UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Mémoire de Maîtrise en médecine No 4449

Is sex specific phenotypic robustness reflected In gene expression data?

Etudiant
Makowka Soline

Tuteur
Bergmann Sven
Dpt de biologie computationnelle

Co-tuteur
Hersch Micha
Dpt de biologie computationnelle

Expert
Kutalik Zoltan
IUMSP

Lausanne, 16.12.2017

Is sex specific phenotypic robustness reflected in gene expression data?

Soline Makowka

Master project, directed by Micha Hersch and Sven Bergmann. University of Lausanne, 2017

Abstract

In Human many non-Mendelian diseases present a biased sex ratio. It is the case for some neurodevelopmental and psychiatric disorders that male are more prone to develop, as schizophrenia, or autism. On the contrary, women are more susceptible to many autoimmune disorders such as systemic lupus erythematous. With regards to immunity, it has also been observed that males are more prone to infections. Phenotypic robustness (PR), defined as the persistence of a phenotype exposed to genetic or environmental constraints could play a role in disease development. Less robust individuals could express cryptic mutations and also be more receptive to environmental stressors. If there would be a difference of robustness between sexes it could also have an impact on sexually dimorphic disease risk. Some molecular mechanisms could buffer genetic and play a role in phenotypic robustness processes and environmental risk, such as Heat shock proteins (HSP). The aim of this project was to test if a sex difference in phenotypic robustness could be investigated through transcriptional data (RNAseq) from lymphobalstoid cell lines (LCL). Higher expression of HSP or other genes could benefits to a stronger robustness. A more narrow transcriptional regulation in one sex of a gene alongside a smaller variance would indicate a more precise regulation and a possible stronger phenotypical robustness.

Gene expression data was generated as RNAseq data LCL from 260 men and 290 women from the CoLaus Study Cohort. 8'924 genes were selected. The statistical analysis was done with R.

268 sex-biased genes with differences of expression were identified, 3.003% of all the tested genes. 31 genes were identified as having a significant difference in variance in expression between sexes, i.e. 0.347% of all the genes tested. Amongst autosomal genes higher variance in expression was observed for males in 8 genes and in females in 10 genes. No significant trend of a conserved sex-biased variance was identified. In X genes 13 genes showed a higher variance in women and none in men. Isoform distribution between males and females was compared for the 31 genes in in order to determine if it was part of the etiology of the variance differences. Significant differences in isoform distribution were found in 1 autosomal gene. 4 HSP genes showed a sex-biased expression, 3 female-biased and one male-biased. 320 Immune-related genes were also specifically screened for differences in expression and in variance in expression. 23 autosomal genes were identified as sex-biased and none with a difference in variance in expression.

Some HSP levels could differ between sexes and play a role in phenotypic robustness process. This study showed small but significant sex-biased expression of HSP in LCL and further analyses of other tissues should be conducted as some HSP might also be regulated by hormonal signals. At the regulation level this RNAseq all-transcriptome screening for sex-specific patterns of regulation did not show a sex-biased trend. Isoforms analysis between sex did not show significant differences in pattern distribution. This mean that the hypothesis of a general sex-specific pattern of regulation leading to a differential robustness mechanisms is unlikely, although the sex-biased expression of specific genes could still play a role. An alternative option would be that LCL could also not be appropriate to measure it.

Keywords:
Sexual dimorphism, Phenotypical robustness, Gene expression, Heat shock proteins, Immune-related genes.

Table of content

Introduction

In Human many non-Mendelian diseases present a biased sex ratio. It is the case for some neurodevelopmental and psychiatric disorders that male are more prone to develop, as schizophrenia, or autism[1,2]. In autism-related disorders women seem to be partially genetically robust to certain mutations that already cause a disease in men[3]. On the contrary, women are more susceptible to many autoimmune disorders such as systemic lupus erythematous, multiple sclerosis, or thyroid autoimmune disorders[4].

With regards to immunity, it has also been observed that males are more prone to infections and die earlier than females in humans but also in many other animal species[5]. Many hypotheses have been formulated to explain this sex-bias. In mammals, males are heterogametic with only one X chromosome and no possibility of compensating for deleterious mutations with a second set of alleles. However, in birds with their ZZ/ZW system, females are heterogametic and they are still more robust against infections[6]. Differences in behaviour and exposure to infection risk could also be important. Sexual hormones such as androgens and oestrogens, and stress hormones such as glucocorticoids, are known to play a great role in the immune response. The Immunocompetence Handicap theory hypothesis that males with good genes can bear high levels of testosterone leading to high development of secondary sexual ornament to attract females while having a less functional immunity as a trade-off [7,8].

However, a difference in sexual hormones levels might be insufficient to explain all the sexual dimorphism in immune defences as this trend has also been shown in some invertebrate species, that lack testosterone as insects, as for example Copepods [9,10].

Phenotypic robustness (PR), defined as the persistence of a phenotype exposed to genetic or environmental constraints could play a role in disease development. Less

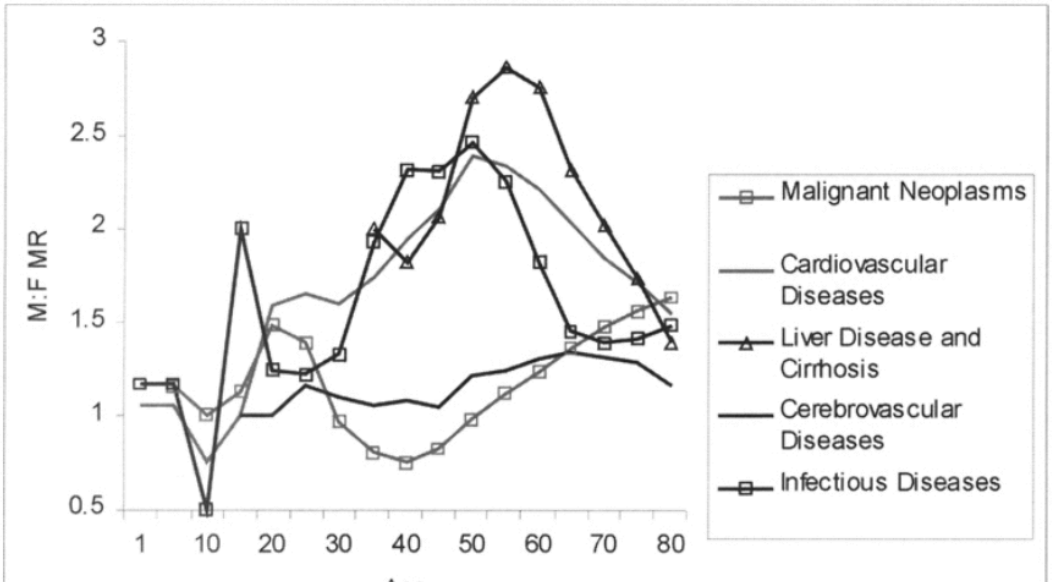

Figure 1. 2001 U.S. M:F Mortality for internal causes. Men are more prone to die from infectious diseases when older than 20 years.

robust individuals could express cryptic mutations and also be more receptive to environmental stressors. For the males which maximize their fitness through competition during reproductive periods their mating success might be more variable than the females and it could make them express cryptic variant revealed by a weaker PR[11]. In *Drosophila* a study showed that mutations affect both sexes but that selection is higher in males[12]. Sexual selection theories predict that females maximize their fitness through longevity and surviving more reproductive seasons. So it could be potentially more advantageous to be more robust against external stressors (genetic mutations or environmental stressor as infections). If there would be a difference of robustness between sexes it could also have an impact on sexually dimorphic disease risk.

Some molecular mechanisms that could buffer genetic and environmental risk, thus increasing robustness, are partially known. Some classes of molecules such as the chaperone HSP90 seem to play a role in buffering both environmental and genetic robustness and could lead to expression of cryptic variant when inactivated in *Drosophila melanogaster* and thus play a role in evolution process[13,14]. Chaperones also known as heat shock proteins (HSP) are protein families that are widely evolutionarily conserved and present in eukaryotic, prokaryotic cells and even in some viruses. They are known to play a role in buffering cellular stress. Many different roles have been described in cellular processes, constitutively and/or in response to environmental stressors (heat, oxidative stress, chemical stressors), through mechanisms such as protein folding and unfolding, aggregation and disaggregation. They also show interactions with DNA repair mechanisms[15].
In 2003, Sex-specific expression of HSP72 the inducible form of HSP70, known as cardioprotectiv was showed in rat heart tissues. Myocardiac levels of HSP72 was measured with a sandwich ELISA in males, females and in ovariectomied females rats (model of menopause). Their results showed that female myocardic tissues expressed twice as much HSP72 as male. Menopause females rats had lower level of expression whereas menopause females with hormonal substitution therapy maintained their HSP72 female level. They concluded that oestrogen played an important role in the HSP72 regulation and that may explain a part of the sex dimorphic cardiac disease risk.

The aim of this project was to test if a sex difference in phenotypic robustness could be investigated through transcriptional data treatment. We did a whole-transcriptome screening looking for genes that were differentially expressed and regulated between sexes in lymphobalstoid cell lines (LCL). A more narrow transcriptional regulation of a gene alongside a smaller variance would indicate a more precise regulation and a possible stronger phenotypical robustness.

Heat-shock proteins may play a fundamental role in robustness processes, and this study tests the hypothesis of a potential differential constitutive expression between the sexes.

Immune-related genes could be more or less precisely regulated in one of the sex and if so it could contribute to robustness against infections or auto-immune disorders. The sex-dependant differences of level of expression will also be investigated as the LCL are not under hormonal control and it will allow us to see if some immune-related genes are constituvly sex-biased.

This screening was done with the RNA-sequencing gene expression data obtained in the CoLaus cohort which contains transcriptional data of Lymphoblastoid Cell lines (LCLs) from 550 people. LCLs are B-lymphocytes cells sampled from the peripheral blood and transformed by the Epstein Barr Virus (EBV) to become « immortalized »[16]. The virus genes are present in cells as episomal but are not thought to cause major changes in the gene expression of the infected cell.

Methodology

Data

Gene expression data was generated as RNAseq data from lymphoblastic cell lines from 260 men and 290 women from the CoLaus Study Cohort. This Cohort was sampled between 2003 and 2006 from inhabitants from Lausanne (CH) who were all caucasians, aged from 35 to 75 years[17].

The data was expressed in RPKM and in FPKM for the isoform data. RPKM is a unity of gene expression used in RNAseq which means "Reads Per Kilobase per million reads Mapped" and is used to normalize in function of the gene length and the sequencing depth. FPKM means "Fragments per Kilobase per million fragments Mapped" and is an analogue to RPKM with as difference that they use fragments and not reads.

Of the 44'000 expressed genes measured, 8'924 were selected. They were only the protein coding, and the ones that passed quality controls (exclusion of: the genes with a RPKM value <1 for more than 90% of the sampled, and genes and then of genes that have no continuous distribution in histogram). Y chromosome-specific genes were also excluded.

The list of 96 heat shock proteins was obtained from Kampinga *et al.*[18].

A list of 844 immunology-related genes was obtained from the InnateDB database on 6th October 2016 with the collected data being sourced from research articles, textbooks and electronic information sources [19].

Software

The data was analysed with R 3.2.2.

Methods

1. To identify the differentially expressed genes between sexes, statistical analyses were performed.
   For each gene the normality of the distribution of the samples in both groups was first tested with a Shapiro test, with the normality hypothesis rejected with $p<0.05$. The difference in expression for each gene between sexes was examined with a Wilcoxon test when the normality of the sample was rejected and a Welch test if not. Sex-biased expressed genes were selected as the ones with a tested p-value below the adjusted cut-off (Bonferroni correction: $p< 0.05/$number of tested genes).

2. Identification of genes with a different variance of expression between sexes was then performed. A Bartlett test was used to test for a difference of variance instead of a F-test, with an adjusted cut-off with the Bonferroni correction ($p< 0.05/$number of tests (genes)). H_0: the variances in each of the groups are the same.

3. We then tested if the difference in variance of expression between sexes was explained by a difference at the isoform level, in their expression pattern. For the genes with a significant difference in variance of expression, isoform expression data was generated.

   The Shannon entropy was used as an indication of disorder: gain or loss of entropy in the expression pattern of isoforms of a given gene would indicate a more or less precise regulation. It was also used to test if a difference in variance of gene expression between sexes could be explained by a different isoform pattern.

   For each gene presenting 2 or more isoforms and each individual, the Shannon's entropy of the isoform's expression distribution was calculated as a measure of regulation of the transcript pattern. For each gene and each individual entropy was calculated.

   H=entropy , X= gene, $P(x_i)$=frequency of the isoform x_i in a given sample.

$$H(X) = -\Sigma P(x_i) \log P(x_i)$$

   For each gene a Wilcoxon rank-sum test was performed to compare the entropy distribution between the two sexes.

4. Differences in expression or in variance of expression between sexes was then tested specifically for the heat shock protein genes.

   An analysis by "candidate gene" was then performed with the 52 HSP genes present in the data. The others 44 HSP genes did not pass quality controls. The first step of the method was applied to test for a difference in expression between sexes but with an adjusted cut-off of 0.05/52.

5. Differences in expression or in variance of expression between sexes was the tested specifically for immune-related genes with two methods:

   An analysis by "candidate gene" was then performed for the 320 immunology-related genes present in the data. The others 524 immune-related genes did not pass quality controls. The first and the 2nd step of the method was applied to test for a difference in expression between sexes but with an adjusted cut-off of 0.05/320.

Results

Genes differentially expressed between sexes

With a Wilcoxon rank-sum test or a Welch test depending on the normality of the sample and a p-value cut-off of 0.05/8'924, 254 in 8'924 genes were identified as differentially expressed between sexes, i.e. 3.003% of all the genes tested. Of the 268, 153 were male-biased and 115 female-biased. From the X chromosome 34 genes were differentially expressed with 7 male-biased genes and 27 female-biased genes.

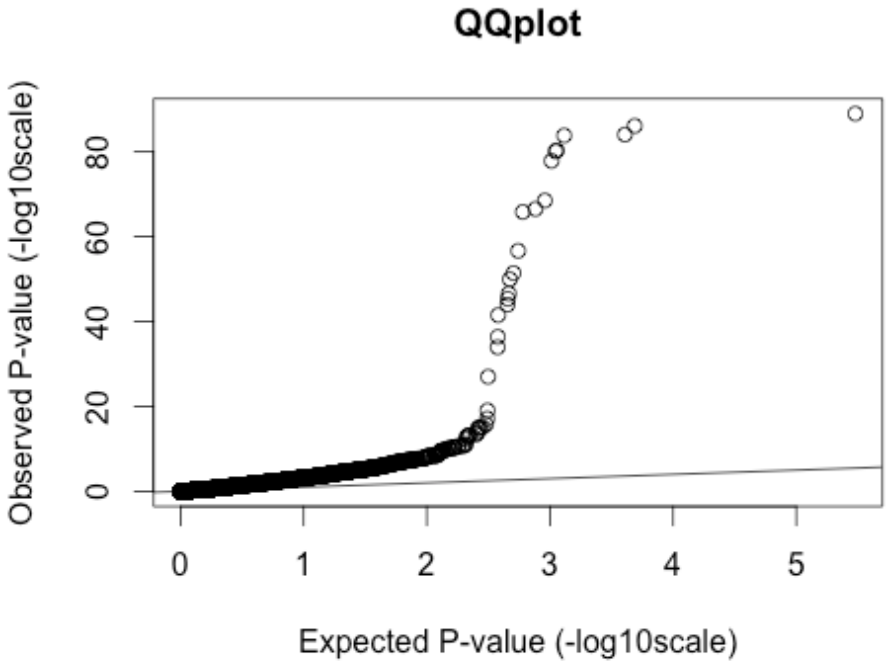

Figure 2 – QQplot of the expected and observed P-values of the test (Welch or Wilcoxon) of the difference in expression between sexes for the 8'924 tested genes.

Genes with a differential expression variance between sexes

With a Bartlett test of difference in variance, and a p-value cut-off of 0.05/8'924, 31 in 8'924 genes were identified as having a significant difference in variance in expression between sexes, i.e. 0.347% of all the genes tested. Amongst autosomal genes higher variance in expression was observed for males in 8 genes and in females in 10 genes. 13 X-genes showed a higher variance in women and none in men.

For the selected genes, autosomal and X genes were considered separately. In each case in order to determine the trend in the difference in variance, a t-test was then applied to the mean of the difference of the log-transformed variance of expression between the sexes. $H_0 : mean(\Delta) = 0$

$$\Delta = \log(Var(male\,\exp ression)) - \log(Var(female\,\exp ression))$$

In the 13 X-genes the hypothesis in a trend in difference of variance could be confirmed with a t-test (t = -14.006, df = 12, p-value = $8.488e^{-09}$)

In the 18 autosomal genes the hypothesis in a trend in difference of variance could not be confirmed with a t-test (t = 0.12942, df = 17, p-value = 0.8985).

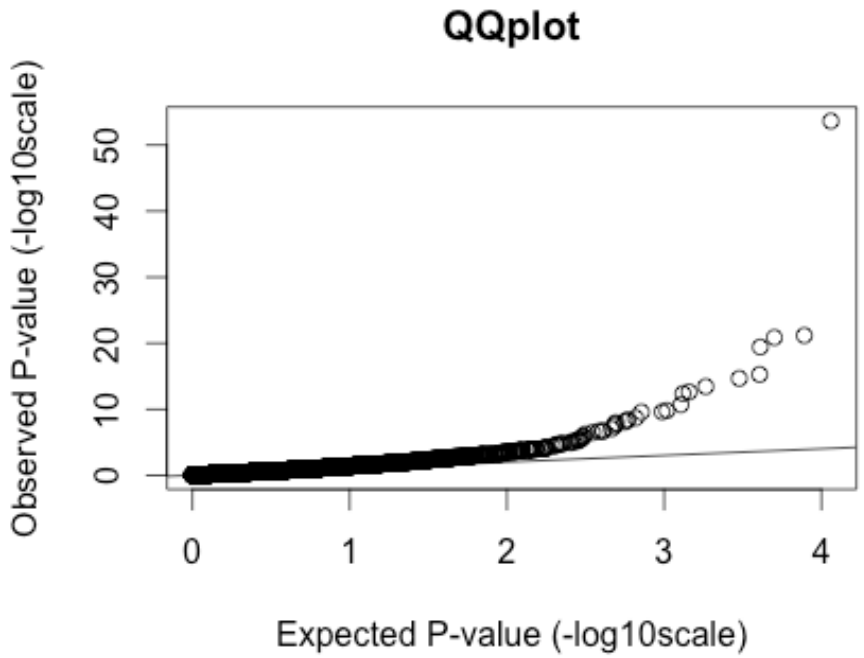

Figure 3- QQplot of the expected and observed P-values of the test of the difference in variance between sexes for the 8'924 tested genes.

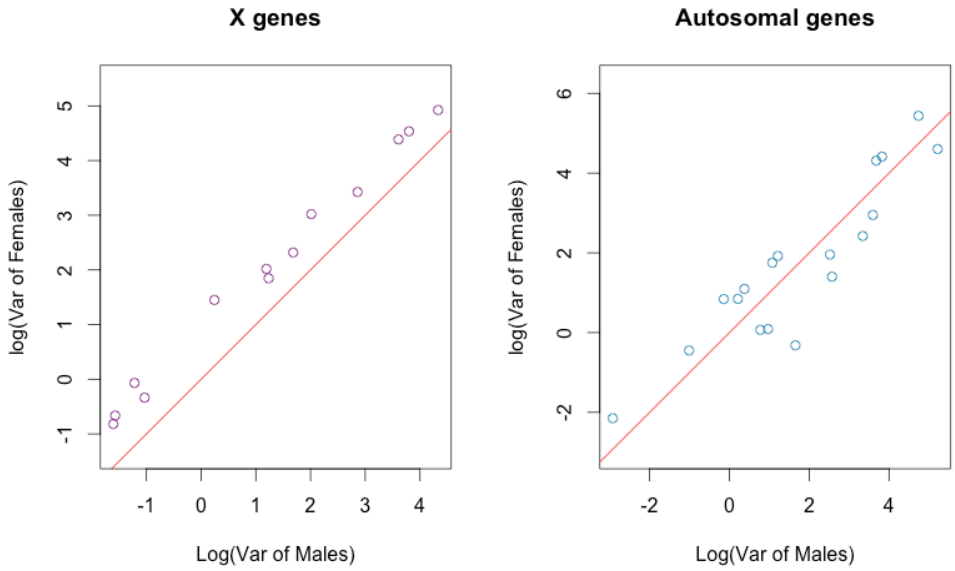

Figure 4- Graphs of the variance in expression between sexes for the 31 significant genes. On the left X genes, on the right Autosomal genes. Each dot corresponds to a gene. Log scale.

Gene	Protein -UniProtKB	Chr	P_value	var Males	var Females
MSL3	Male-specific lethal 3 homolog	X	3.7750e-06	17.4629	30.7141
ZFX	Zinc finger X-chromosomal protein	X	1.3356e-08	0.3558	0.7145
SLC39A9	Zinc transporter ZIP9	14	3.5987e-08	2.9332	5.7635
SMC1A	Structural maintenance of chromosomes protein 1A	X	5.5095e-16	7.4878	20.4955
P2RX5	P2X purinoceptor 5	17	1.2656e-08	114.2542	229.6768
STS	Steryl-sulfatase	X	2.0273e-07	5.3789	10.1626
ALG13	Putative bifunctional UDP-N-acetylglucosamine transferase and deubiquitinase ALG13	X	3.8276e-20	0.2956	0.9346
NRN1	Neuritin	6	9.1050e-08	36.4262	19.0966
SEPT6_	septin 6	X	2.4740e-10	36.8899	80.3904
GTF2F1	General transcription factor IIF subunit 1	19	1.3848e-21	13.0804	4.0645
EIF2S3	Eukaryotic translation initiation factor 2 subunit 3	X	1.2810e-06	76.1332	137.6128
WDR74	WD repeat-containing protein 74	11	2.2299e-54	5.2243	0.7266
PSAT1	phosphoserine aminotransferase 1	9	2.1694e-07	39.5697	74.6421
MMP7	matrix metallopeptidase 7	11	3.7543e-14	28.2662	11.2641
ZNF473	zinc finger protein 473	19	2.7412e-10	0.0536	0.1167
GPR174	Probable G-protein coupled receptor 174	X	5.4295e-07	3.4336	6.3377
HIST1H1E	Histone H1.4	6	4.2480e-13	2.6317	1.0915
EIF1AX	Eukaryotic translation initiation factor 1A, X-chromosomal	X	3.1292e-06	12.4157	7.0725
ZNF101	Zinc finger protein 101	19	2.1187e-11	3.2953	7.5267
FAM101B	Refilin-B	17	4.0727e-06	0.3636	0.6383
PRKX	cAMP-dependent protein kinase catalytic subunit PRKX	X	3.3180e-07	184.6806	99.6881
MAFF	Transcription factor MafF	22	6.8076e-22	1.2746	4.2637
CHM	Rab proteins geranylgeranyltransferase component A 1	X	2.2380e-07	1.2360	2.3298
MPHOSPH8	M-phase phosphoprotein 8	13	1.6179e-10	0.2006	0.4409
TRAPPC2	Trafficking protein particle complex subunit 2	X	8.2610e-09	3.3641	6.8246
MAFG	Transcription factor MafG	17	2.4254e-13	0.2078	0.5144
SLC25A29	Mitochondrial basic amino acids transporter	14	4.4339e-09	1.4552	2.9913
DDX3X	ATP-dependent RNA helicase DDX3X	X	5.3261e-09	2.1694	1.0707
TMSB4XP4	Pseudogene	9	2.1185e-09	44.6900	93.2786
NEDD8-MDP1	Protein NEDD8-MDP1	14	2.1759e-15	0.8653	2.3171
MAL	Myelin and lymphocyte protein	2	1.4389e-06	45.8105	82.5651

Table 1 Table of the 31 genes with a significant difference of variance between sexes. Gene and protein name Chr.: chromosome. P_value: P_value of the Bartlett test of difference in variance between sexes. Var Males: male variance. Var Females: female variance. In blue men higher variance, in red women-higher variance.

Isoform distribution pattern between sexes

23 of the 31 genes with a significant difference in variance between sexes had multiple isoforms. With a Wilcoxon rank-sum test and a P-value cutoff of 0.05/23, 5 genes were selected as having a significant difference in entropy distribution between sexes. 4 of the 5 were X-linked and only one, SLC39A9 was an autosomal.

Gene	P_value	Median Male entropy	Median Female entropy
SLC39A9	0.000473	1.1318	1.0487
SMC1A	6.6357e-05	0.22473	0.2903
STS	6.4788e-05	1.02599	1.0592
TRAPPC2	2.0211e-05	1.17118	1.2191
DDX3X	0.001407	0.84592	0.9138

Table 2 Isoforms expression pattern analysis. P_value: p_value of the Wilcoxon test comparing the male and female entropy distribution for one gene.

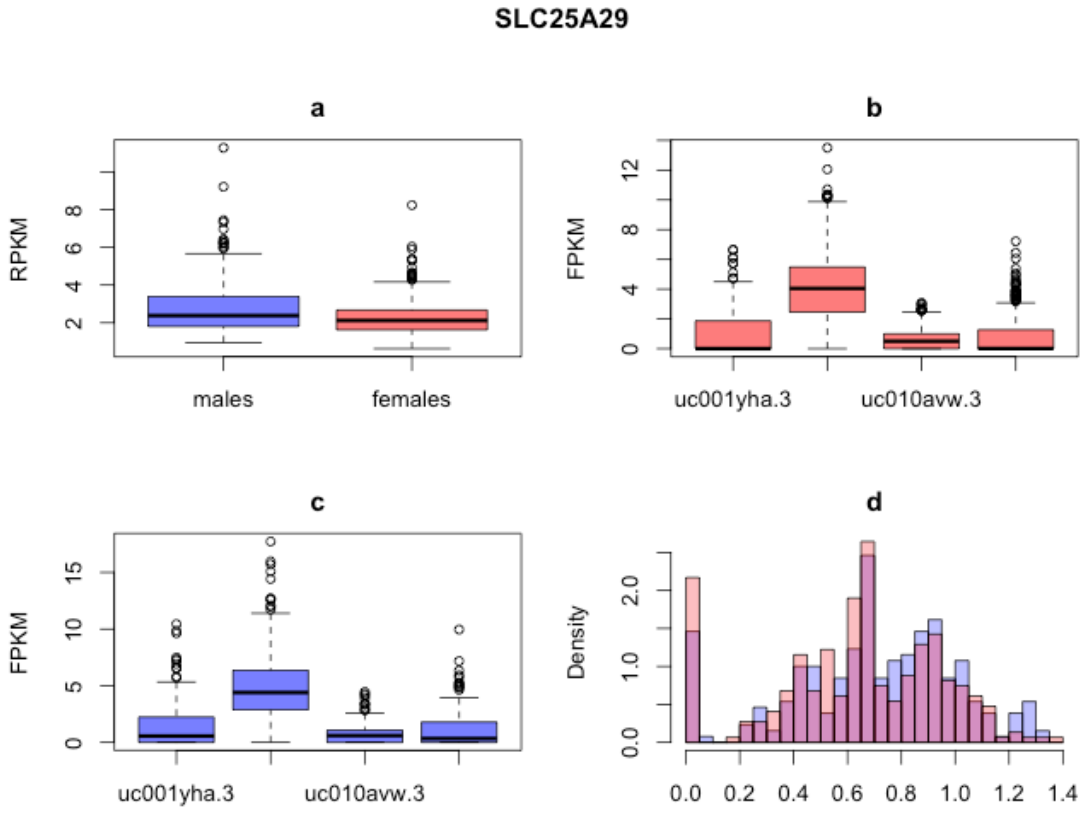

Figure 5 Expression of SLC25A29 gene and its Isoforms. a : expression in men and women. b isoforms expression in females. c: isoforms expression in males. d: distribution of entropy of the isoforms expression between sexes (blue = men , red=women).

Heat shock proteins

Of the 52 HSP tested genes, only 1 was present in the list of the genes with a significant difference in variance.

The "candidate gene" approach with an adjusted p-value cut-off of 0.05/52 showed 4 autosomal genes with a significant difference in gene expression between sexes, with 3 being female-biased and 1 male-biased.

Gene and protein name	Chr.	P_value	Males [RPKM]	Females [RPKM]	Difference Males-Females
DNAJC1	10	2.6345e-07	7.2977	7.898	-0.6003
DNAJC18	5	7.6795e-06	1.52455	1.7537	-0.22915
DNAJB5	9	9.8980e-05	2.2653	2.073	0.1923
HSPA13b	21	6.2204e-06	16.361	17.714	-1.353

Table 3- Table of the 4 HSP genes with a significant difference of expression between sexes. Gene and protein name Chr.: chromosome. P_value: P_value of the Wilcoxon test of expression between sexes. Males: male median expression [RPKM]. Females: female median expression [RPKM]. Difference: difference of the males median and the females median of expression [RPKM] in blue men-biased genes, in red women-biased genes.

HSPA13 also named *STCH* is part of the HSPA (or HSP70) family that all contain a ATPase domain. *HSPA13* is known as an intracellular, microsomal-associated HSP, constitutively expressed in different amounts in all tested human tissues[20 21].

DNAJB5, DNAJC1 and DNAJC18 are part of the DNAJ HSP family (HSP40) that all have a conserved J domain whose role is to regulate the HSPA proteins activity through interaction with their ATPase domains.

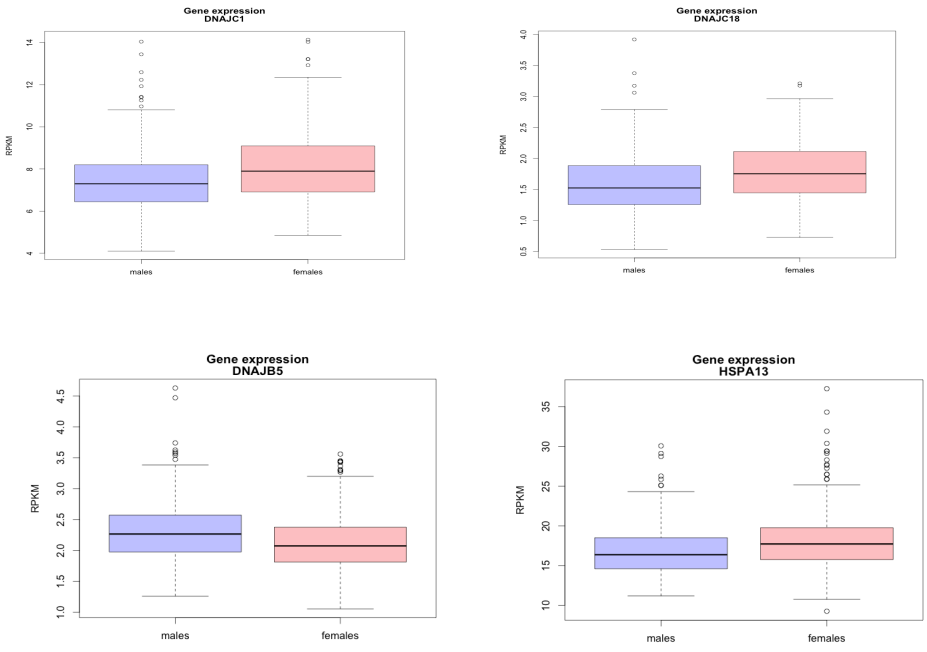

Figure 6- Gene expression of the 4 sex-biased HSP genes

Immune-related genes

The overlap between the 844 immunology-related genes list and our list of genes was 320 genes. Of the 320 tested genes, none were present in the list of the genes with a significant difference in variance. 7 were present in the list of genes with a significant difference in expression. The "candidate gene" approach with an adjusted p-value cut-off of 0.05/320 showed 23 autosomal genes with a significant difference in gene expression between sexes, with 15 being female-biased and 8 male-biased. The significance of the result of higher expression in females (16 out of 22) was tested with a binomial test and was close to significance (p = 0.05248). The "candidate gene" approach did not show any significant immune-related genes with a significant difference in variance between sexes.

Gene	Protein	Chr.	P_value	Males	Females	Diff M-F
IFNAR2	Interferon alpha/beta receptor 2	21	5.8235e-05	24.683	26.468	-1.785
TRAF7	E3 ubiquitin-protein ligase TRAF7	16	2.3213e-05	21.1255	20.469	0.6565
CD38	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 1	4	8.5909e-05	34.1345	36.789	-2.6545
CIITA	MHC class II transactivator	16	2.2583e-06	9.7344	10.743	-1.0086
ITGAE	Integrin alpha-E	17	3.3950e-08	1.51535	1.6556	-0.1402
PSTPIP1	Proline-serine-threonine phosphatase-interacting protein 1	15	2.0968e-05	2.09285	2.2904	-0.1975
PIK3CG	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma isoform	7	3.4083e-07	3.17505	3.6919	-0.5168
NFATC3	Nuclear factor of activated T-cells, cytoplasmic 3	16	8.3925e-08	10.7915	11.285	-0.4935
IRF2	Interferon regulatory factor 2	4	1.4640e-06	32.157	33.473	-1.316
CD74	HLA class II histocompatibility antigen gamma chain	5	0.00011644	623.095	661.36	-38.265
LCP2	Lymphocyte cytosolic protein 2	5	6.2752e-06	3.58505	3.2673	0.3177
CD300A	CMRF35-like molecule 8	17	4.6643e-05	41.779	37.643	4.136
CASP1	Caspase-1	11	6.7978e-09	6.82675	5.5296	1.2971
CXCR4	C-X-C chemokine receptor type 4	2	2.2257e-08	35.286	40.444	-5.158
CASP10	Caspase-10	2	7,7773e-05	4.81115	5.1112	-0.30005
SLC44A1	Choline transporter-like protein 1	9	1,7459e-06	5.68	6.3041	-0.6241
HLA-DMA	HLA class II histocompatibility antigen, DM alpha chain	6	6.5070e-06	37.727	40.58	-2.8529
ACKR3	Atypical chemokine receptor 3	2	3.7391e-06	16.788	12.905	3.883
IFNLR1	Interferon lambda receptor 1	1	7.2110e-07	3.30035	3.7966	-0.4962
IGSF8	Immunoglobulin superfamily member 8	1	3.0528e-10	6.07315	6.6738	-0.6006
LAX1	Lymphocyte transmembrane adapter 1	1	4.1704e-05	3.23105	4.099	-0.86795
NDUFS3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial	11	5.1781e-05	15.457	14.957	0.5

Table 2 – Table of the 22 immune-related genes with a difference of expression between sexes. Gene: name of the gene. Protein: protein coded by the gene. Chr.: chromosome. P_value: P_value of the Wilcoxon test in expression between sexes. Males: male median expression [RPKM]. Females: female median expression [RPKM]. Diff M-F: difference of the males median and the females median of expression [RPKM] in blue men-biased genes, in red women-biased genes.

Discussion

Genes differentially expressed between sexes

The result of 3.003% of all the tested genes with a significant difference is very close to the 3.2% found by another study where RNAseq from the Geuvadis consortium was analysed [22]. Sex-biased genes on autosomes do not show a trend of sex-bias. That does not exclude that some of the genes could participate in some robustness processes.

For a part of the X-linked genes the female-biased gene expression could be explained by the fact that about 15% of the X-linked genes[23] seem to escape the X-inactivation process. A study showed that there could be variation genes escaping between individuals in this process[24].

Very small differences in mean or median expression between the sexes could become statistically significant with small effect size leading to the risk of no real biological effect in the cell. And this raises the question of which level of difference of RPKM is needed to really lead to a biological effect? It is probably protein specific.

Genes with a differential expression variance between sexes and Isoforms analysis

In autosomal genes, no significant trend in variance difference between sexes was found which does not give any arguments for a more generalised precise regulation in any sex. However, some of the 31 genes could still contribute individually to processes in robustness. Isoforms analysis did not show great differences in expression pattern between sexes and can probably not explain the differences in variance of expression even if there is on gene with a significant difference in isoforms distribution pattern between sexes.

Differences in expression variance between the sexes in X-linked genes could also be explained by the escape from the X-inactivation.

The Bartlett test used is rather sensitive to departures from normality of the samples and it might have been a better option to use, for example a Levene's test. However, the P-value QQplot suggest that the tests assumption were respected and the data independent.

Heat shock proteins

The four HSP genes with a difference in expression between sexes have been partially replicated by Perrine Steffe in another student project on the Geuvadis cohort, which is a good indication of the robustness of the result. Out of the four genes, one was replicated as significance level and the three other showed the same trend.

However, the effect size is small, ranging from 0.19 to 1.3 RPKM and this raises the question of the biological effects of theses differences in expression. The constitutive level of HSP expression in the LCL cells was tested and it would be interesting to test if theses differences in expression represent a real effect of higher robustness against stressor between sexes .

Differences in tissue expression of HSP genes seem to occur and as an example, DNAJB5 seems to have a quiet reduced expression in LCL than in the heart (mean 17.5RPKM) or brain tissue (mean 13.1 RPKM)[25]. Indeed tissue-specific differences in expression and the kinetics of the induction of HSP could also be interesting to investigate and compare between sexes.
The HSP transcription factors named heat-shock factors (HSFs) seem to play an important role in the HSP expression regulation by binding to promoters, heat shock elements, upstream of the HSP genes[26]. They are known to show tissue-specific patterns of expression, to interact with other mediators and could play an important role not only in stress response but also in development [10,26,]. Their constitutive levels and response to stressors could also be compared between sexes.

Immune-related genes

There was no immune-related gene with a significant difference of variance of expression. This mean that in LCL theses analyse can't show a difference of regulation and thus robustness between sexes.
22 sex-biased immune-related genes were selected, mainly female-biased, which give an indication that even without any external stimulus there are differences in immune-genes expression. Being more expressed does not mean that the immune reaction would be more efficient, some genes can be overexpressed but they can be at the biological level inhibitory of the immune–system or present some complexes pleiotropic effects. As an example, the CD300A gene coding for the CMRF35-like molecule 8 was men-biased in the analyse is an inhibitory receptor that may be involved in Natural Killer cell and mast cell degranulation down regulation[27].
As the LCL show a kind of constitutive state, without external stimulus it could not be the right method to investigate the immune-system differences. Indeed some differences in regulations could exist only in Vivo, due to sex hormones or other systems. It would be interesting to test it in different tissues ton incorporate in the analyse complexity off the immune system and its interactions with other systems as hormones, and even microbioma.

Conclusion

Some HSP levels could differ between sexes and play a role in phenotypic robustness process. This study showed small but significant sex-biased expression of HSP in LCL and further analyses of other tissues should be conducted as some HSP might also be regulated by hormonal signals. At the regulation level this RNAseq all-transcriptome screening for sex-specific patterns of regulation did not show a sex-biased trend. Isoforms analysis between sex did not show significant differences in pattern distribution. This mean that the hypothesis of a general sex-specific pattern of regulation leading to a differential robustness mechanisms is unlikely, although the sex-biased expression of specific genes could still play a role. An alternative option would be that LCL could also not be appropriate to measure it.

Other mechanisms as miRNA, which are non-coding RNAs that might play a role in robustness process with a buffering effect on gene expression could also be investigate between sexes[28].

References

[1] Fombonne, (1999). The epidemiology of autism: a review. *Psychol Med*, Jul;29(4):769-86.

[2] Mcgrath, J., Saha, S., Chant, D., & Welham, J. (2017). Schizophrenia : A Concise Overview of Incidence , Prevalence , and Mortality, *Epidemiologic Reviews, 30*(December), 67–76.

[3] Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., … Eichler, E. E. (2014). A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. *American Journal of Human Genetics*, *94*(3), 415–425.

[4] Whitacre, C. C. (2001). Sex differences in autoimmune disease. *Nature Immunology*, *2*(9), 777–80. http://doi.org/10.1038/ni0901-777

[6] Kruger, D. J., & Nesse, R. M. (2006). An evolutionary life-history framework for understanding sex differences in human mortality rates. *Human Nature*, *17*(1), 74–97.

[6] Clutton-Brock, T. H., & Isvaran, K. (2007). Sex differences in ageing in natural populations of vertebrates. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1629), 3097–3104.

[7] Folstad, I., & Karter, A. J. (1992). Parasites, Bright Males, and the Immunocompetence Handicap. *The American Naturalist*, *139*(3), 603–622. http://doi.org/10.1086/285346

[8] Muehlenbein, M. P., & Bribiescas, R. G. (2005). Feature Article Testosterone-Mediated Immune Functions and Male Life Histories. *American journal of human biology, 558*(December 2004), 527–558.

[9] Nunn, C. L., Lindenfors, P., Pursall, E. R., & Rolff, J. (2009). On sexual dimorphism in immune function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1513), 61–69.

[10] Wedekind, C., & Jakobsen, P. (1998). *Male-Biased Susceptibility to Helminth Infection: An Experimental Test with a Copepod*. *Oikos* (Vol. 81).

[11] Wade MJ (1979). Sexual selection and variance in reproductive success. *Am Nat* 114: 742–746.

[12] Mallet, M. A., & Chippindale, A. K. (2010). Inbreeding reveals stronger net selection on Drosophila melanogaster males : implications for mutation load and the fitness of sexual females. *Heredity*, *106*(6), 994–1002.

[13] Rutherford, S. L., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution, *Nature, 396*(November).

[14] Rutherford, S. L., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution, *Nature, 396*(November).

[15] Sottile, M. L., & Nadin, S. B. (2017). Heat shock proteins and DNA repair mechanisms: an updated overview. *Cell Stress and Chaperones*.

[16] Hussain, T., & Mulherkar, R. (2012). Lymphoblastoid Cell lines: a Continuous in Vitro Source of Cells to Study Carcinogen Sensitivity and DNA Repair. *International Journal of Molecular and Cellular Medicine*, *1*(2), 75–87.

[17] Firmann, M., Mayor, V., Vidal, P. M., Bochud, M., Pécoud, A., Hayoz, D., … Vollenweider, P. (2008). The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovascular Disorders*, *8*(1), 6.

[18] Kampinga, H. H., Hageman, J., Vos, M. J., Kubota, H., Tanguay, R. M., Bruford, E. A., … Hightower, L. E. (2009). Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress and Chaperones*, *14*(1), 105–111.

[19] Breuer, K., Foroushani, A. K., Laird, M. R., Chen, C., Sribnaia, A., Lo, R., … Lynn, D. J. (2017). InnateDB : systems biology of innate immunity and beyond — recent updates and continuing curation, *Nucleic Acids Research ,41*(November 2012), 1228–1233.

[20] Otterson, G. A., Flynn, G. C., Kratzkel, R. A., Coxon, A., Johnston, P. G., & Kaye, F. J. (1994). Stch encodes the 'ATPase core' of microsomal stress7O protein, *The EMBO Journal,* vol.13*13*(5), 1216–1225.

[21] Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., … Uhlén, M. (2014). Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Molecular & Cellular Proteomics : MCP*, *13*(2), 397–406.

[22] Shen, J. J., Wang, T.-Y., & Yang, W. (2017). Regulatory and evolutionary signatures of sex-biased genes on both the X chromosome and the autosomes. *Biology of Sex Differences*, *8*, 35.

[23] Carrel, L., & Willard, H. F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, *434*, 400. Retrieved from http://dx.doi.org/10.1038/nature03479

[24] Zhang, Y., Castillo-Morales, A., Jiang, M., Zhu, Y., Hu, L., Urrutia, A. O., … Hurst, L. D. (2013). Genes That Escape X-Inactivation in Humans Have High Intraspecific Variability in Expression, Are Associated with Mental Impairment but Are Not Slow Evolving. *Molecular Biology and Evolution*, *30*(12), 2588–2601.

[25] Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., … Uhlén, M. (2014). Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Molecular & Cellular Proteomics : MCP*, *13*(2), 397–406.

[26] Stephanou, A., & Latchman, D. S. (2011). Transcriptional Modulation of Heat-Shock Protein Gene Expression. *Biochemistry Research International*, *2011*, 238601.

[27] Kim, E.-J., Lee, S.-M., Suk, K., & Lee, W.-H. (2012). CD300a and CD300f differentially regulate the MyD88 and TRIF-mediated TLR signalling pathways through activation of SHP-1 and/or SHP-2 in human monocytic cell lines. *Immunology*, *135*(3), 226–235.

[28] Siciliano, V., Garzilli, I., Fracassi, C., Criscuolo, S., Ventre, S., & di Bernardo, D. (2013). MiRNAs confer phenotypic robustness to gene networks by suppressing biological noise. *Nature Communications*, *4*, 2364.