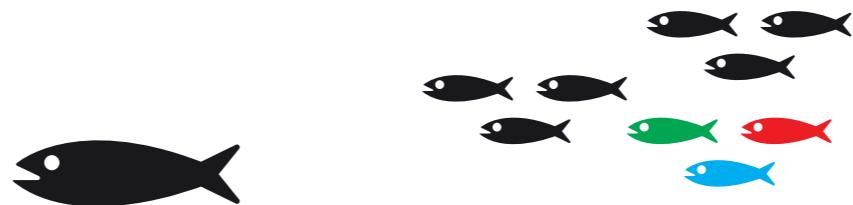
SbfI flanking sequence

↓ + barcode (one per individual)



AATTT TGCGG TGATGACTGGTACACATGTTATGACGAT...

barcode SbfI sequence



Chr. I

Chr. II

↓ + restriction enzyme (SbfI)

TGCGG TGATGACTGGTACACATGTTATGACGAT...

SbfI flanking sequence

↓ + barcode (one per individual)



AATTT TGCGG TGATGACTGGTACACATGTTATGACGAT...

barcode SbfI sequence

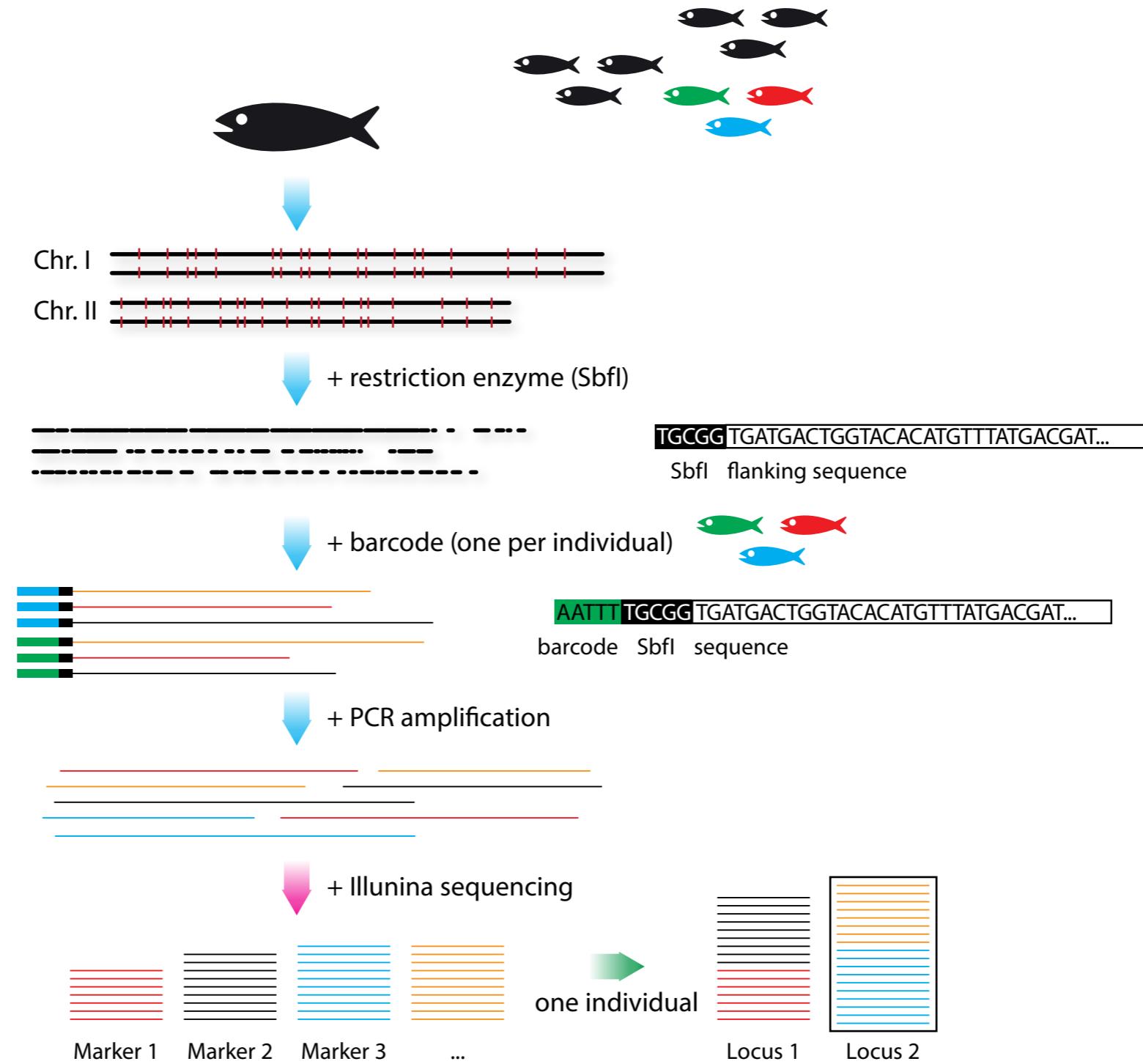
↓ + PCR amplification

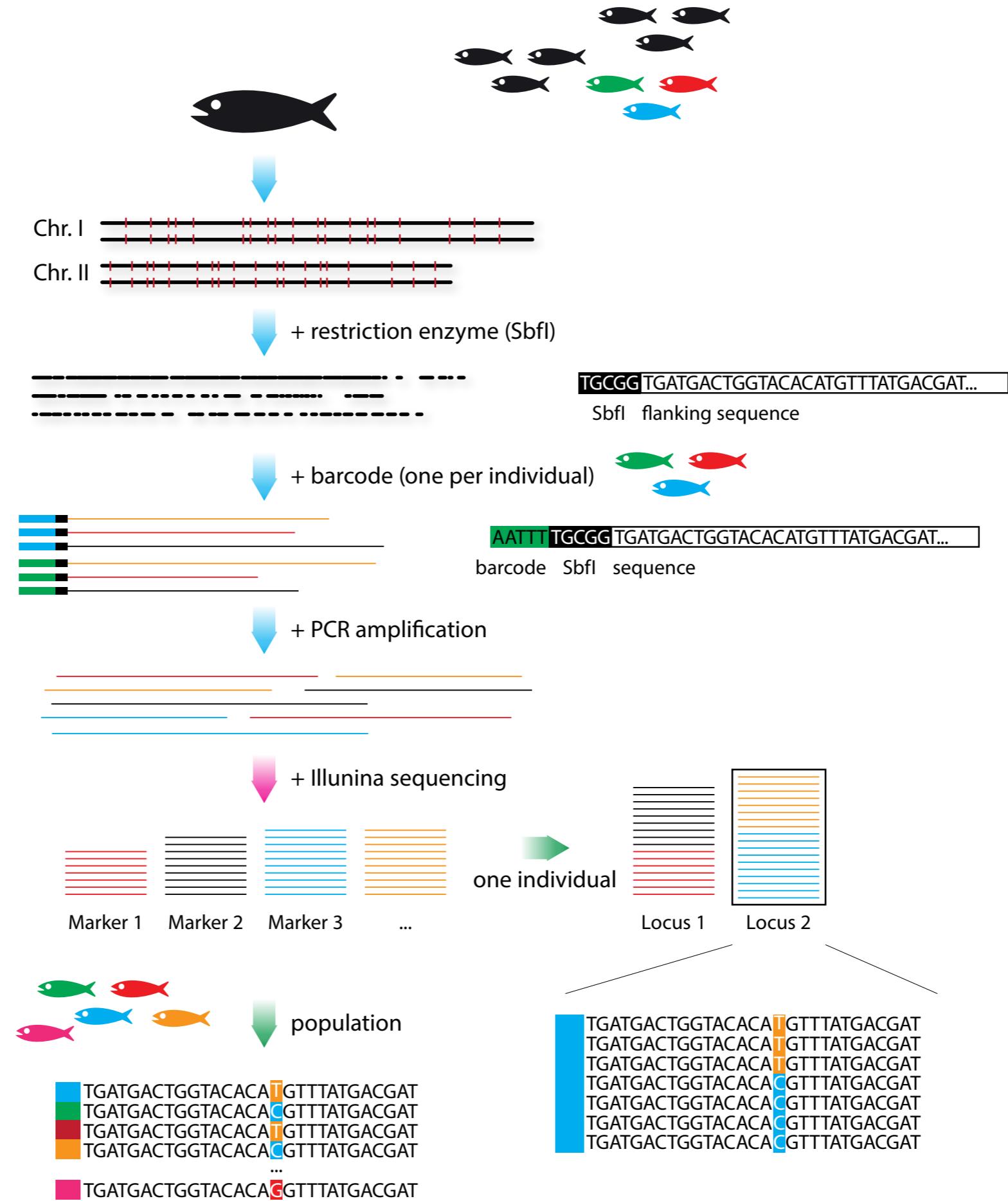
.....

.....

.....

.....







(c) Kate Hobbs Photography

RAD-tag sequencing

- Base on **high-throughput sequencing** technology. (Baird et al., PLoS One, 2008)
- Powerful technique to generate **markers**, **linkage maps**, **SNP discovery** and **QTL analysis**.
- Can generate linkage maps with **thousands of SNP** markers in a **very short time** and **fairly cheaply**:
 - ◆ >300 M reads of 150 bases for less than **€1,000**
 - ◆ generates **~100,000** (up to 1.2 M) loci, **~20,000 informative SNPs**
- RAD analysis can be done in species with **no prior genomic information**.
- Good first step in analysis of sex determination in a **"new" species** , or **improved** with **very high density genetic map** for MAS and GAS.



© Reynir SkarsgArd

E.g.: Sex-determination in aquaculture

- Fish display an enormous diversity in sex differentiation and determination.
- As aquaculture has developed, it has become apparent that for many species, **control over sex ratio** (and **age** at maturation) is desirable.
- The list of species includes:
 - ◆ **Salmonids** (females mature later, female triploids sterile)
 - ◆ **Tilapias** (all-male culture prevents reproduction before harvest; males also grow faster)
 - ◆ **Flatfish**, halibut (females mature later and grow larger)
 - ◆ **Grass carp** (used for aquatic plant control in USA; triploids used to prevent reproduction)

Monosex male: Tilapia

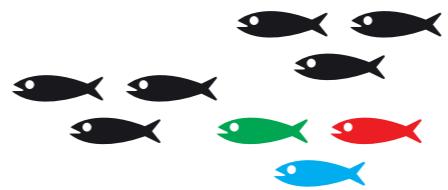
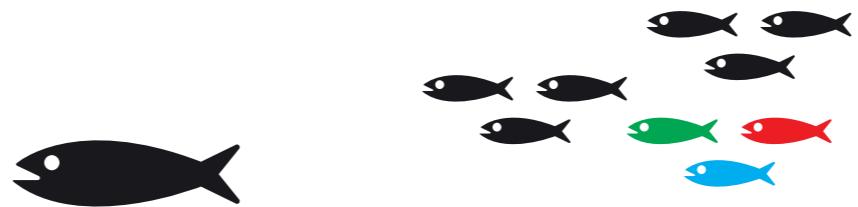
- Tilapias rank **second** in global aquaculture
- Production dominated by **Nile tilapia** (*Oreochromis niloticus*), but many other species
- **Mixed-sex** culture generally results in maturation and **unwanted reproduction** before harvest
- All male culture prevents this
- Techniques have included interspecies **hybrids**, direct **hormonal treatment** (dominates industry), **YY x XX crosses** (smaller commercial use), selection for increased **thermosensitivity** in sexual differentiation (experimental).



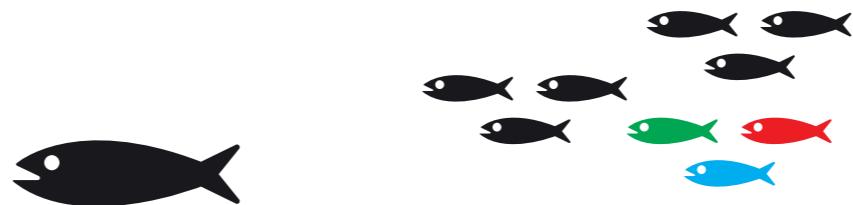
mixed sex



monosex male



Chr. I Two horizontal black lines with red vertical markers.
Chr. II Two horizontal black lines with red vertical markers.



↓
Chr. I ||| ||| ||| ||| ||| |||
Chr. II ||| ||| ||| ||| ||| |||

↓ + restriction enzyme (SbfI)

TGCGG TGATGACTGGTACACATGTTATGACGAT...

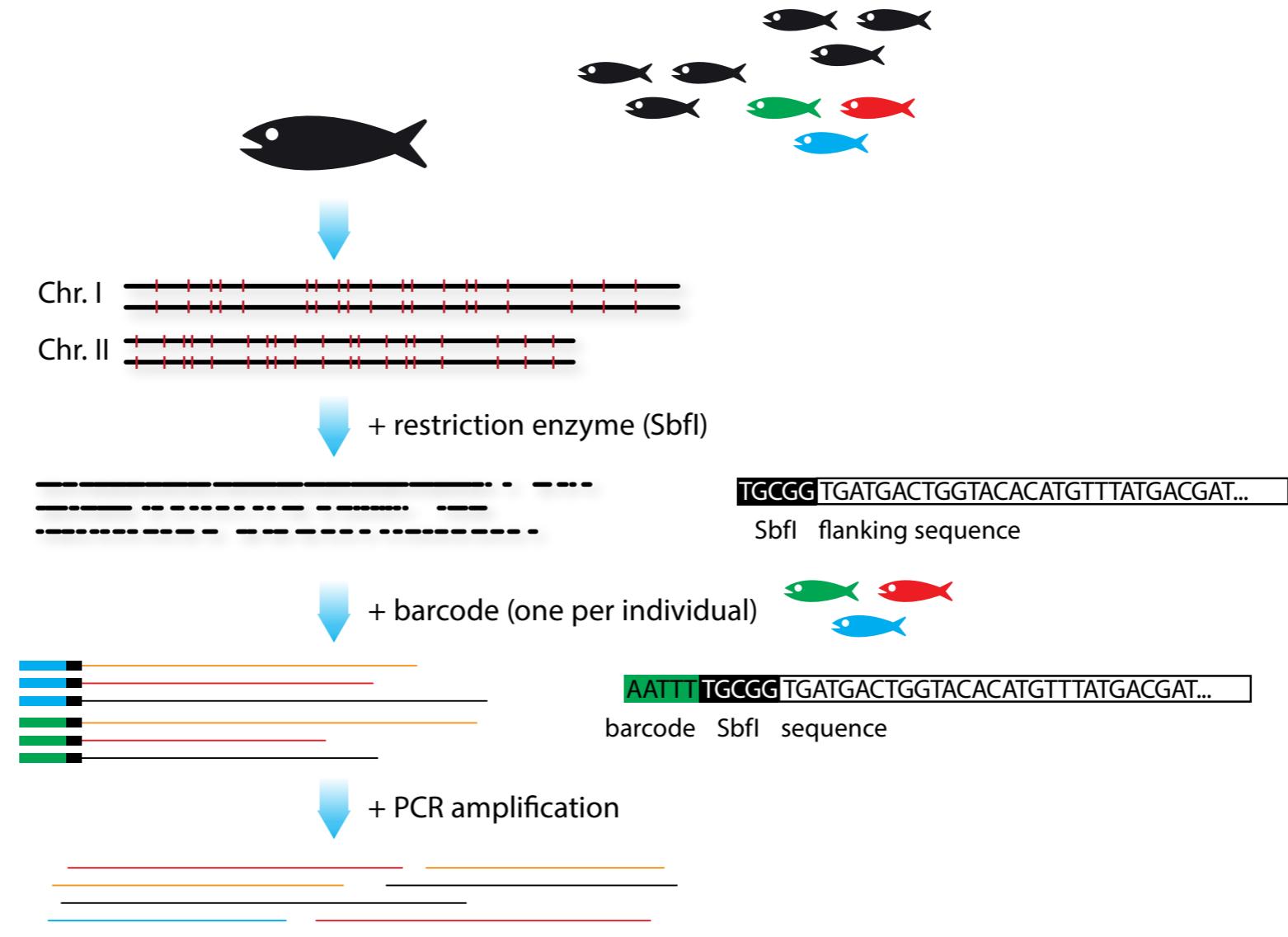
SbfI flanking sequence

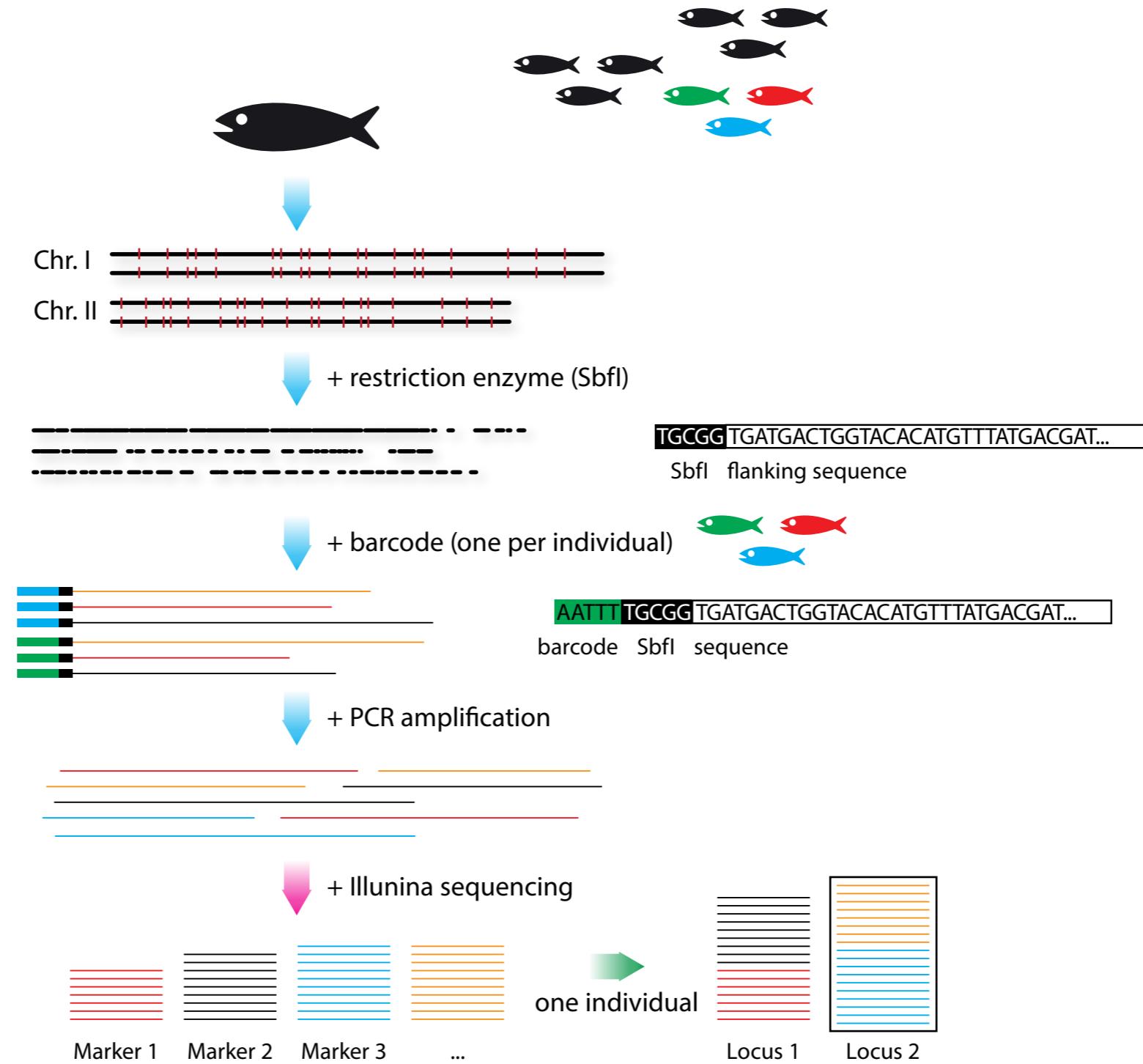
↓ + barcode (one per individual)

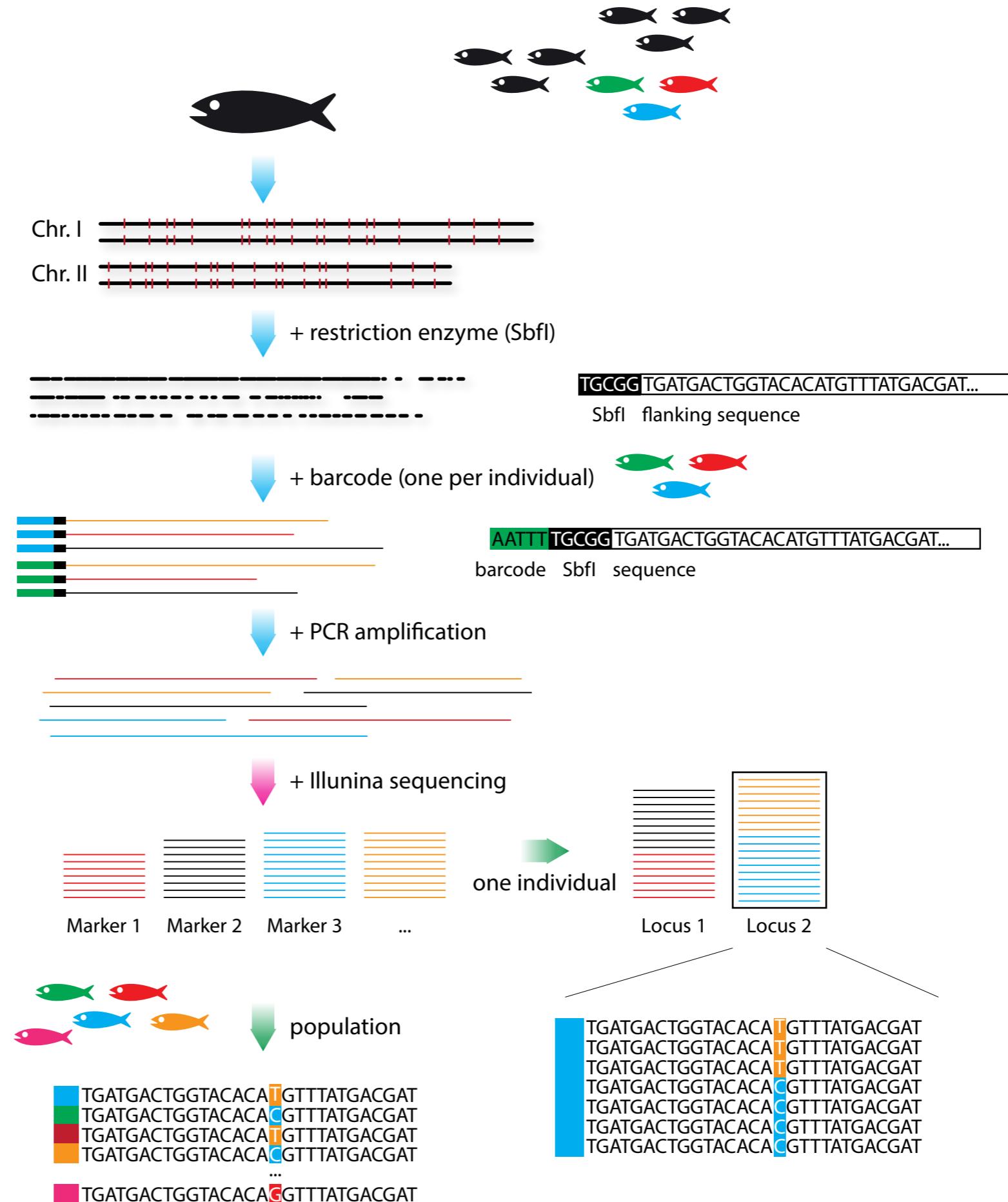


AATTT TGCGG TGATGACTGGTACACATGTTATGACGAT...

barcode SbfI sequence







Notation

Read/Sequence

A raw sequence that comes from a sequencing machine.

TAGGTCAGATAACCGTAGATA

Tag = Marker or RAD-marker or RAD-Tag

Several reads having the same sequence can be merged to one tag.

TAGGTCAGATAACCGTAGATA

TAGGTCAGATAACCGTAGATA

TAGGTCAGATAACCGTAGATA

Locus

A unique position/location on the genome where Reads were generated, and Tags align to.



High-throughput sequences



© Calgary Zoo

Nile tilapia

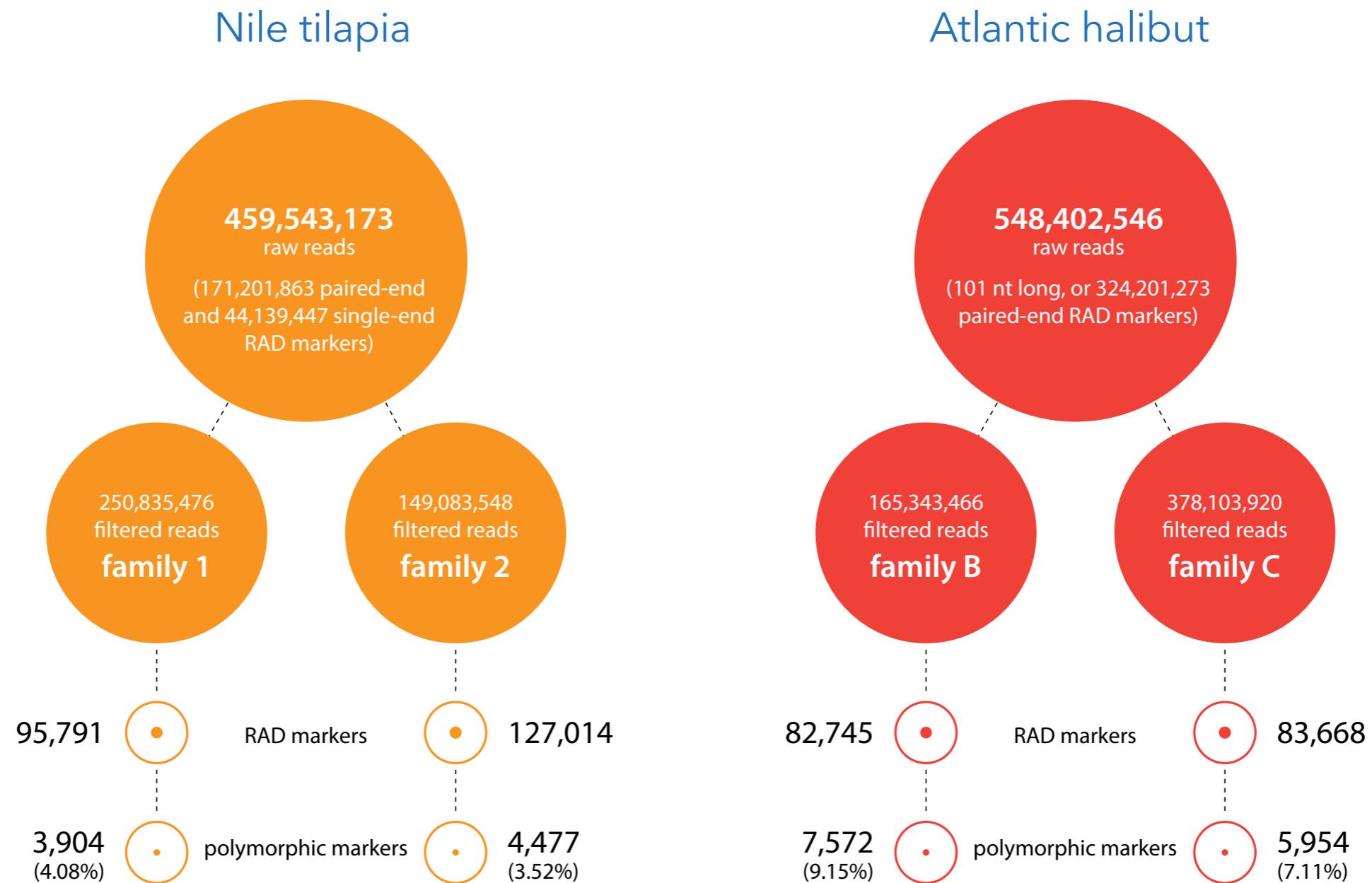
- Monosex **males**
- XX/XY “plus”
- LG1?
- Draft genome sequence (female)
- Clonal “**isogenic**” line **female** x XY male mapping families (n = 68, 20)



Atlantic halibut

- Monosex **females**
- XX/XY (confirmed)
- Unknown LG
- **No genome** sequence
- **Wild** female x **wild** male mapping families (n = 70, 30)

High-throughput sequences

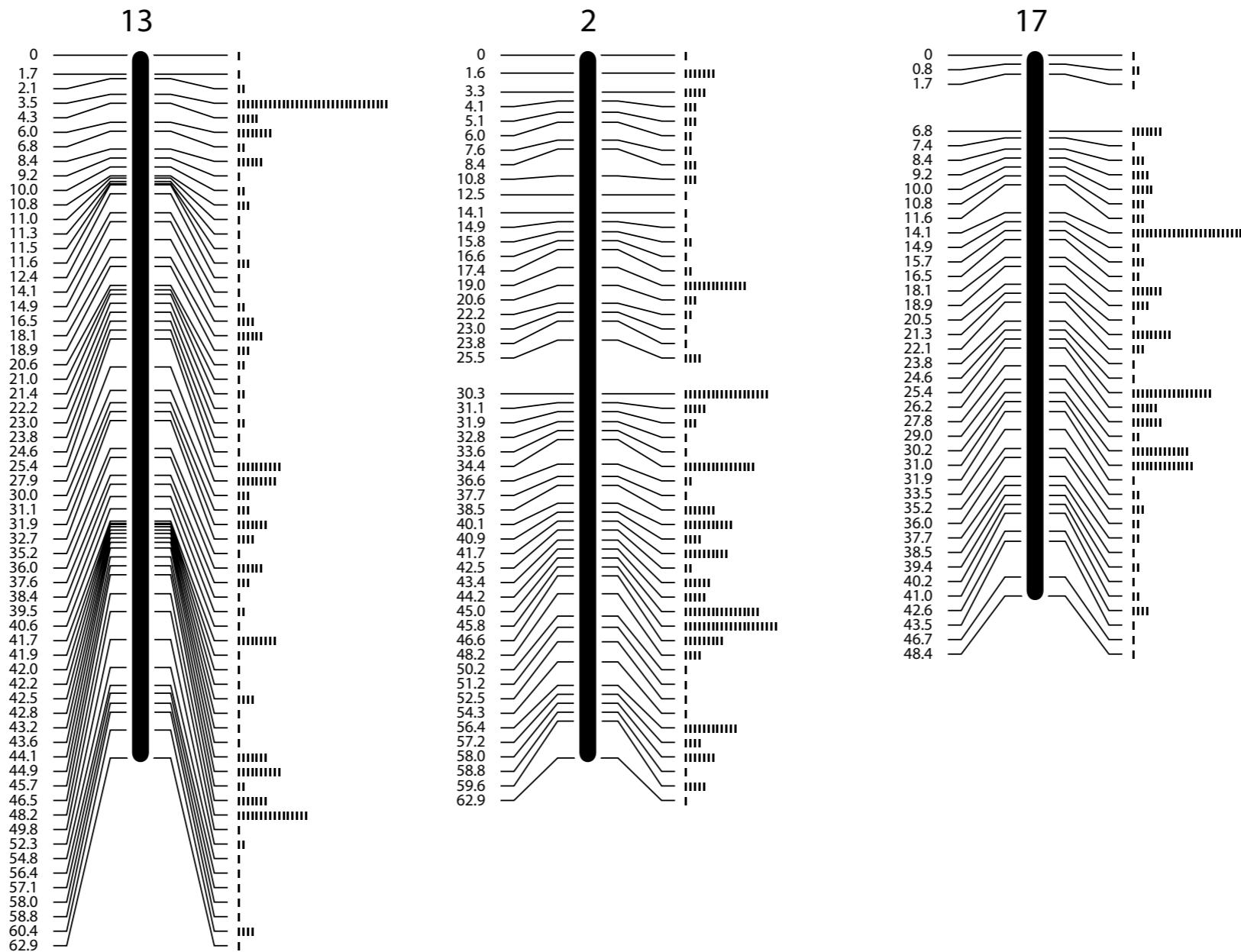


Genetic maps: Tilapia

459,543,173 sequences
4,477 polymorphic markers

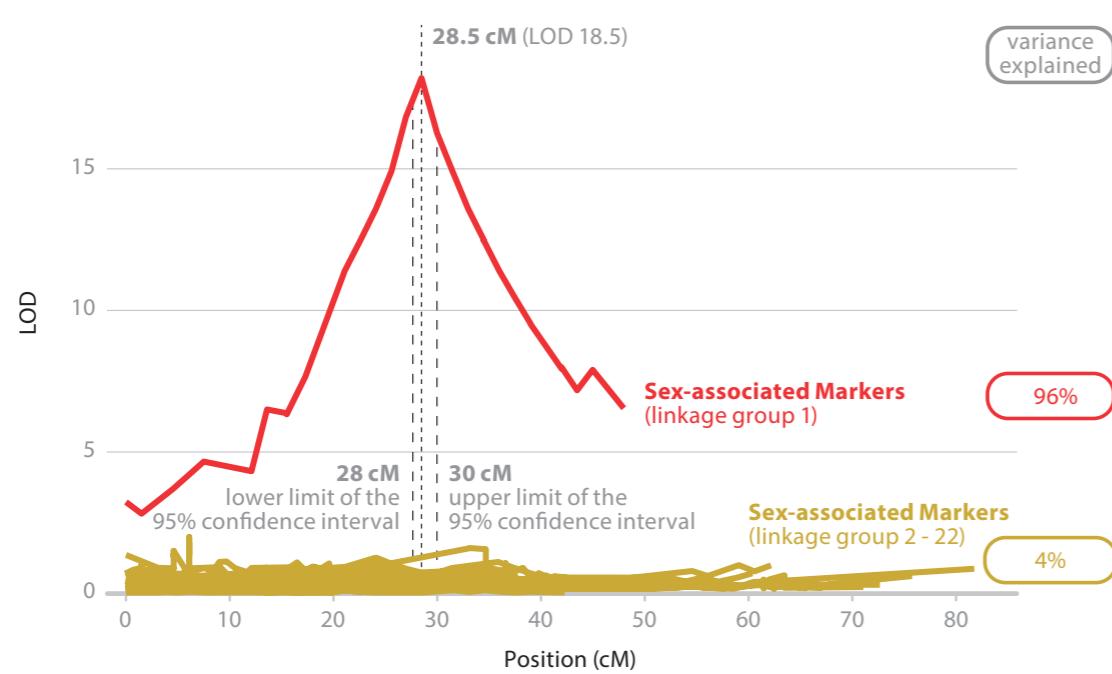
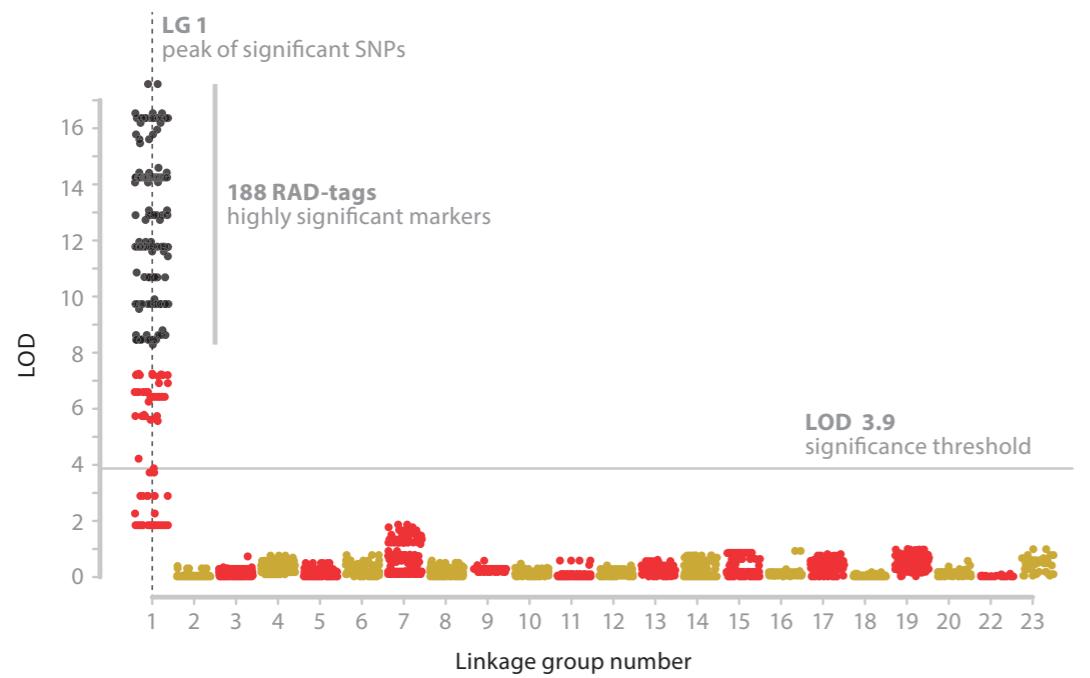


22 linkage groups
(chromosomes)

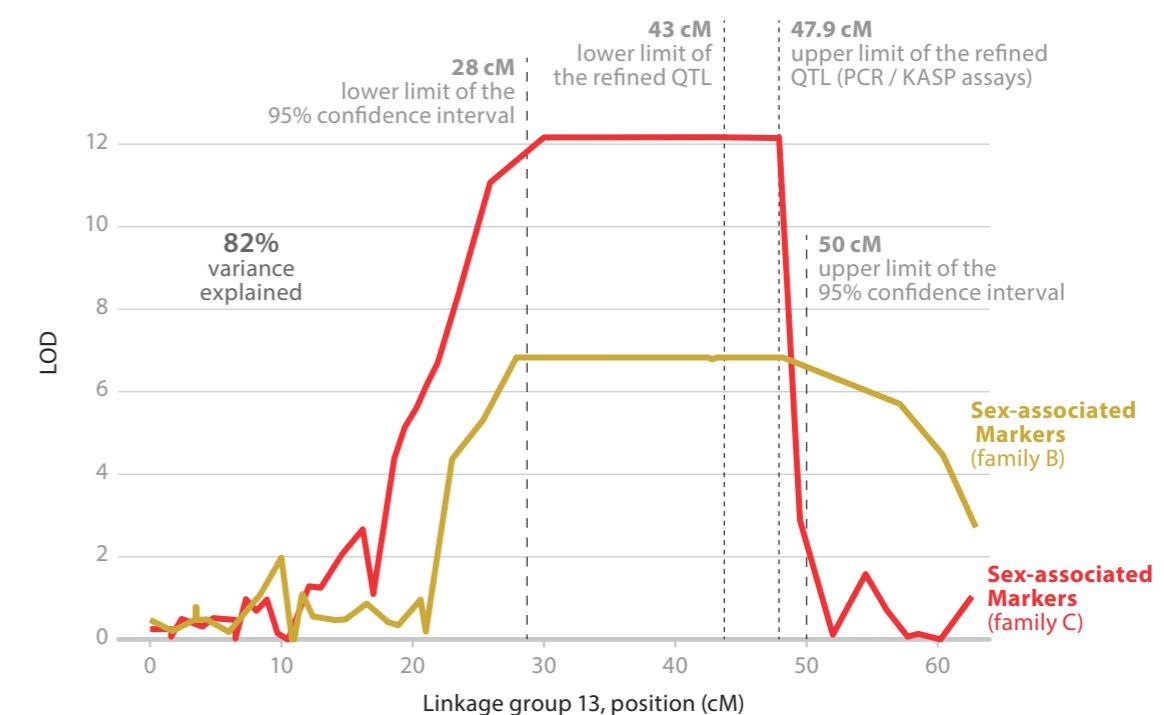
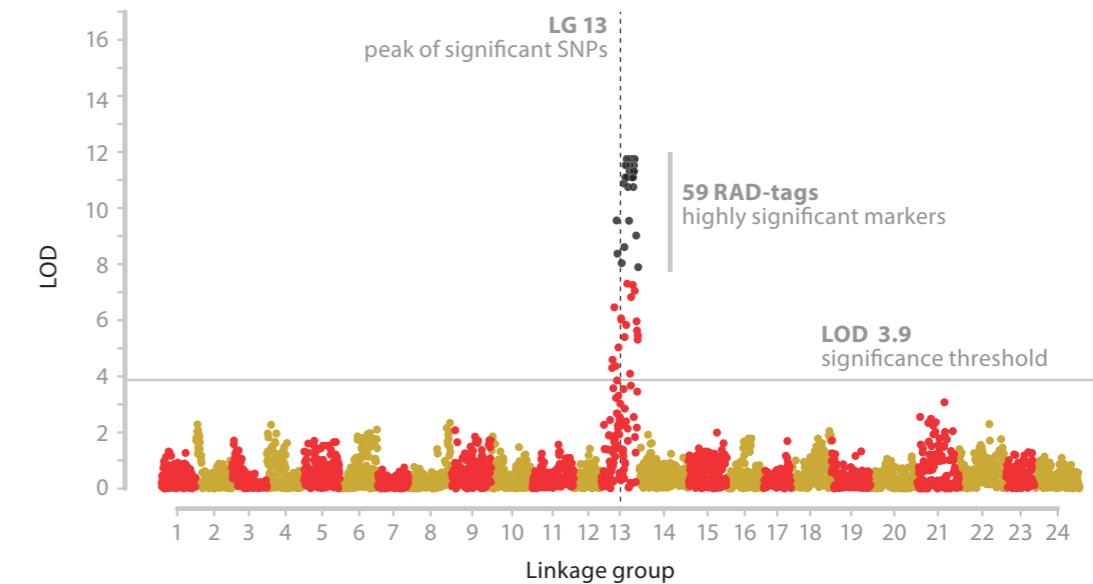


Sex-determining genes: QTL analysis

Nile tilapia



Atlantic halibut



Reminder

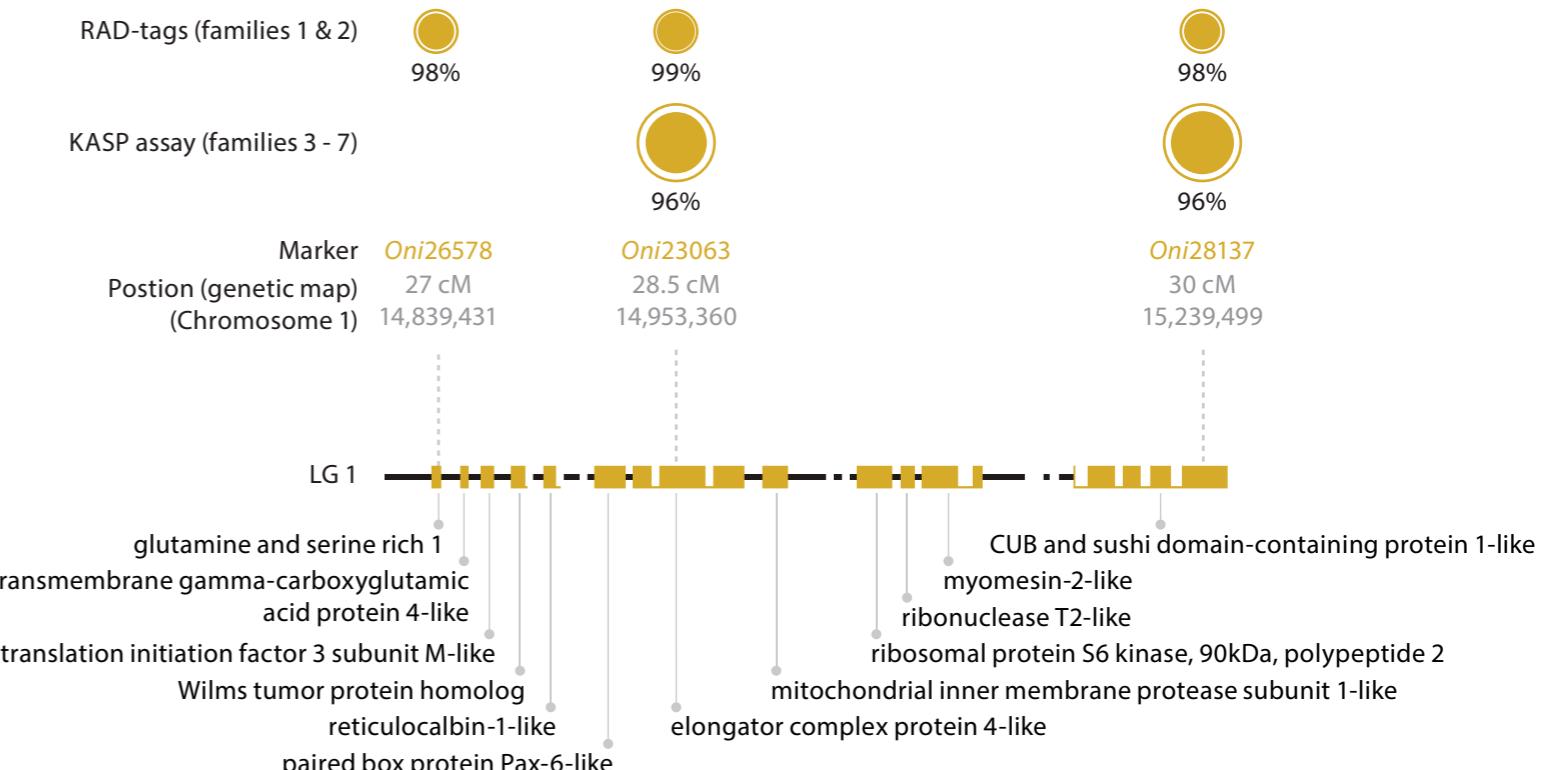
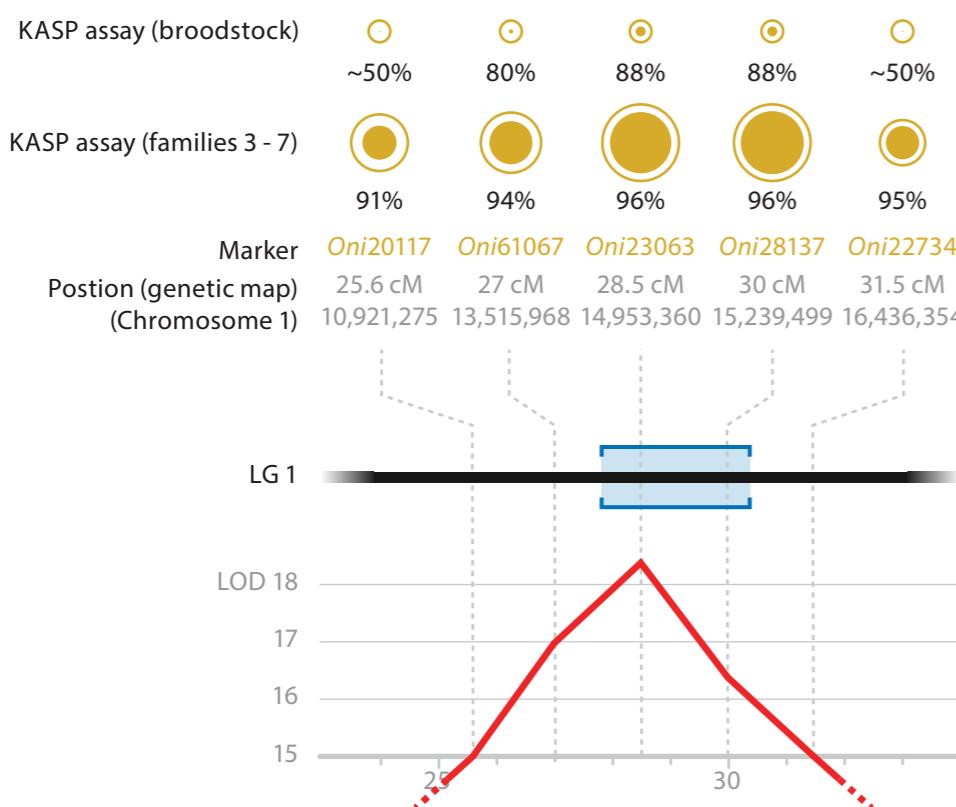
A Quantitative Trait Locus (QTL) is a region in a chromosome, that was identified (by sequencing). And a sequence change in this region (often a polymorphism) has been statistically associated with a phenotype.

A QTL is a marker associated with a phenotype. This is **not** a Gene. This is a fragment of DNA close to a functional region involved in the phenotype.

Sex-determining genes: Fine mapping

Nile tilapia

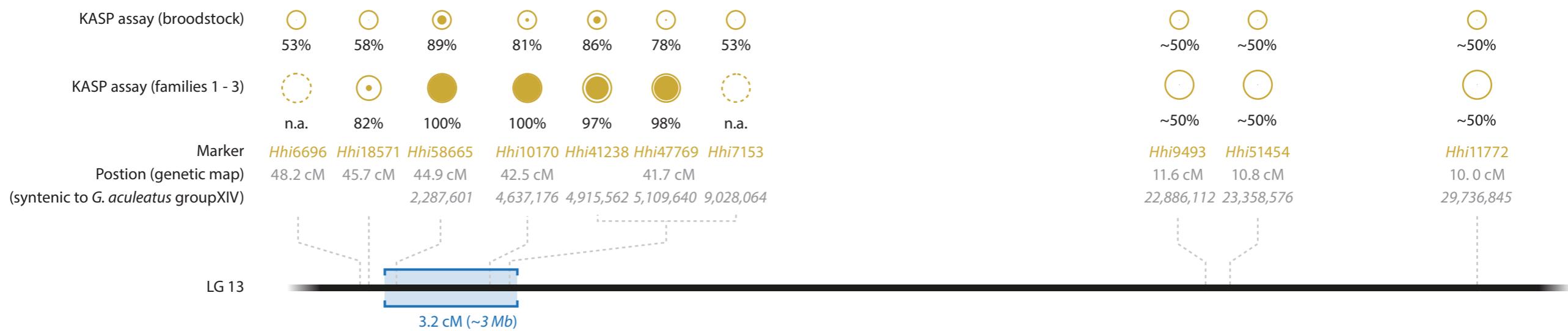
Markers > 96% association
440 kb region



Sex-determining genes: Fine mapping

Atlantic halibut

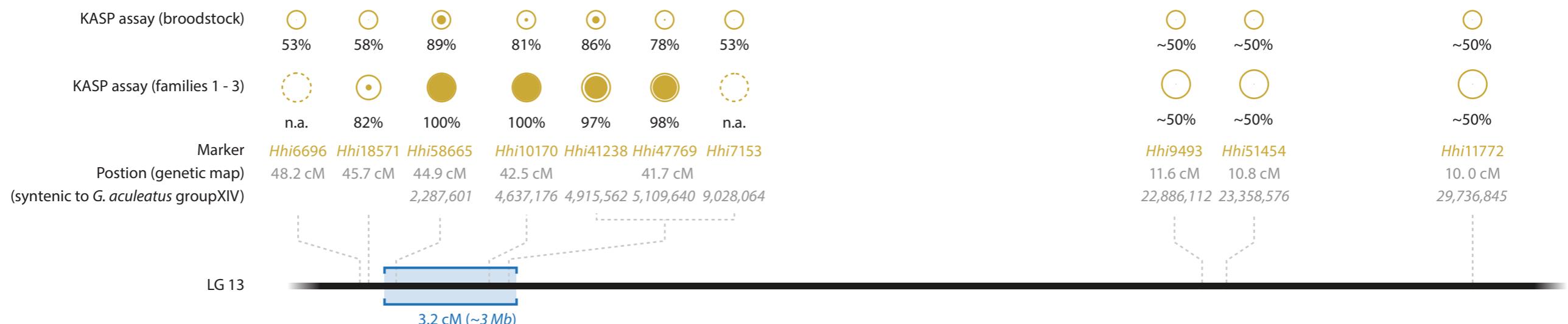
Markers ~89% association
~3 Mb region



Sex-determining genes: Fine mapping

Atlantic halibut

Combined markers > 97% prediction
~3 Mb region



Predicted

M F

if [*Hhi58665*] is C/G (heterozygote) and [*Hhi41238*] is G/T (heterozygote) then Male

Observed	M	F
M	43	3
F	0	47

if [*Hhi58665*] is C/G (heterozygote) and [*Hhi41238*] is G/G (homozygote) and [*Hhi10170*] is G/G (homozygote) then Male
all other alleles Female

Sex-determining genes

- Space for improvement
- For some species with simple (XX/XY) sex determination systems, monosex female culture was developed fairly quickly and has been very successful (e.g. rainbow trout).
- Success in aquaculture depends on need, cost, difficulty.
- For complex/new sex determination systems, MAS may help but GAS would be better for developing sex control!

RAD experimental design and wet lab considerations

- Families are better than populations (map genotypes, build genetic map)
- Paired-ends provide extra information in case of gene duplication, repetitive regions and can be post-processed for extra SNPs
- Coverage > 50 reads per RAD-tag (to be able to see all alleles with confidence)
- Number of samples and the restriction enzyme used are key!

Tutorial