CiBreed
Center for Integrated Breeding Research

Understanding the regulatory mechanisms for sex determination and colour pattern genes in zebrafish

S. Hosseini[1,2] , M. Gültas[2,3] , L. Hahn[3] , A.O. Schmitt[2,3] , H. Simianer[1,2] A.R. Sharifi[1,2]

[1]Animal Breeding and Genetics Group, Department of Animal Sciences, University of Goettingen, 37075 Goettingen

[2]Center for Integrated Breeding Research, University of Goettingen, 37075 Goettingen

[3]Breeding Informatics Group, Department of Animal Sciences, University of Goettingen, 37075 Goettingen

Introduction

> **Phenotypic differences between males and females have been a subject of intense research since Darwin**

> **The regulation of sex-biased gene expression plays a major role in phenotypic dimorphism in many taxa** (Small et al., 2009)

> **Sex-biased gene expression revealed that sex determination genes might be associated with colour pattern genes in respect to sexual dimorphism** (Sharma et al., 2014)

Introduction

> **In zebrafish, males show a more intense yellow colouration compared to females, particularly upon sexual activity** (Singh & Nüsslein-Volhard, 2015)

> **The pronounced phenotypic differences in colouration between males and females were observed in the caudal fin** (Hosseini et al., 2018)

> **Can we identify specific and common transcriptional regulatory mechanism for sex determination and colour pattern genes?**

Introduction
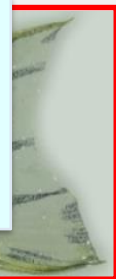
> **The regulation of gene expression in living organism is controlled by transcription factors (TF) and their combinatorial interplay** (Meckbach et al., 2015)

> **Transcription factors are a special class of gene regulatory proteins that bind to the promoter or enhancer of DNA in a cooperative manner to regulate the transcription of a gene**
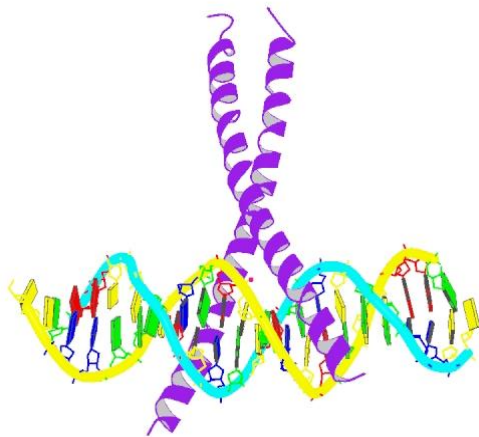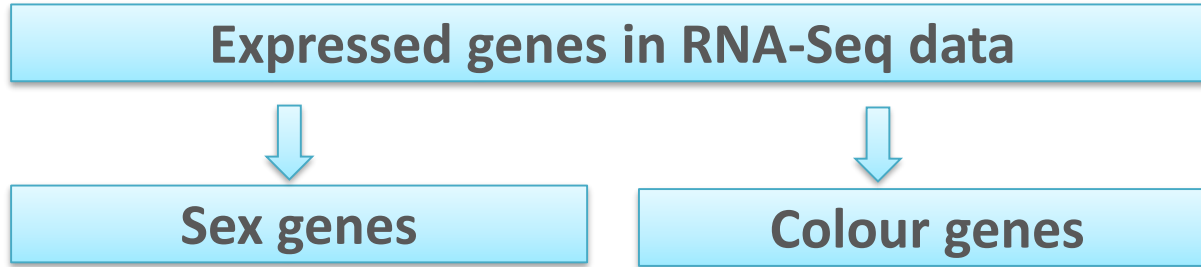

Consensus binding sequence

Enhancer

Promoter

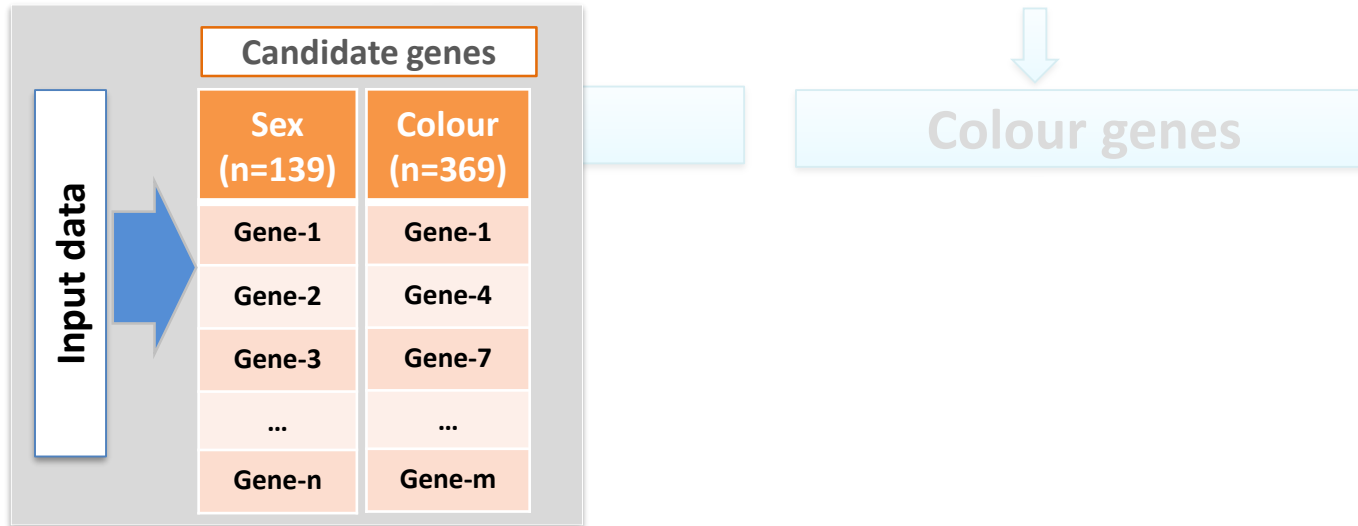**TSS: Transcription Start Site**
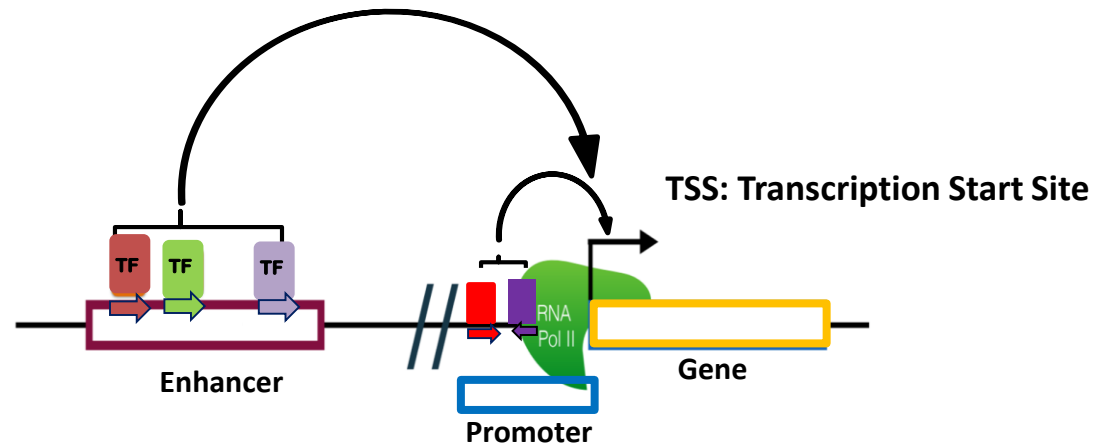
Gene

The main goal of this study was:

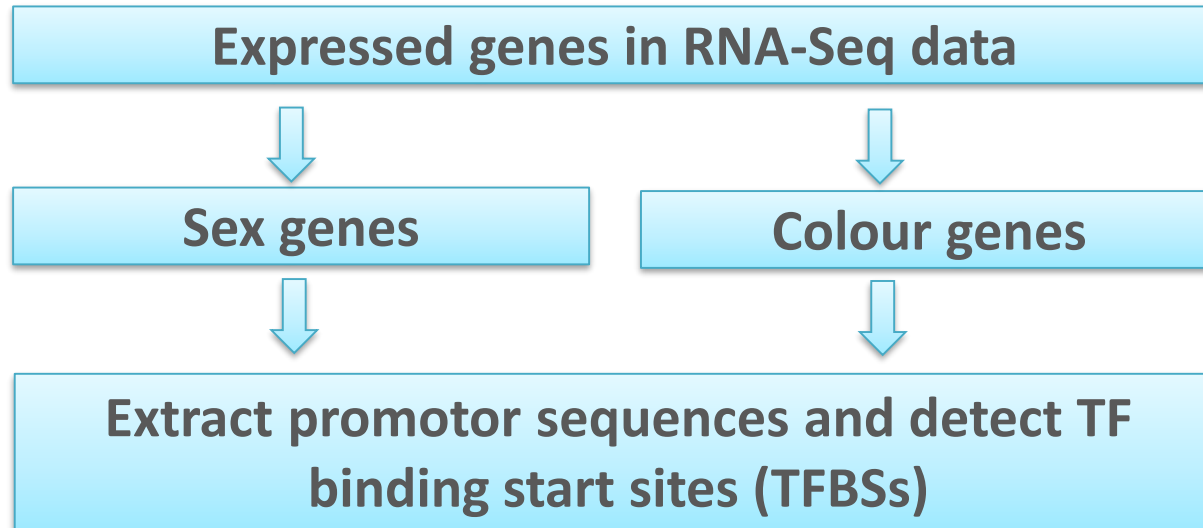➢ **to investigate the specific and common TFs cooperation in the regulation mechanism of sex determination and sex-associated colour pattern genes in zebrafish**

➢ **to gain new insights into the complex transcriptional mechanisms of sex determination and colour pattern genes in zebrafish as a model animal**

Material and Methods

Expressed genes in RNA-Seq data

Sex genes

Colour genes

Material and Methods

Expressed genes in RNA-Seq data

Candidate genes	
Sex (n=139)	**Colour (n=369)**
Gene-1	Gene-1
Gene-2	Gene-4
Gene-3	Gene-7
...	...
Gene-n	Gene-m

Input data

Colour genes

Material and Methods

Expressed genes in RNA-Seq data

Sex genes

Colour genes

Extract promotor sequences and detect TF binding start sites (TFBSs)

Expressed genes in RNA-Seq data

Input data

Candidate genes

Sex (n=139)	Colour (n=369)
Gene-1	Gene-1
Gene-2	Gene-4
Gene-3	Gene-7
...	...
Gene-n	Gene-m

Promoter analysis

TFBSs prediction

Material and Methods

Expressed genes in RNA-Seq data

Sex genes

Colour genes

Extract promotor sequences and detect TF binding start sites (TFBSs)

Apply PC-TraFF to identify the specific TF cooperation

Mapping the cooperative TFs in the Cytoscape platform

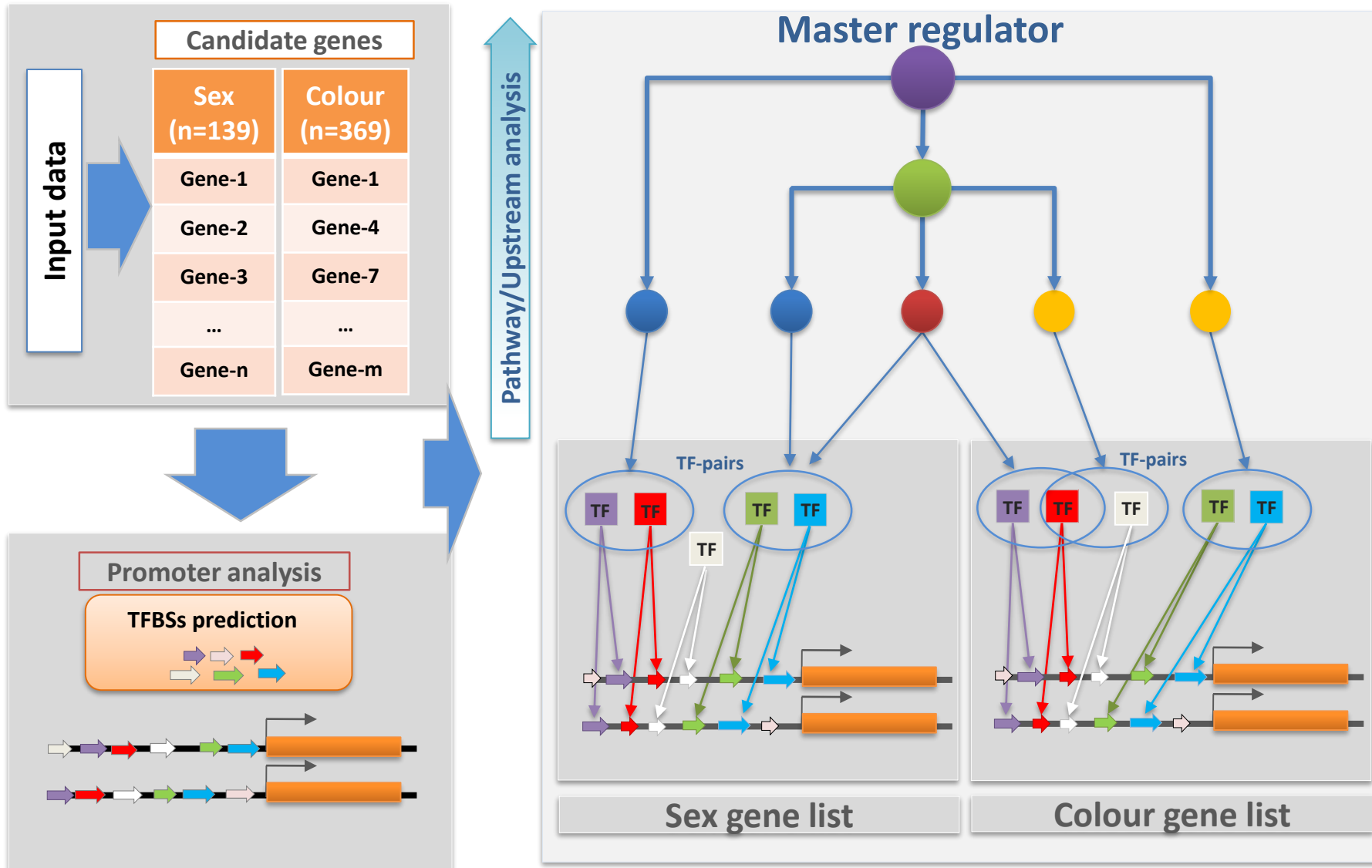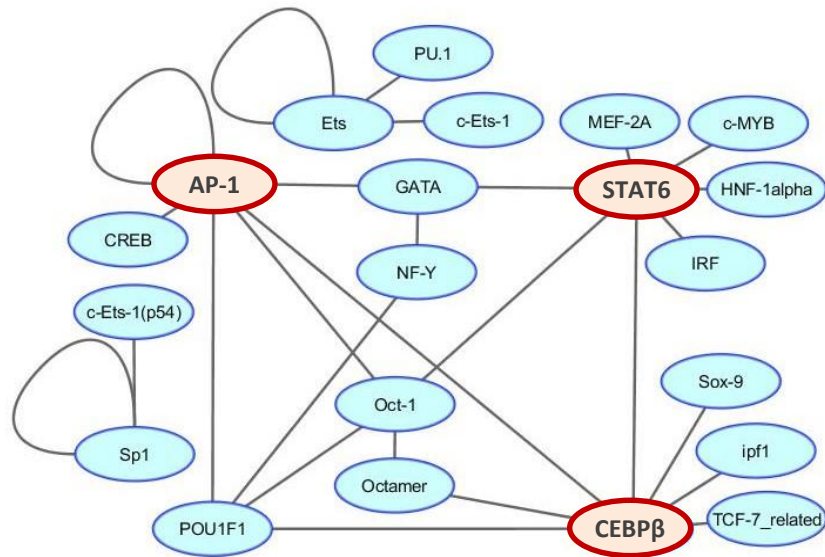## Identification of cooperative TFs and master regulator

Result

The identified specific transcription factor (TF) cooperation for sex determination and colour pattern genes

Specific TF cooperation for sex determination genes

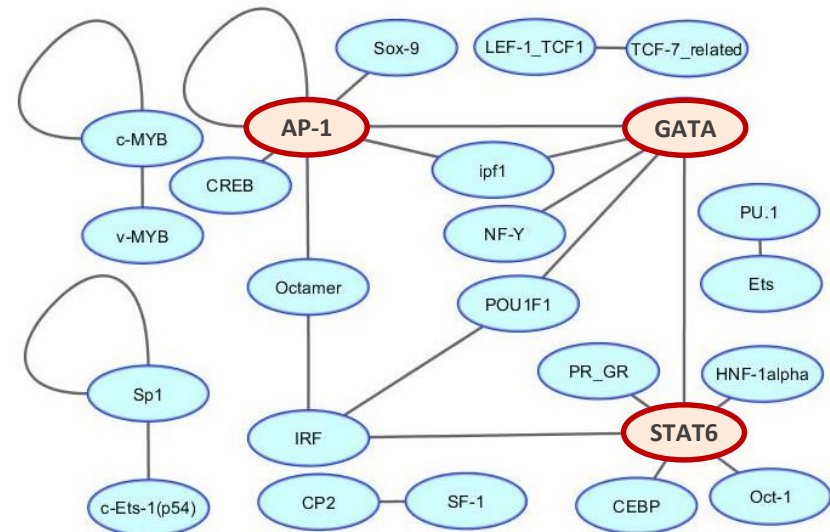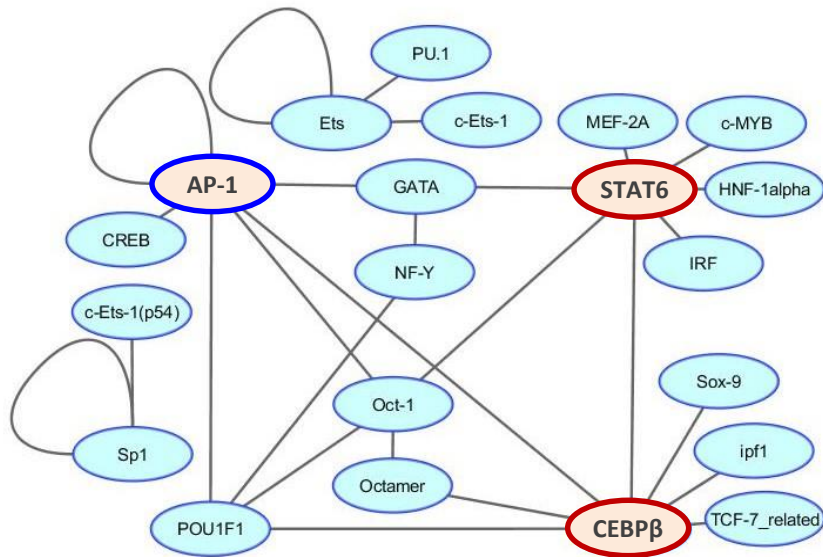Specific TF cooperation for colour pattern genes

Result

The identified specific transcription factor (TF) cooperation for sex determination and colour pattern genes

Specific TF cooperation for sex determination genes

Specific TF cooperation for colour pattern genes

- AP1 plays pivotal role in regulation of the expression of several genes involved in steroidogenesis and male sex differentiation pathway

- AP1 coordinates transport and delivery of melanin synthesizing enzymes (e.g. *tyrp1*) to maturing melanosomes in the cells

Result

The identified specific transcription factor (TF) cooperation for sex determination and colour pattern genes

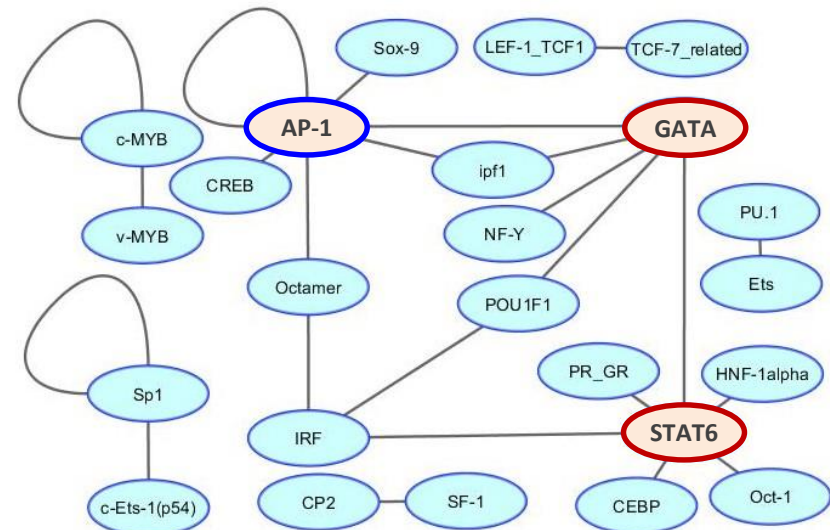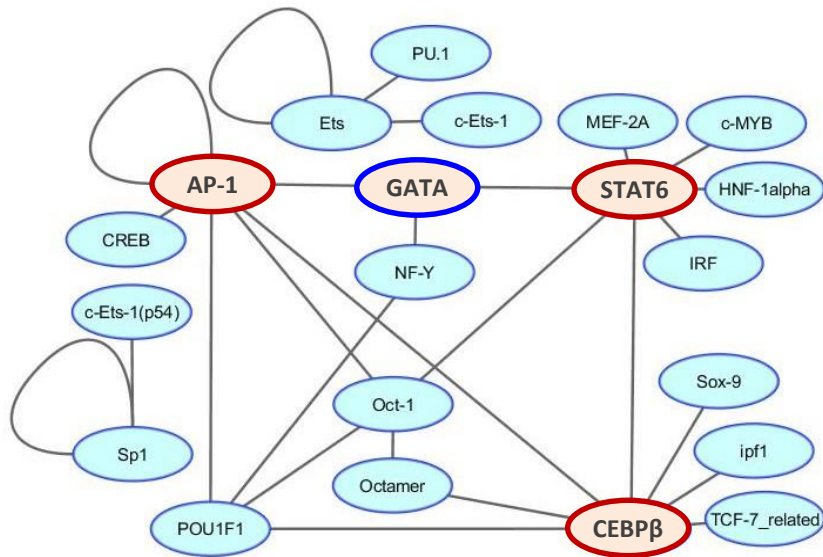Specific TF cooperation for sex determination genes

Specific TF cooperation for colour pattern genes

- GATA4 is associated with the process of sex determination and is involved in testicular differentiation and spermatogenesis

- GATA family members are essential for formation of pigmentation. GATA3 has an important role in the development of melanocytes

The identified specific transcription factor (TF) cooperation for sex determination and colour pattern genes
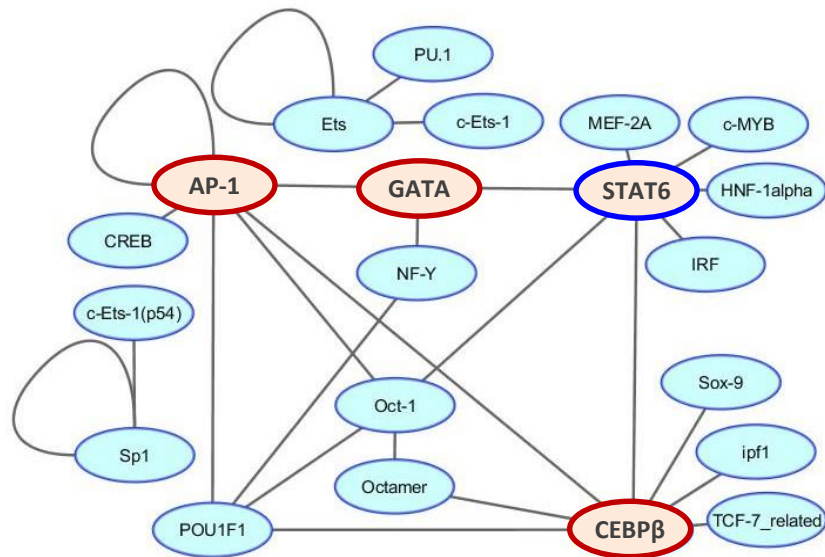
Specific TF cooperation for sex determination genes

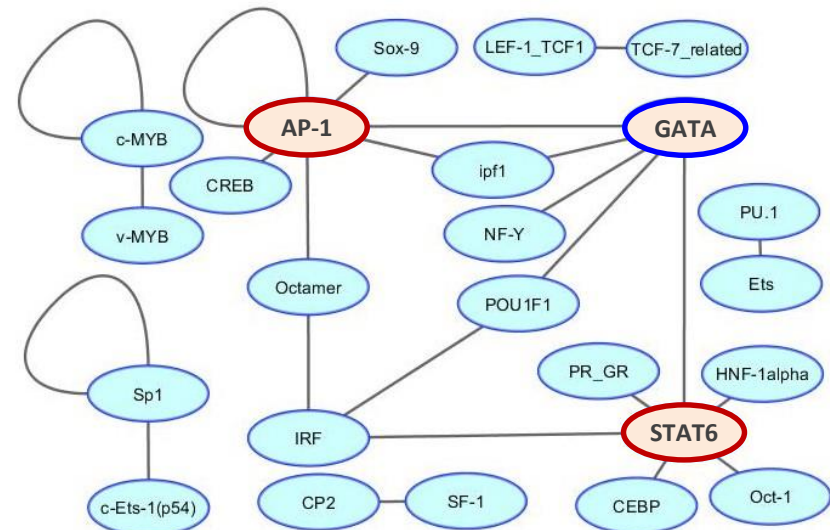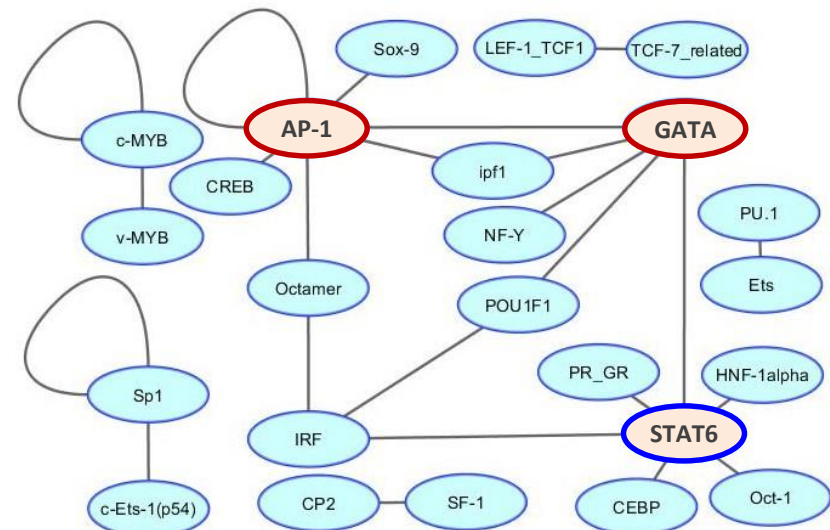Specific TF cooperation for colour pattern genes

- STAT6 plays an important role in reproductive function, particularly in the process of sperm activation. It is also involved in PI3K signaling pathway activity during ovarian development

- STAT6 is one of the most important factors to regulate the expression of genes responsible for melanin pigmentation in melanogenesis process

Result

The identified common transcription factor (TF) cooperation for sex determination and colour pattern genes
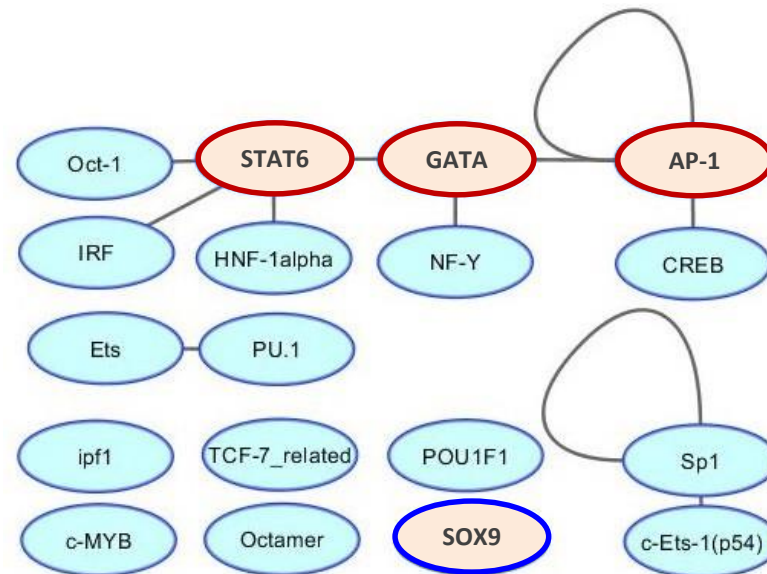
Common single TFs and their cooperation for sex and colour genes

- Activation of SOX9 leads to disassembly of follicles, oocyte degeneration and conversion of follicle cell during juvenile ovary–to-testis transformation resulted in testis formation in zebrafish

- SOX9 is a key player in the differentiation and pigmentation of melanocyte during pigment cell development

Result

The identified master regulator (PKACA) for sex determination and colour pattern genes

Pathway analysis (Colour genes)

Pathway analysis (Sex genes)

Shared key regulator

- PKACA is a protein, which encodes by *PRKACA* gene, and is involved in many intracellular signaling events

Result
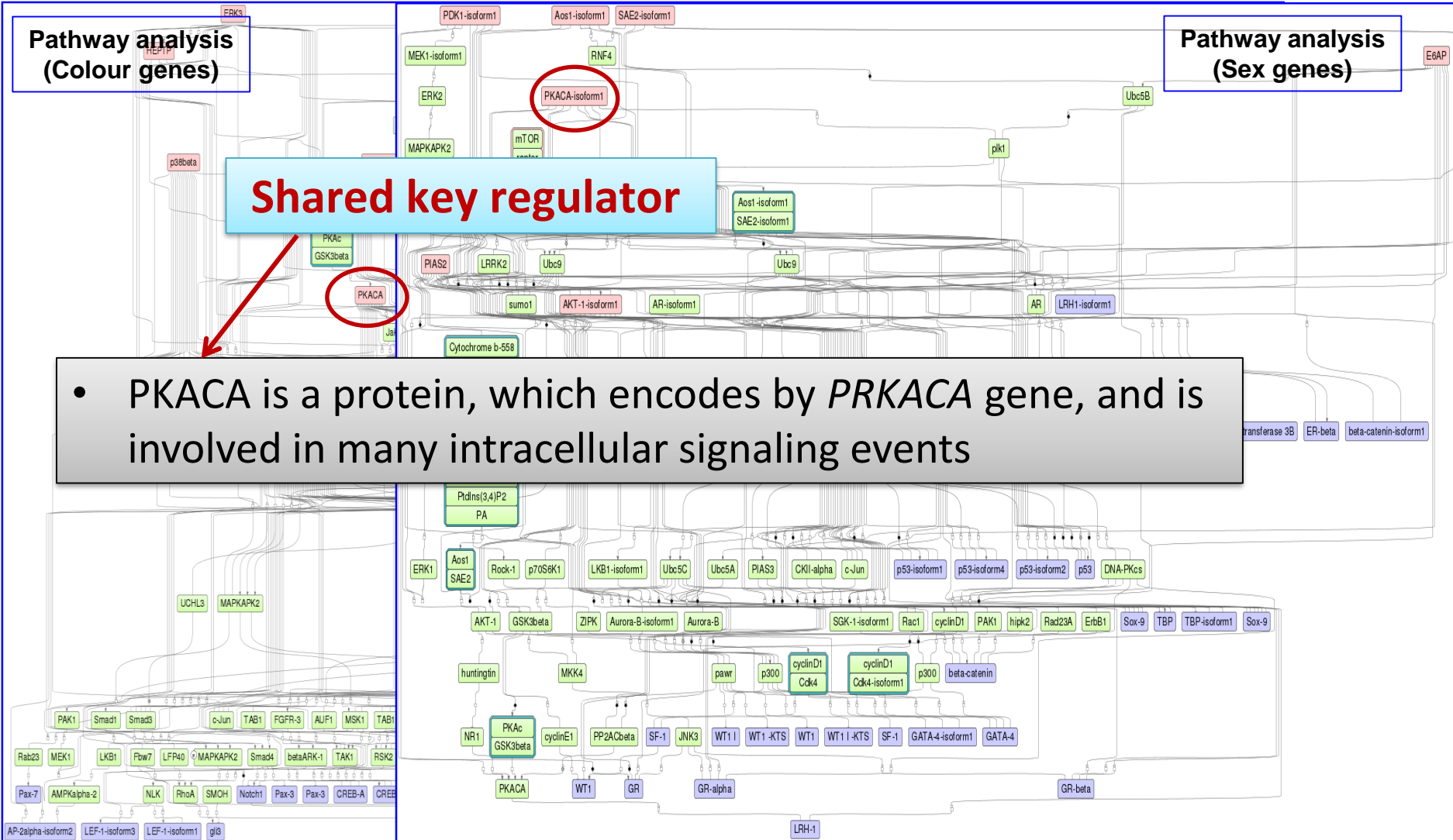
PRKACA involved in different pathways

Pathway	Melanogenesis	
	Colour	Sex
Gene	➢ *creb1a* ➢ *mitfa* ➢ *kita*	➢ *wnt4a* ➢ *ctnnb1* ➢ *wnt11*

Pathway	Wnt signaling	
	Colour	Sex
Gene	➢ *mycb* ➢ *apc* ➢ *lef1*	➢ *wnt11* ➢ *ctnnb1* ➢ *tp53*

PRKACA

Pathway	MAPK signaling	
	Colour	Sex
Gene	➢ *casp3a* ➢ *egfra* ➢ *map2k1*	➢ *tp53* ➢ *pdgfra*

Pathway	Tight junction	
	Colour	Sex
Gene	➢ *tjp1b* ➢ *ltgb1a* ➢ *itgb1b.2*	➢ *gata4* ➢ *tuba4l* ➢ *tuba7l*

Result
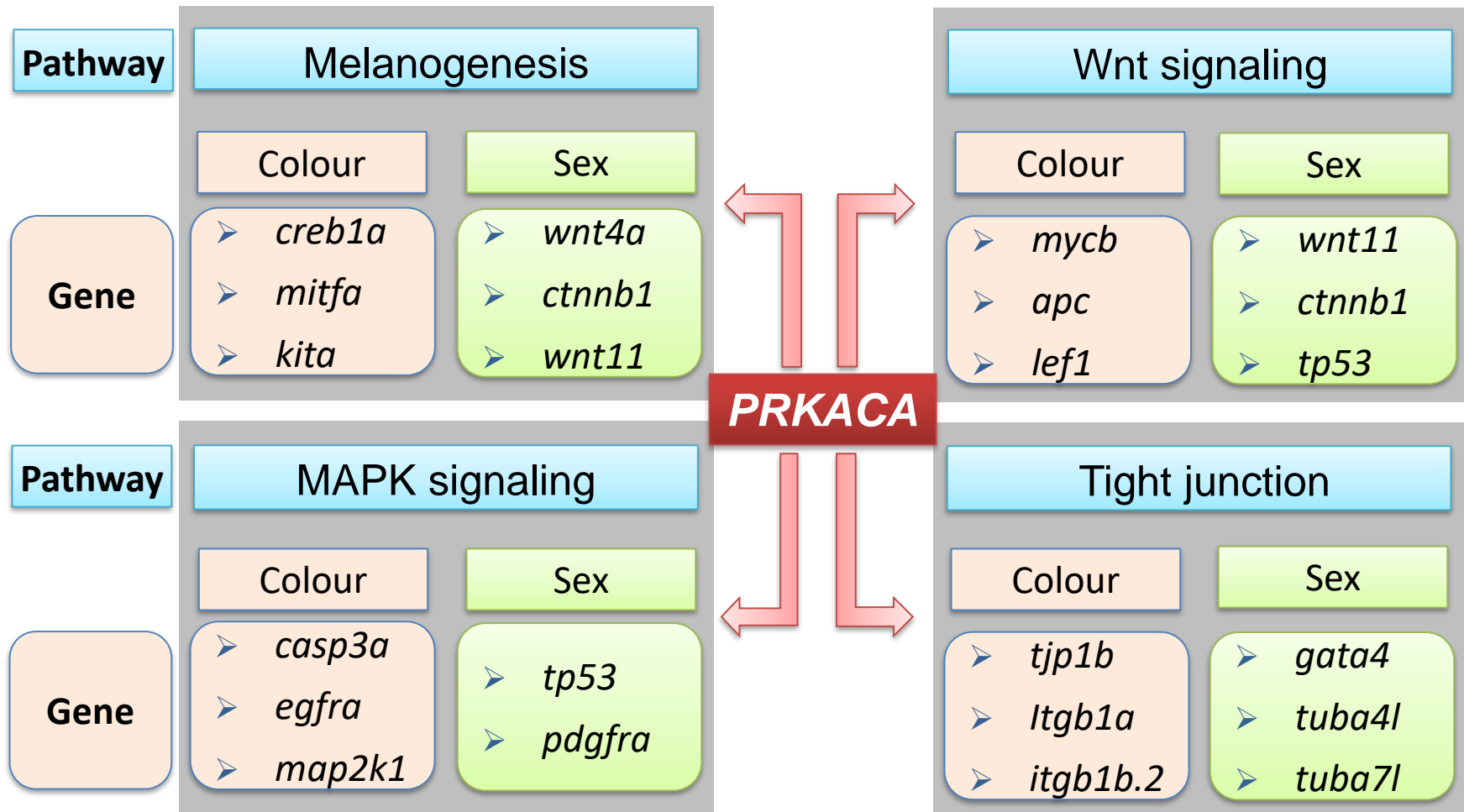
Protein-protein interaction network

- Red colour nodes illustrate Melanogenesis pathway

- Violet colour nodes illustrate MAPK signaling pathway

- Green colour nodes illustrate Wnt signaling pathway

Conclusion
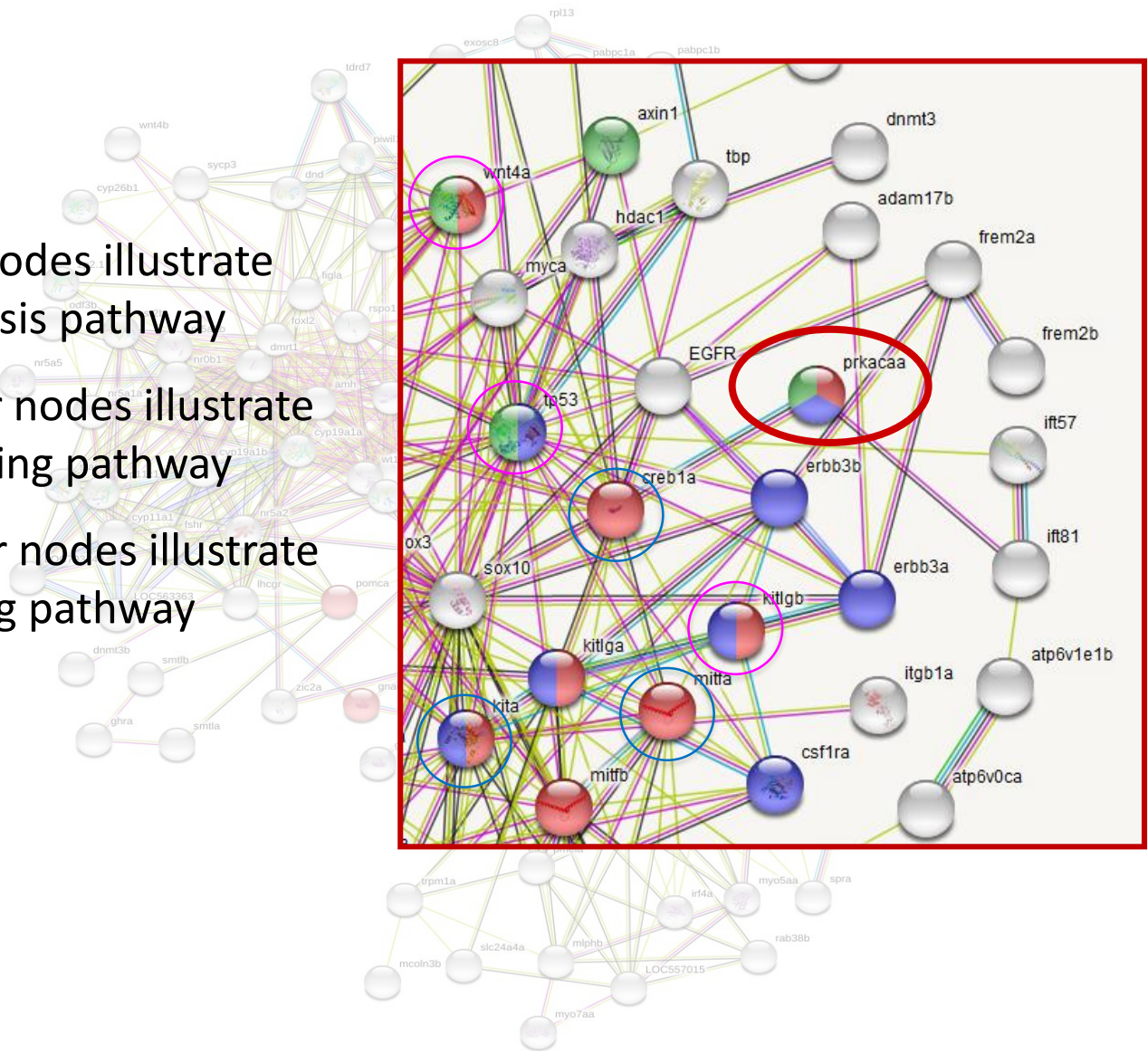
- ➢ **We identified several specific cooperative TFs for sex determination and colour pattern genes**

- ➢ **The intersection between both TF networks (sex and colour) demonstrated the common TF cooperation for both gene-sets**

- ➢ **Upstream/pathway analysis resulted in identification of a master regulator (PKACA) for both phenotypes**

- ➢ **Our finding in this study could provide insight into the functional background of the association between sex and colour genes in zebrafish inducing phenotypic dimorphism**

THANK YOU VERY MUCH FOR YOUR ATTENTION

http://e360.yale.edu/images/slideshow/Tank-of-zebrafish-YaleE360_800.jpg

BMC Genomics

RESEARCH ARTICLE   **Open Access**
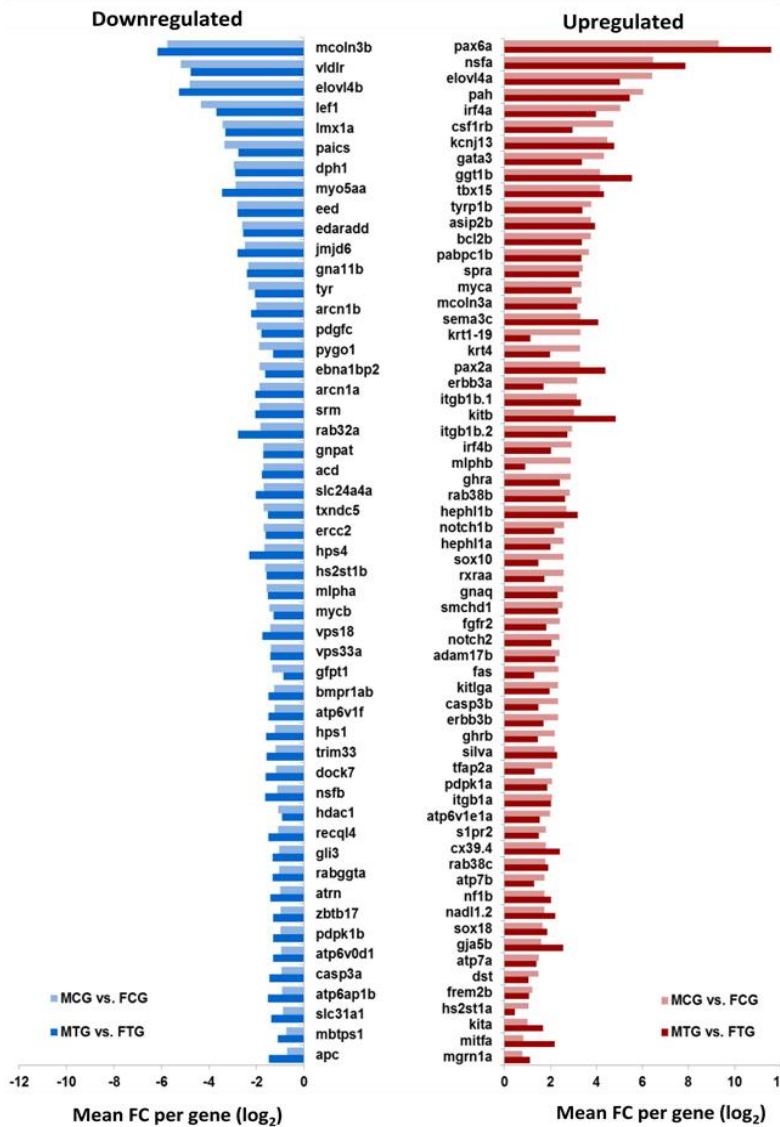
Genetic mechanism underlying sexual plasticity and its association with colour patterning in zebrafish (*Danio rerio*)

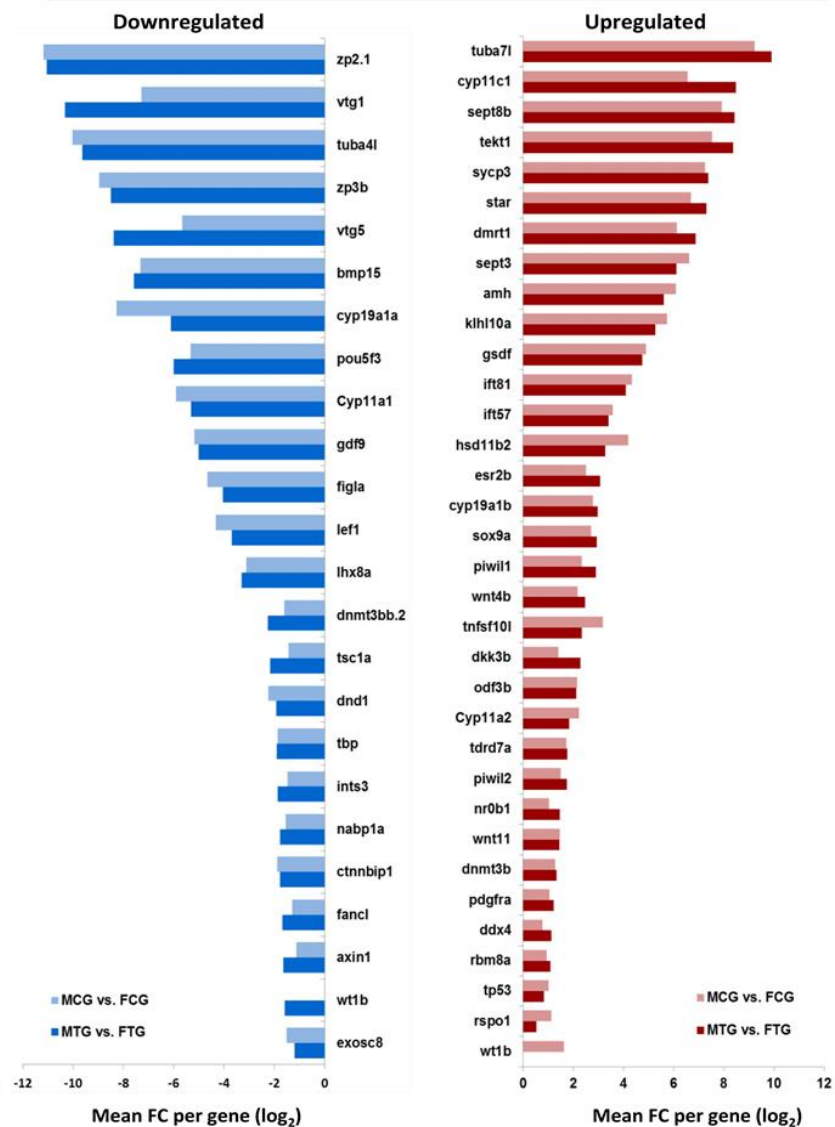Shahrbanou Hosseini[1,2]*, Ngoc-Thuy Ha[1,2], Henner Simianer[1,2], Clemens Falker-Gieske[1,2], Bertram Brenig[1,2,3], Andre Franke[4], Gabriele Hörstgen-Schwark[1], Jens Tetens[1,2], Sebastian Herzog[5,6] and Ahmad Reza Sharifi[1,2]
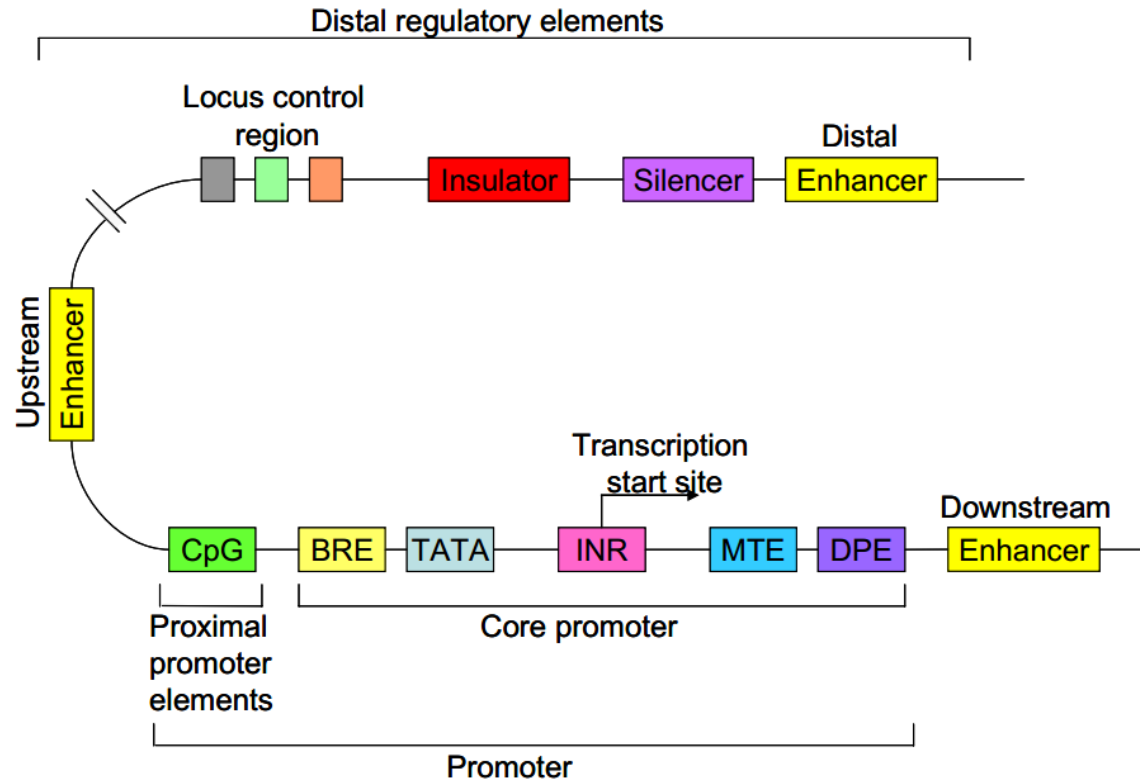
Colour pattern genes

Downregulated

mcoln3b, vldlr, elovl4b, lef1, lmx1a, paics, dph1, myo5aa, eed, edaradd, jmjd6, gna11b, tyr, arcn1b, pdgfc, pygo1, ebna1bp2, arcn1a, srm, rab32a, gnpat, acd, slc24a4a, txndc5, ercc2, hps4, hs2st1b, mlpha, mycb, vps18, vps33a, gfpt1, bmpr1ab, atp6v1f, hps1, trim33, dock7, nsfb, hdac1, recql4, gli3, rabggta, atrn, zbtb17, pdpk1b, atp6v0d1, casp3a, atp6ap1b, slc31a1, mbtps1, apc

Upregulated

pax6a, nsfa, elovl4a, pah, irf4a, csf1rb, kcnj13, gata3, ggt1b, tbx15, tyrp1b, asip2b, bcl2b, pabpc1b, spra, myca, mcoln3a, sema3c, krt1-19, krt4, pax2a, erbb3a, itgb1b.1, kitb, itgb1b.2, irf4b, mlphb, ghra, rab38b, hephl1b, notch1b, hephl1a, sox10, rxraa, gnaq, smchd1, fgfr2, notch2, adam17b, fas, kitlga, casp3b, erbb3b, ghrb, silva, tfap2a, pdpk1a, itgb1a, atp6v1e1a, s1pr2, cx39.4, rab38c, atp7b, nf1b, nadl1.2, sox18, gja5b, atp7a, dst, frem2b, hs2st1a, kita, mitfa, mgrn1a

Sex determination genes

Downregulated

zp2.1, vtg1, tuba4l, zp3b, vtg5, bmp15, cyp19a1a, pou5f3, Cyp11a1, gdf9, figla, lef1, lhx8a, dnmt3bb.2, tsc1a, dnd1, tbp, ints3, nabp1a, ctnnbip1, fancl, axin1, wt1b, exosc8

Upregulated

tuba7l, cyp11c1, sept8b, tekt1, sycp3, star, dmrt1, sept3, amh, klhl10a, gsdf, ift81, ift57, hsd11b2, esr2b, cyp19a1b, sox9a, piwil1, wnt4b, tnfsf10l, dkk3b, odf3b, Cyp11a2, tdrd7a, piwil2, nr0b1, wnt11, dnmt3b, pdgfra, ddx4, rbm8a, tp53, rspo1, wt1b

Legend: MCG vs. FCG, MTG vs. FTG

X-axis: Mean FC per gene (log$_2$), scale from -12 to 12

Position weight matrix (PWM) for TFSBs

GAGGTAAAC
TCCGTAAGT
CAGGTTGGA
ACAGTCAGT
TAGGTCATT
TAGGTACTG
ATGGTAACT
CAGGTATAC
TGTGTGAGT
AAGGTAAGT

position frequency matrix (PFM)

$$M = \begin{array}{c} A \\ C \\ G \\ T \end{array} \begin{bmatrix} 3 & 6 & 1 & 0 & 0 & 6 & 7 & 2 & 1 \\ 2 & 2 & 1 & 0 & 0 & 2 & 1 & 1 & 2 \\ 1 & 1 & 7 & 10 & 0 & 1 & 1 & 5 & 1 \\ 4 & 1 & 1 & 0 & 10 & 1 & 1 & 2 & 6 \end{bmatrix}.$$

position probability matrix (PPM)

$$M = \begin{array}{c} A \\ C \\ G \\ T \end{array} \begin{bmatrix} 0.3 & 0.6 & 0.1 & 0.0 & 0.0 & 0.6 & 0.7 & 0.2 & 0.1 \\ 0.2 & 0.2 & 0.1 & 0.0 & 0.0 & 0.2 & 0.1 & 0.1 & 0.2 \\ 0.1 & 0.1 & 0.7 & 1.0 & 0.0 & 0.1 & 0.1 & 0.5 & 0.1 \\ 0.4 & 0.1 & 0.1 & 0.0 & 1.0 & 0.1 & 0.1 & 0.2 & 0.6 \end{bmatrix}.$$

probability of the sequence S = GAGGTAAAC given the above PPM **M** can be calculated:

$$p(S|M) = 0.1 \times 0.6 \times 0.7 \times 1.0 \times 1.0 \times 0.6 \times 0.7 \times 0.2 \times 0.2 = 0.0007056.$$

This is equivalent to multiplying each column of the PPM by a Dirichlet distribution and allows the probability to be calculated for new sequences

$$M = \begin{array}{c} A \\ C \\ G \\ T \end{array} \begin{bmatrix} 0.26 & 1.26 & -1.32 & -\infty & -\infty & 1.26 & 1.49 & -0.32 & -1.32 \\ -0.32 & -0.32 & -1.32 & -\infty & -\infty & -0.32 & -1.32 & -1.32 & -0.32 \\ -1.32 & -1.32 & 1.49 & 2.0 & -\infty & -1.32 & -1.32 & 1.0 & -1.32 \\ 0.68 & -1.32 & -1.32 & -\infty & 2.0 & -1.32 & -1.32 & -0.32 & 1.26 \end{bmatrix}.$$

Pointwise mutual information (PMI) for TF cooperation

The PMI of a pair of outcomes x and y belonging to discrete random variables X and Y quantifies the discrepancy between the probability of their coincidence given their joint distribution and their individual distributions, assuming independence. Mathematically:

$$\mathrm{pmi}(x; y) \equiv \log \frac{p(x, y)}{p(x)p(y)}$$

Pathway

| KEEG Pathways name	Sex determination genes	Colour pattern genes
Calcium signaling pathway	*lhcgr, pdgfra*	*ednrba, egfra, erbb3a, erbb3b, gna11a, gnaq*
Cell cycle	*tp53*	*hdac1, mycb, zbtb17, rb1*
Gap junction	*tuba4l, tuba7l, pdgfra*	*drd2a, egfra, drd2b, map2k1, gna11b, tjp1b, pdgfc, gna11a, gnaq*
Cytokine-cytokine receptor interaction	*amh*	*ghra, bmpr1ab, csf1ra, csf1rb, fas, eda*
Tight junction	*gata4, tuba4l, tuba7l, prkcz*	*itgb1b.2, itgb1a, tjp1b*
Wnt signaling pathway	*wnt4a, ctnnb1, tp53, lef1, axin1, ctnnbip1, wnt11*	*wnt1, apc, mycb, lef1*
MAPK signaling pathway	*tp53, pdgfra*	*egfra, casp3a,nf1a, fgfr2, mycb, map2k1, traf6, nf1b, casp3b, fas*
Melanogenesis	*wnt4a, ctnnb1, lef1, wnt11*	*tyrp1a, mitfa, mitfb, tyr, kita, ednrba, pomca, tyrp1b, kitlga, kitlgb, gna11a, gnaq, creb1a*
Mitophagy - animal	*tp53*	*mitfa, mitfb*
mRNA surveillance pathway	*rbm8a*	*pabpc1a, pabpc1b,*
Apoptosis	*tp53, tuba4l, tuba7l*	*atp6v1e1b,atp6v1h, wnt1, pdpk1b, map2k1, atp6v1f*
p53 signaling pathway	*tp53*	*fas, casp3a, casp3b*
RNA transport	*eef1a1l1, rbm8a*	*pabpc1a, pabpc1b,*

Enriched KEGG database pathways
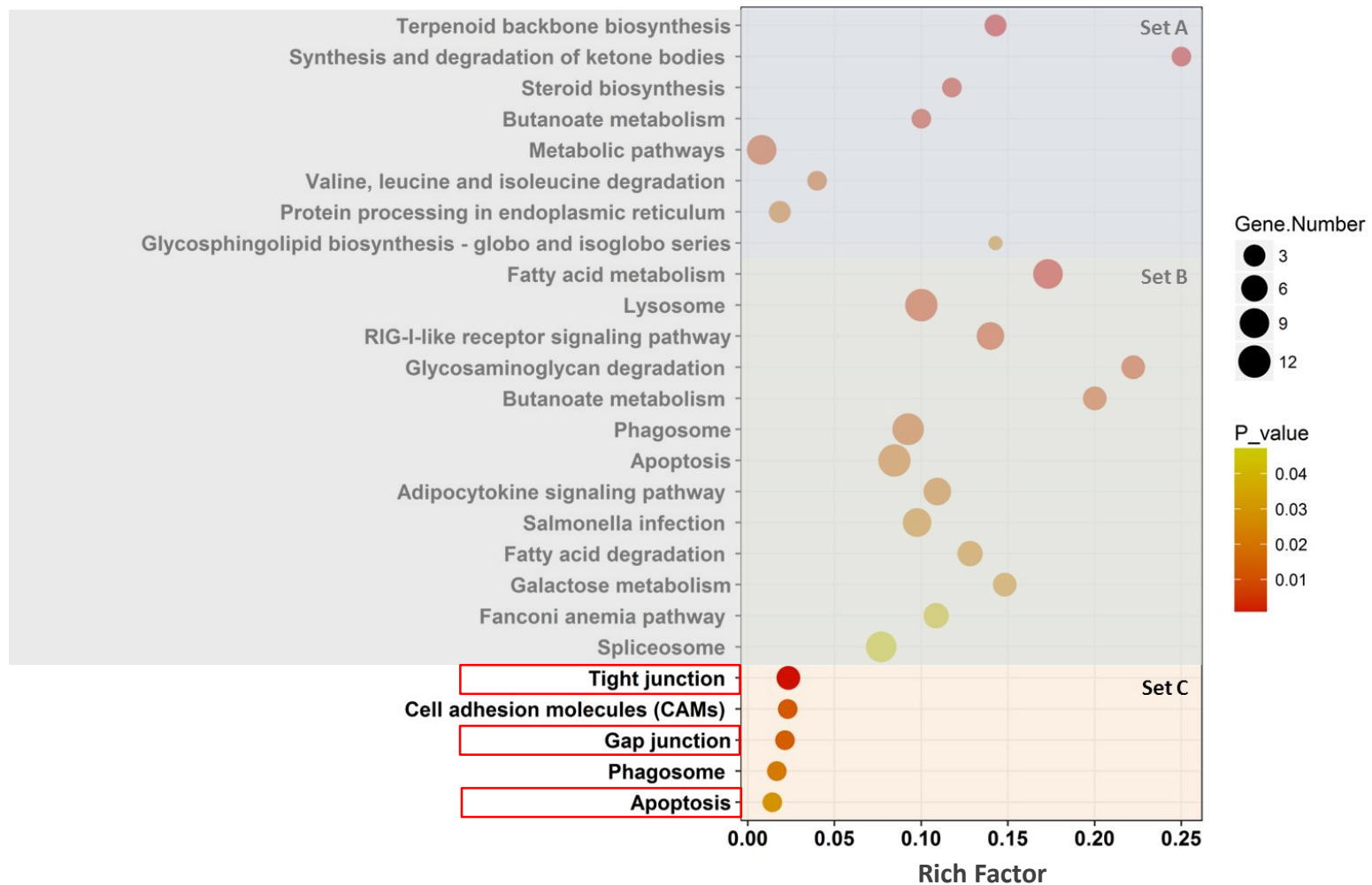
Set A: enriched GOs considering significant DEGs in MTF vs. FTF (119 genes)
Set B: enriched GOs considering significant DEGs in MTG vs. FTG (762 genes)
Set C: enriched GOs considering significantly DEGs overlapping between four comparison groups (108 genes)

Cooperation between transcription factors
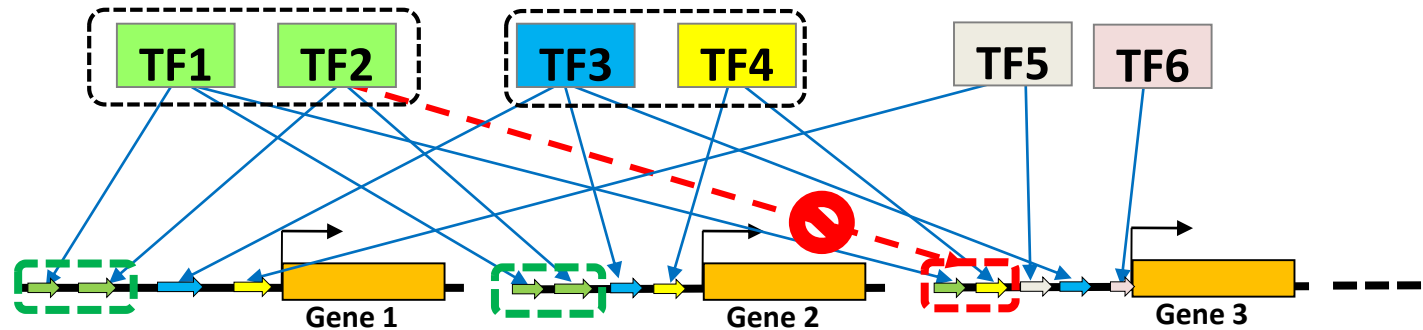
- TFs bind to DNA in a cooperative manner
  - Partner choice is not random
  - It depends on the biological functions

- Cooperation of TFs
  - Between homotypic or heterotypic TFs
  - **Synergistic** or **antagonistic** interactions