Recent Advances in Analysing Sex Determination in Fish

David Penman

Institute of Aquaculture

University of Stirling

Scotland, UK

UNIVERSITY OF
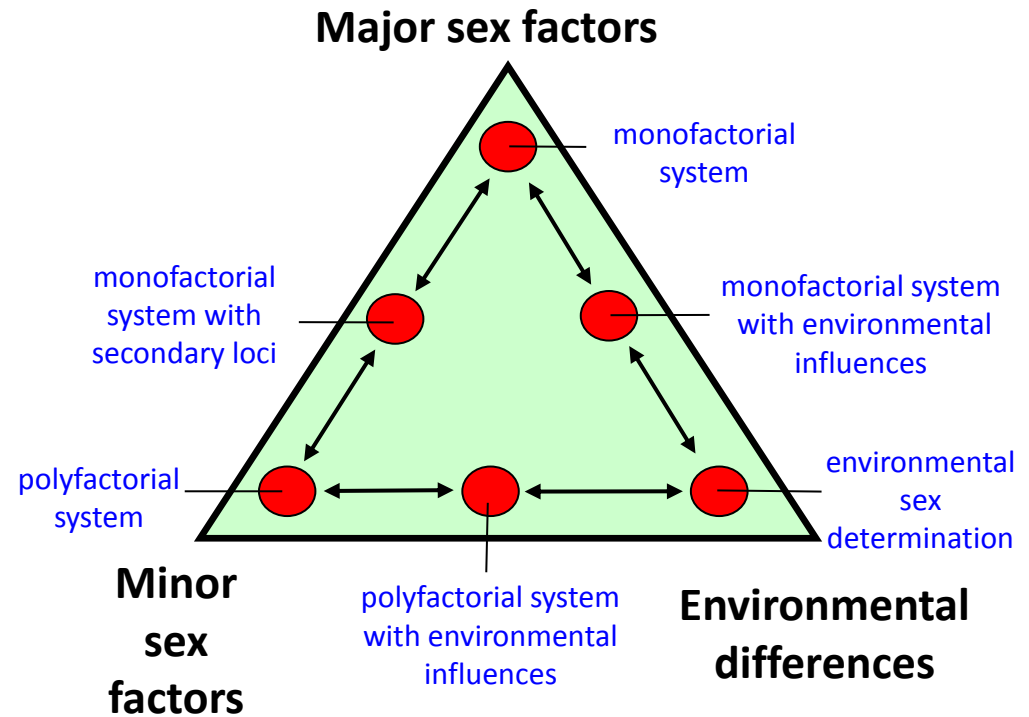STIRLING

Acknowledgements

Background

- As aquaculture has developed, it has become apparent that for many (but not all) species, control over sex ratio is desirable.

- Fish (the vertebrates that didn't leave the water) display an enormous diversity in sex differentiation and determination (so do crustaceans!)

Hermaphroditic   **Gonochoristic**

Protandrous
Protogynous
Simultaneous

Major sex factors

monofactorial system

monofactorial system with secondary loci

monofactorial system with environmental influences

polyfactorial system

environmental sex determination

Minor sex factors

polyfactorial system with environmental influences

Environmental differences

Penman and Piferrer (2008)

Where do (did) we start?

Sex reversal

Progeny testing

Environmental effects on sex ratio

Gynogenesis (Androgenesis)

Hybridization

Sex-associated markers

Karyotype

Where do (did) we start?

Sex reversal

Progeny testing

Environmental effects on sex ratio

Gynogenesis (Androgenesis)

Hybridization

Sex-associated markers

Karyotype

Very few species of aquaculture interest show differentiated sex chromosomes – examples in other fish species

Power of DNA marker/genomics techniques has greatly increased over last decade

NGS and GBS

- Some (10?) years ago, a clear starting point for the analysis of sex determination in a "new" species was not easy

- In particular, molecular methods were still very laborious and expensive, and not always very productive regarding sex determination (some success)

- For example, generating enough microsatellite markers to produce a moderate (low!) density linkage map (200-500) and then genotyping 100 progeny (one family) was a very significant task

- NGS has made huge differences to the time/cost and number of markers involved in analysing sex determination (and many other traits)

- GBS (genotyping by sequencing) is becoming key in this – several methods

NGS and GBS

- Some (10?) years ago, a clear starting point for the analysis of sex determination in a "new" species was not easy

- In particular, molecular methods were still very laborious and expensive, and not always very productive regarding sex determination (some success)

- For example, generating enough microsatellite markers to produce a moderate (low!) density linkage map (200-500) and then genotyping 100 progeny (one family) was a very significant task

- NGS has made huge differences to the time/cost and number of markers involved in analysing sex determination (and many other traits)

- GBS (genotyping by sequencing) is becoming key in this – several methods

(Only 3 presentations featuring RADseq at ISGA XI – how many at ISGA XII?)

RADseq

- Baird et al. (2008) and subsequent variations
- Powerful technique to generate and genotype markers, linkage maps and analyse sex determination
- Can generate linkage maps with thousands of SNP markers in a very short time and fairly cheaply (~300 M reads of 100 bases for a few thousand pounds/dollars/euros -> ~60-90 k loci -> ~4-8 k informative SNPs)
- RADseq (and ddRADseq) analysis can be done in species with no prior genomic information and often takes considerably less time than producing the biological material
- Good first step in analysis of sex determination in a "new" species -> XX/XY, WZ/ZZ or something else?
- We are currently working on species including Atlantic halibut, Nile tilapia and European sea bass, also hapuku, arapaima

Standard RADseq

N = 100   *Sbf*I   3500-6500 SNPs
(barcodes)   (e.g. per family)
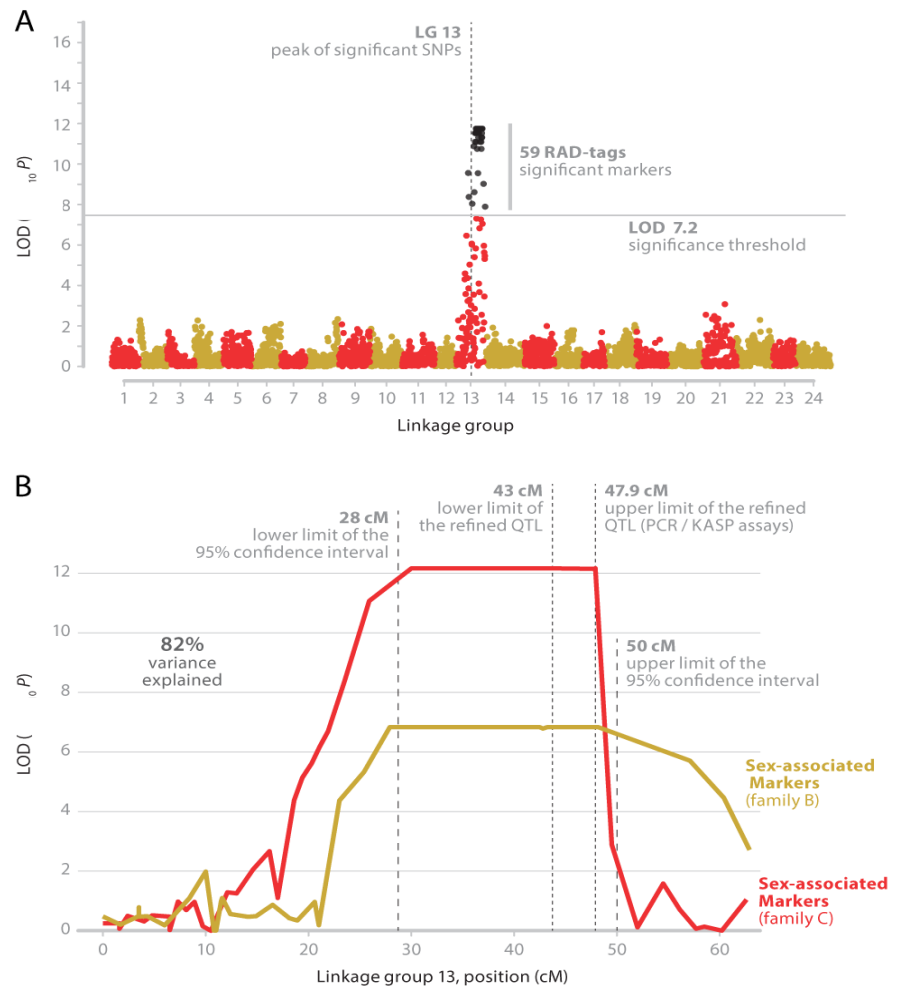2 lanes HiSeq   <€12 k ~ 2 weeks seq.

e.g. repeat analysis with different enzyme

Standard RADseq

ddRADseq

BSA-ddRADseq?

N = 100   *Sbf*I , ??   >>10,000 SNPs
(barcodes)   (e.g. per family)
4 lanes HiSeq   <€24 k   ~ 4 weeks seq.

N = 100   *Sbf*I   3500-6500 SNPs
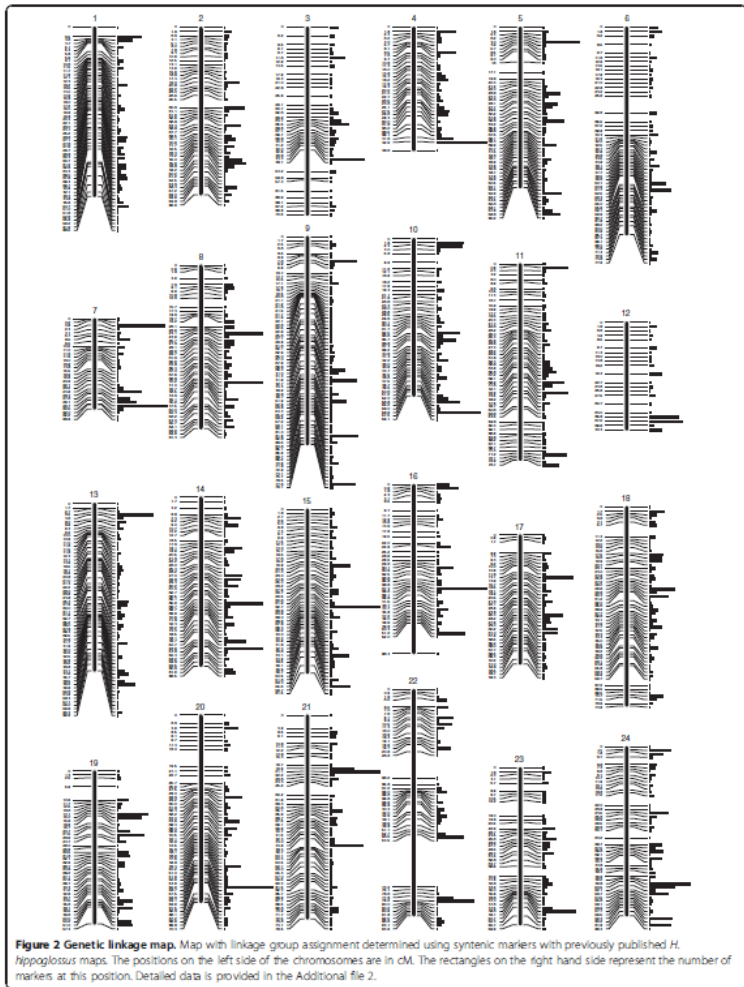(barcodes)   (e.g. per family)
2 lanes HiSeq   <€12 k ~ 2 weeks seq.

N = 100+ e.g.*Sbf*I+*Sph*I 500-1500 SNPs
(coverage variance)   (e.g. per family)
2 lanes MiSeq   <€2.5 k   2 days seq.

N = 25 families *Sbf*I/*Sph*I 500-1500 SNPs
(sire, dam, male and female progeny pools)
2 lanes MiSeq   <€2.5 k   2 days seq.

Atlantic halibut – XX/XY (no genome sequence)

Figure 2 Genetic linkage map. Map with linkage group assignment determined using syntenic markers with previously published *H. hippoglossus* maps. The positions on the left side of the chromosomes are in cM. The rectangles on the right hand side represent the number of markers at this position. Detailed data is provided in the Additional file 2.

5764 SNPs in 24 linkage groups (1514 cM); sex determining locus in LG13

Palaiokostas et al. (2013a)

Sex association verification and sex prediction – halibut

Tightly sex-linked SNPs can be used to predict phenotypic sex in broodstock – now being applied in monosex production (testing of potential neomales)
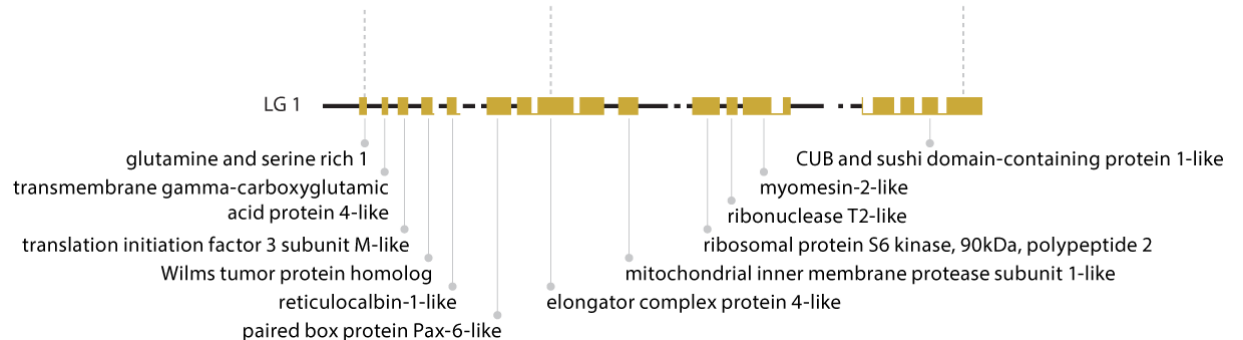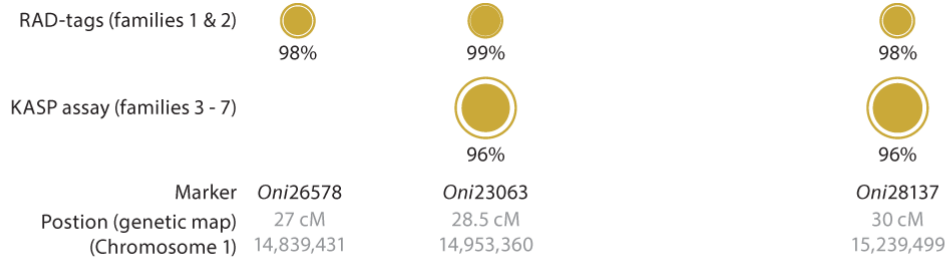
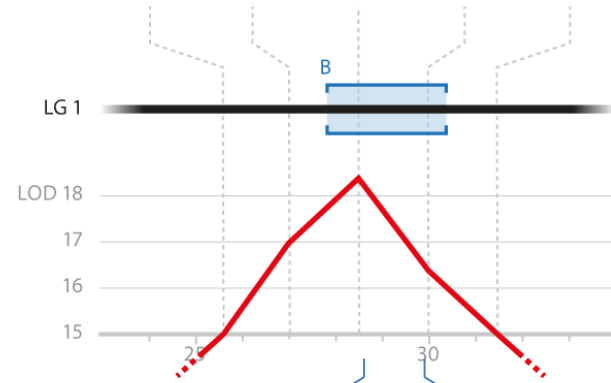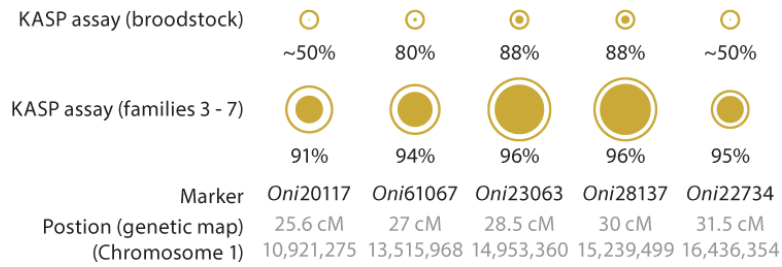# Nile tilapia - oligogenic sex determination?

Cross with balanced sex ratio:

- Male map only (isogenic XX female used)
- 3802 SNP markers
- 1176 cM
- XX/XY locus in LG1*
- Draft genome assembly available

*but in LG23 in some strains!

Palaiokostas et al. (2013 b)
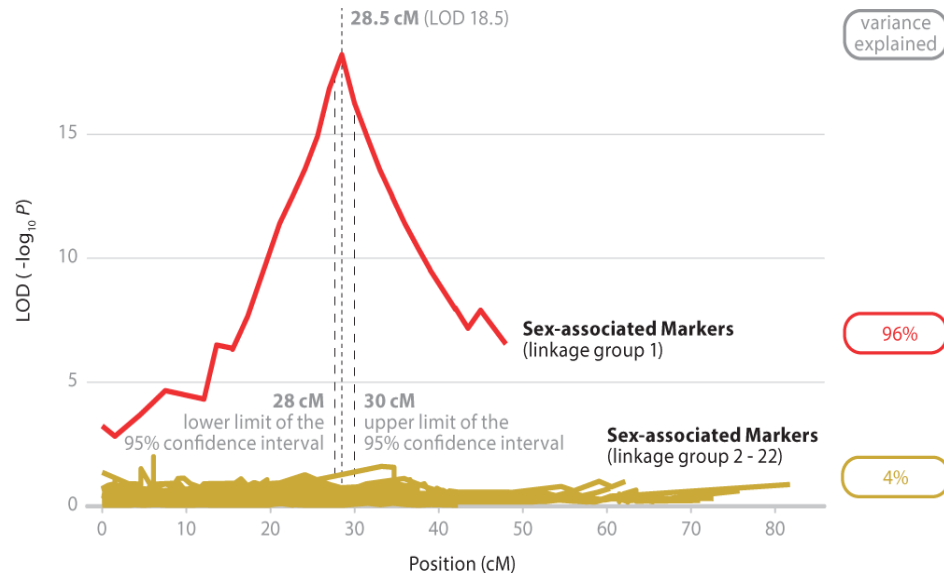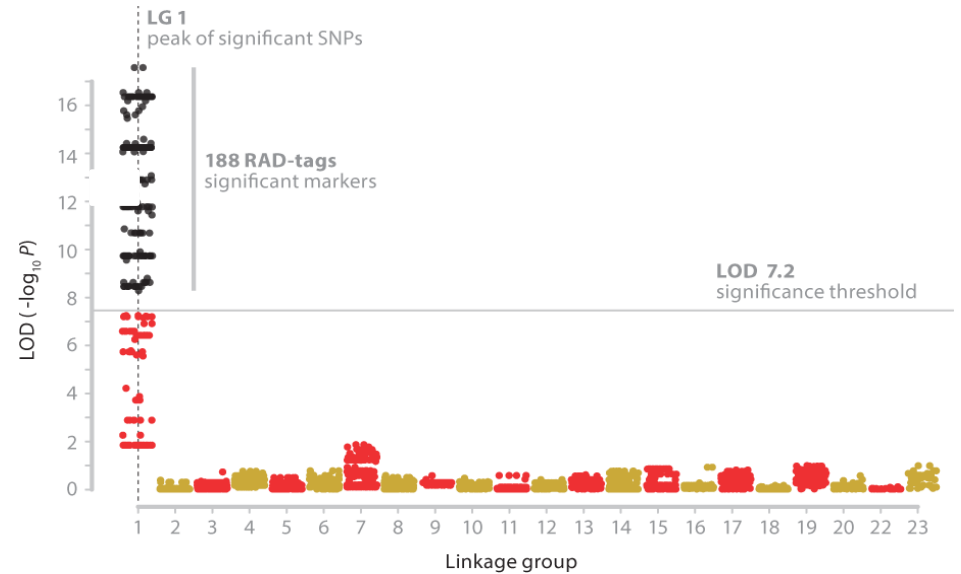
Verification of LG1 sex association in Nile tilapia

Comparison to draft genome – but SD gene (Y) may not be present in XX genome!

Nile tilapia – skewed sex ratios and temperature effects

- ddRADseq (mod. from Peterson et al. 2012) used in this study (641 SNPs)
- Upper part shows QTL analysis of a family with a skewed sex ratio (64% males), lower part based on reduced dataset (XY males removed)
- Second sex determining locus detected in LG20
- QTL on LG20 also showed chromosome-wide effect in another family with a high temperature effect on sex ratio (also coincidence of GSD/TISD loci in LG1, 3, 23; cf Luhmann et al. 2012)

(Palaiokostas et al. 2015)

Nile tilapia: selection for response to temperature effects on sex ratio
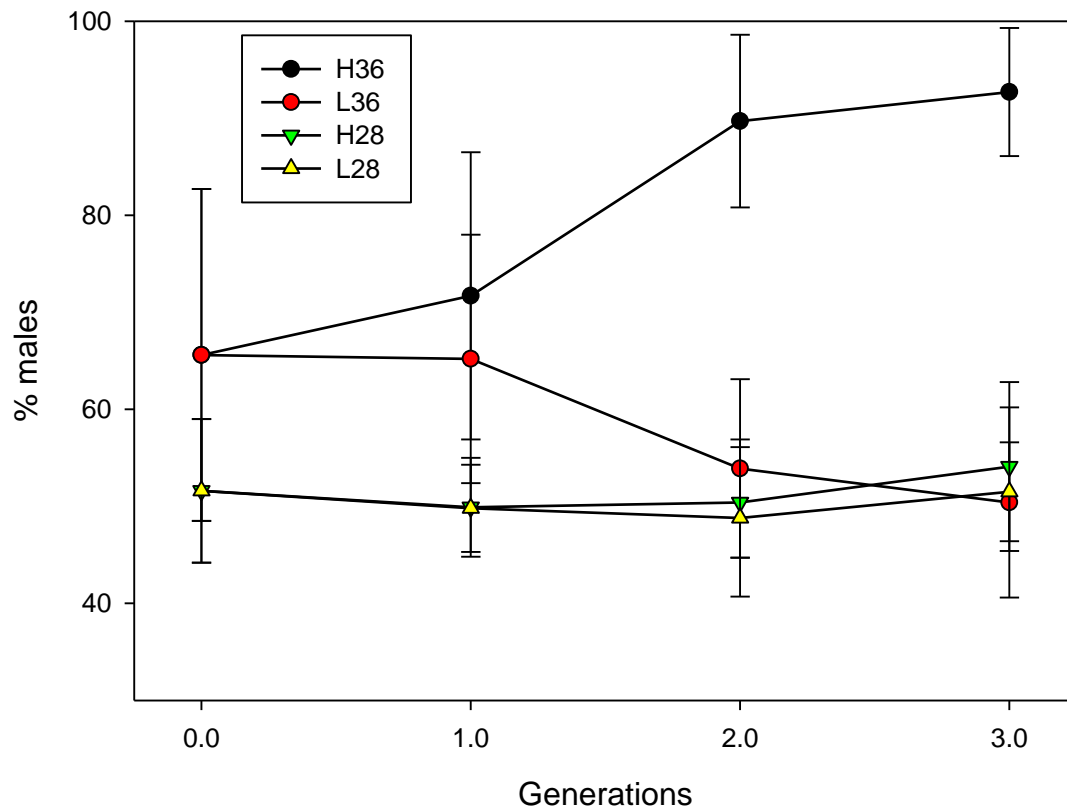
Nile tilapia shows genetic polymorphism for response of sex ratio to early rearing (~10d) at high temperature (36°C).
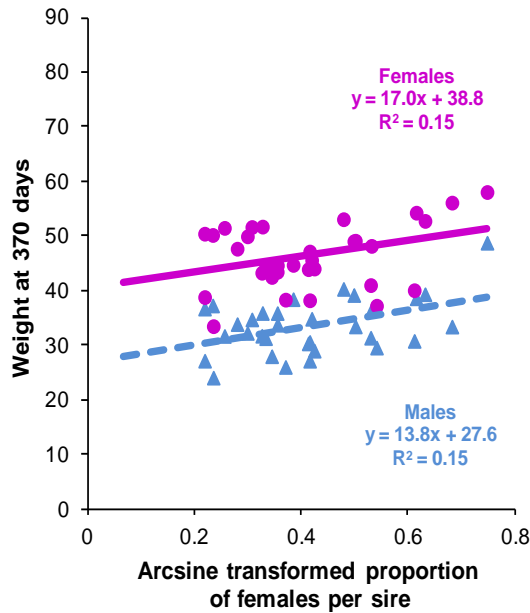
Three generations of bidirectional selection on a L. Manzalla population produced 93 % males in H line and 50 % males in L line at 36°C (both lines close to 50% at 28°C).

Wessels and Hörstgen-Schwark (2011)

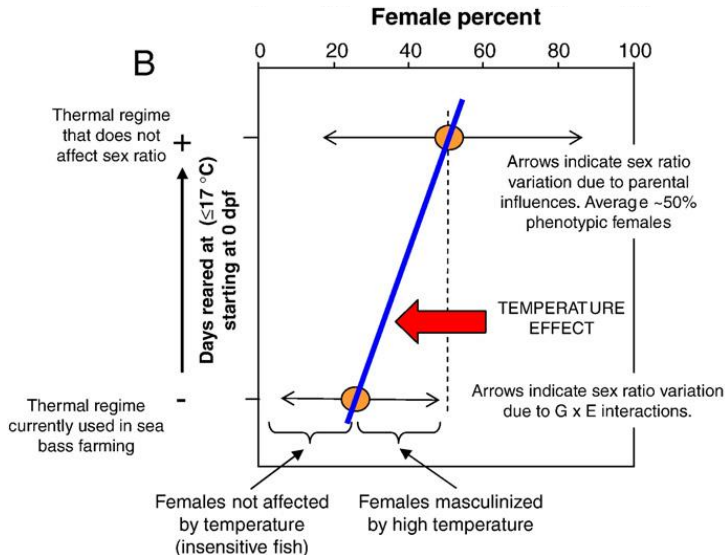# European seabass: polygenic + temperature effects

allegromar.com

h² of sex tendency= 0.62±0.12
Gen corr with BW: r_A= 0.50 ± 0.09

Select for faster growth

Change early thermal regime

Increased proportion of females
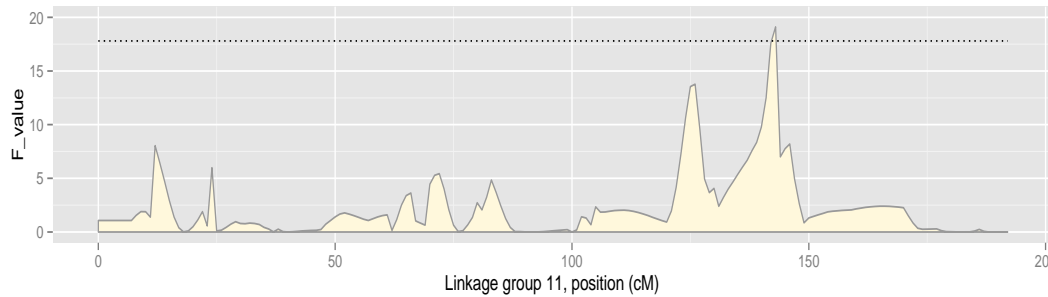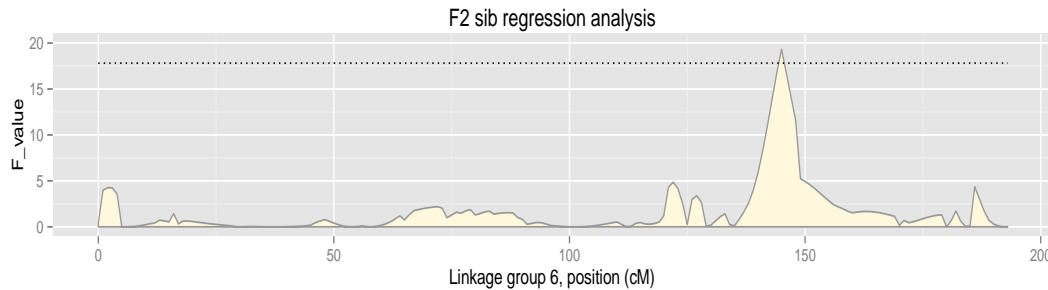
Females grow faster and mature later

Vandeputte et al. (2007)
Navarro-Martin et al. (2009)

# European Seabass sex QTL analysis (RADseq)

F2 sib regression analysis
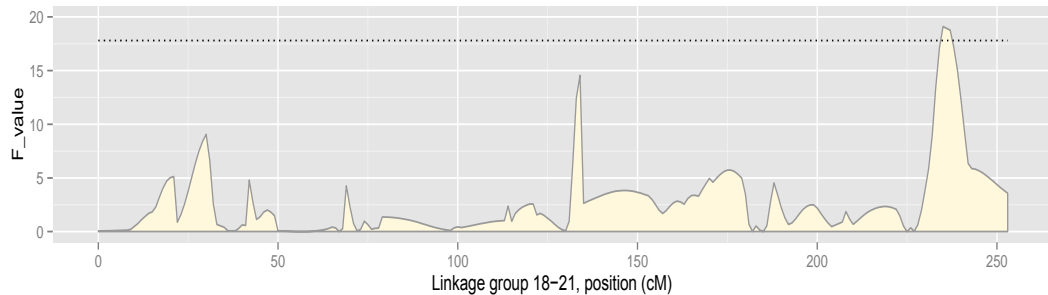
- F2 cross from extreme sex ratio F0s
- Linkage map - 6706 markers (4817 cM)
- Minor effects of three QTLs
- Supports polygenic hypothesis (presentation later)

The curious case of zebrafish sex determination
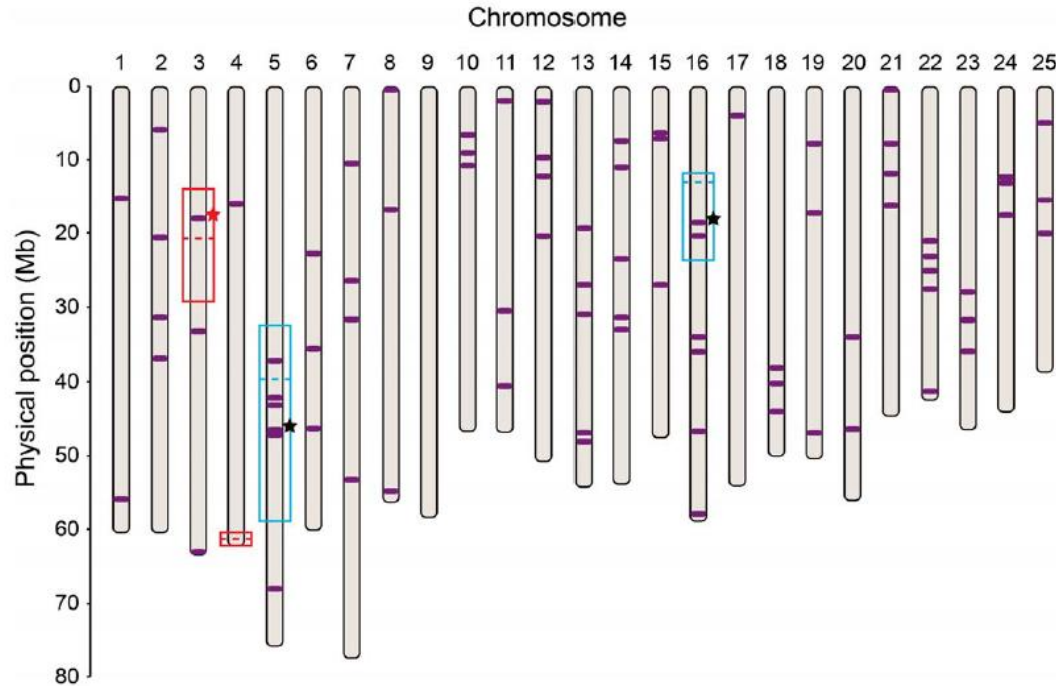
- Sex determination in zebrafish poorly understood until recently, despite its importance as model organism – variable and often highly distorted sex ratios
- SNP-based mapping (RADseq and others) suggested polygenic sex determination in domesticated zebrafish (QTLs shown by red and blue boxes above: Anderson et al. 2012)
- Subsequent analysis of natural strains (using GWAS-RAD) showed WZ/ZZ system based in chromosome 4 (Wilson et al. 2014) – function (almost) lost in domesticated strains

Not just fish – sea lice too
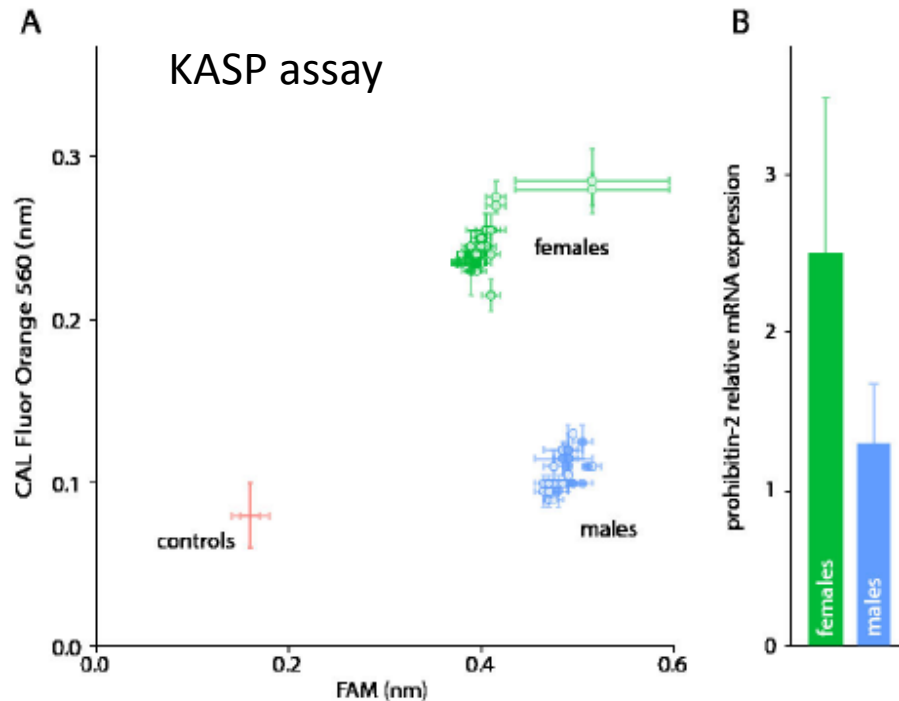
- RADseq used to isolate SNP markers in salmon louse (*L. salmonis*) strain
- One SNP showed complete association with sex in this and three other strains (A: heterozygous in females, homozygous in males – WZ/ZZ)
- SNP is in prohibitin-2 gene, which shows sex-dependent expression (B)

(Carmichael et al. 2013)

From mapping to sex determining genes

- Fine mapping, breakpoint analysis, etc
- Comparison to genome assembly (if available), synteny, etc
- Candidate genes
- Expression studies (targetted genes or broader transcriptomic approaches)
- Epigenetics
- Knockout (e.g. CRISPR), gene transfer
- Confirmation of candidate genes
- GAS rather than MAS for monosex production

(Candidate) sex-determining genes in fish (all XX/XY spp.)

| Species	SD gene	Origin	Genomic structure
Medaka [1] (*Oryzias latipes*)	*DMY*	Duplicated *Dmrt1b*	Y-specific
Luzon ricefish [2] (*O. luzonensis*)	*Gsdf*		Allelic on Y and X
Fugu [3] (*Takifugu rubripes*)	*Amhr2*		Allelic on Y and X (one SNP difference)
Patagonian pejerrey [4] (*Odontesthes hatcheri*)	*amhy*	Duplicated *amh*	Y-specific
Rainbow trout [5] (*Oncorhynchus mykiss*)	*sdY*	Similar to C-terminal of *IRF 9*	Y-specific
Nile tilapia [6] (*Oreochromis niloticus*)	*amhy*		LG23 (in this popn)

Takifugu rubripes (smithlabdb.usc.edu)

Odontesthes hatcheri (irresistibleflyshop.com.ar)

[1]Matsuda et al., 2003; [2]Hattori et al., 2012; [3]Kamiya et al., 2012; [4]Hattori et al., 2012; [5]Yano et al., 2012; [6] Eshel et al. (2012, 2014)

Atlantic salmon "sex test"
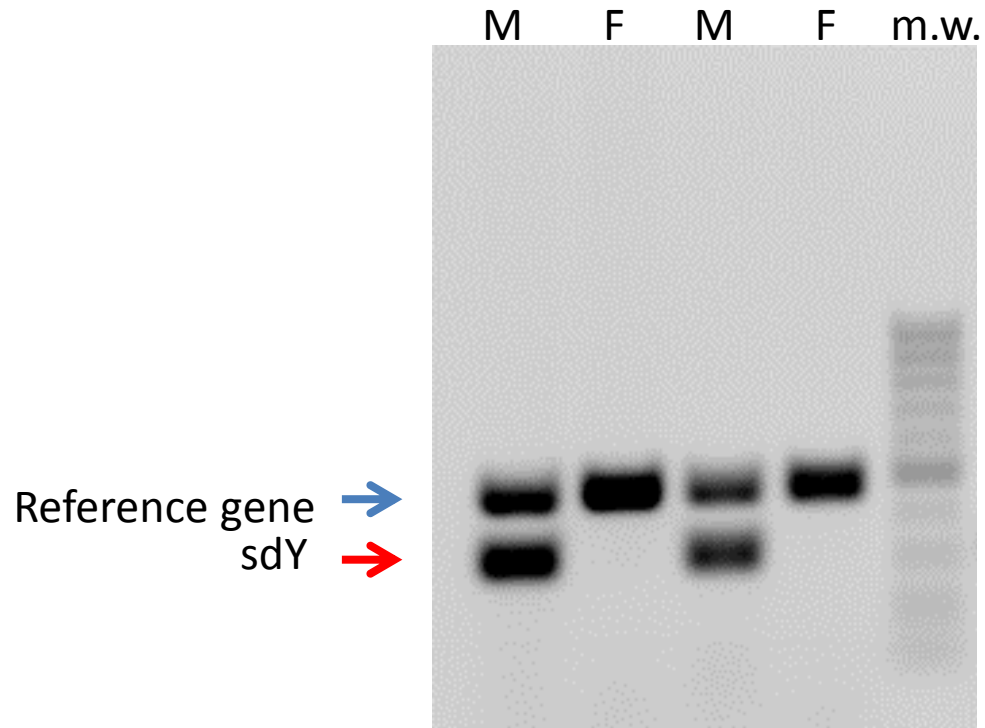
(Image from John Taggart)

Concluding remarks

- GBS offers a scalable, fairly quick and cost-efficient way for an initial assessment of the sex-determining system in "new" species

- May lead to isolation of markers for MAS for monosex development, identification of regions to search for candidate genes, etc in sample sex determining systems (XX/XY, WZ/ZZ)

- Can also help in elucidating more complex systems (oligogenic, polygenic, GSD/TISD)