The Influence of Sex on Mutational Processes in Pediatric Cancers: Preliminary Results

Jieun Oh[1,4,9,10], Constance H. Li[2,3,4], Brandon L. Tsai[4,5,9,10], Paul C. Boutros[3,4,6,7,8,9,10]

[1]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region. [2]Computational Biology Program, Ontario Institute for Cancer Research, Toronto, ON, Canada. [3]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada. [4]Department of Human Genetics, University of California, Los Angeles, CA, USA. [5]David Geffen School of Medicine, Medical Scientist Training Program, University of California, Los Angeles, CA, USA. [6]Department of Pharmacology & Toxicology, University of Toronto, Toronto, ON, Canada. [7]Vector Institute for Artificial Intelligence, Toronto, Canada. [8]Department of Urology, University of California, Los Angeles, CA, USA. [9]Jonsson Comprehensive Cancer Center, University of California, Los Angeles, CA, USA. [10]Institute for Precision Health, University of California, Los Angeles, CA, USA.

Overseas Research Fellowship (ORF) 2020 for Science Students

Oh Jieun
Major in Biochemistry
Poster No. E5
External Supervisor:
Dr. Paul C. Boutros
Internal Supervisor:
Dr. Jason W.H. Wong

INTRODUCTION

Sex differences have been observed in various diseases, in the form of epidemiology, prognosis, susceptibility to disease and other factors (1,2). Cancer is not an exception and it has been observed that males generally have a higher incidence and death rate than women for all cancers combined (3). However, these differences change depending on cancer type and women have higher incidence rates in some cancers, such as thyroid cancer (3). The reasons for these differences are complex, but social behaviors that are more common in one sex, such as smoking and alcohol-consumption, differences in hormone levels, sex chromosomes and sex-biased molecular disparities are commonly thought to be the causes of these sex-biases (4). However, studies have shown that sex differences in survival still prevail even after adjusting for external risk factors, suggesting that sex differences in biology play an important role (5,6). Although studies on sex differences on a genetic level have been carried out in adult cancers, sex differences in childhood cancers have not been documented as thoroughly (7,8). Therefore in this study, sex differences in childhood cancers are being studied, including acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), wilms tumour (WT), rhabdoid tumour (RT), neuroblastoma (NBL) and osteosarcoma (OS).

METHODOLOGY

Data:

- Somatic single nucleotide variants, RNA and copy number data from Therapeutically Applicable Research to Generate Effective Treatments (TARGET) were used in this study (9).

Single Nucleotide Variant (SNV):

- To see if there was a sex bias in the presence of non-synonymous SNVs in genes, each sample was given a binary value depending on whether it had a non-synonymous SNV in a gene. Proportion tests were used to see if there was a difference. P-values were then FDR-adjusted to account for multiple hypothesis testing. Genes were also filtered at a higher threshold where only genes with three or more samples with SNVs or five more samples with SNVs in either male or female samples were selected.
- SNV burden on a sample level was investigated next by calculating the sum of non-synonymous SNVs in a sample and using a Mann-Whitney U test.

Genome Instability:

- Proportion of the genome with a copy number aberration (PGA) was analysed on autosomal chromosomes using Mann-Whitney U tests. This was repeated on a chromosomal level for autosomal chromosomes where the p-values were FDR-adjusted.

Copy Number Aberration (CNA):

- Analysis on copy number aberrations was performed using proportion tests and p-values were FDR-adjusted.

RNA:

- RNA-seq data, normalized by FPKM-UQ, were analysed by performing Mann-Whitney U tests. Since the results are only preliminary, more work on RNA-seq needs to be done as some genes have several samples which have zero as their FPKM-UQ value and additional analyses will be done to confirm the results.

Additional Notes:

- All analyses were done on R statistical environment (v3.6.3). Plots were produced using the BPG (v6.0.1) (10).
- Not all tumour types had harmonized data (or available data) hence they will be analysed separately.

RESULTS

Single Nucleotide Variant (SNV)

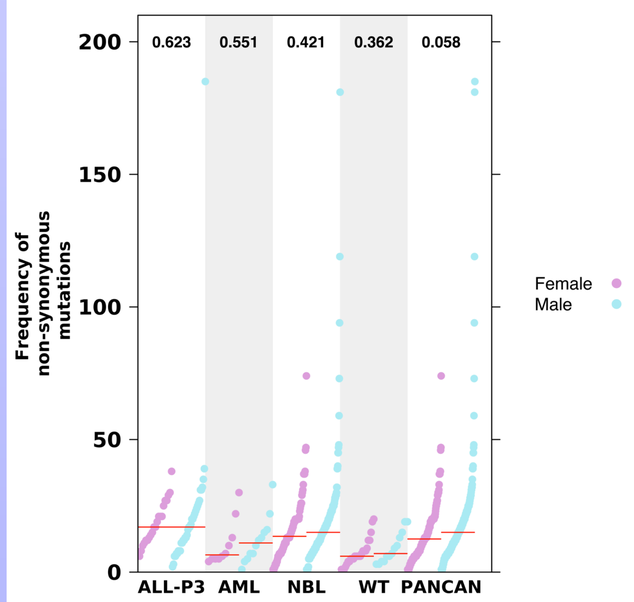

Figure 1. Total number of non-synonymous SNVs per sample in each tumour type.
The black numbers at the top of the plot indicate the p-values. There were no significant results in this analysis. Red lines indicate the median value.

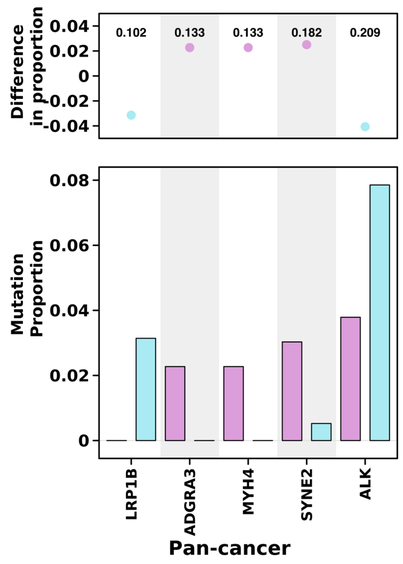

Figure 2. The top plot shows the difference in proportions of several genes as named at the bottom. Similarly, the bottom bar plot shows the proportion of samples containing a non-synonymous SNV for the corresponding gene in each sex. The numbers at the top are adjusted p-values. There were no significant results in this analysis.

RNA-seq

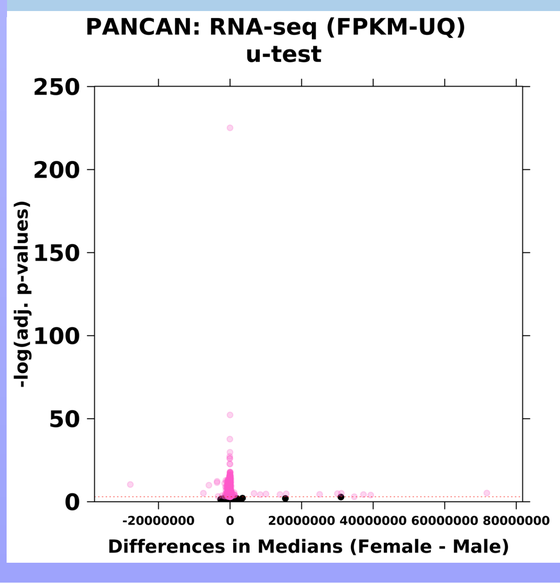

Figure 3.
Volcano plot for pan-cancer (ALL-P1, ALL-P2, ALL-P3, AML, CCSK, NBL, OS, RT) RNA-seq. Transparent pink points indicate significant results. Note that many insignificant genes are also present, but may look hidden as the pink points become more concentrated.

Dotted line indicates –log(0.05). 0.05 is the cutoff point for adjusted p-values.

Several genes have been identified as significant across tumour types. Therefore, further analyses will be carried out for confirmation.

RESULTS

Copy Number Aberration and Proportion of the Genome with a Copy Number Aberration

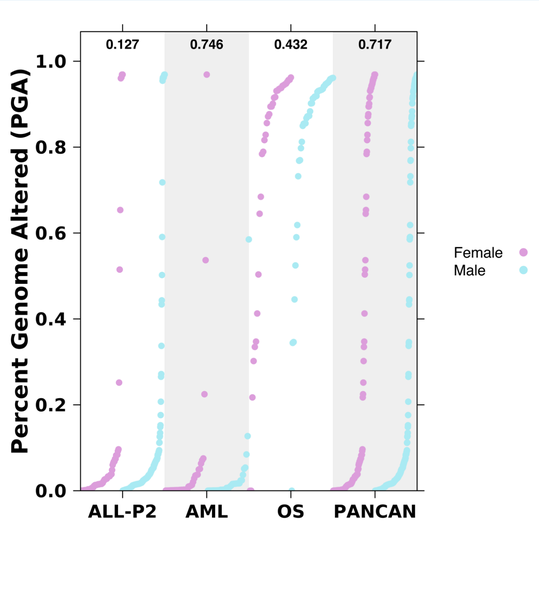

Figure 4.
Scatter plot of proportion of the genome with a copy number aberration (PGA) for each available tumour type. Black letters at the top are p-values.

No significant results from autosomal chromosomes together and individually.

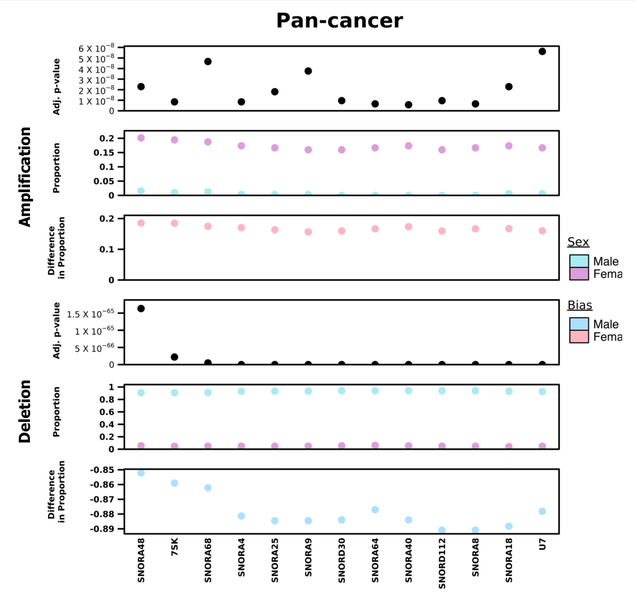

Figure 5.
Adjusted p-values, proportions and differences in proportions for some significant genes in pan-cancer copy number aberrations (adj p. <0.05), once for amplification and once for deletion. However, this is not a comprehensive plot as many genes, including pseudogenes have been filtered out.

Several genes have been identified as significant across tumour types. These genes will be investigated in more detail.

Tumour types and Number of Samples used for SNVs and PGA

Sex\Tumour Subtype	ALL-P3	AML	NBL	WT	PANCAN
Female	21	10	76	25	132
Male	35	12	131	13	191

Table 1. Number of samples for SNV analyses (SNV burden and proportions)

Sex\Tumour Subtype	ALL-P2	AML	CCSK	OS	PANCAN
Female	72	37	0	35	144
Male	221	36	11	46	314

Table 2.
Number of samples for proportion of the genome with a copy number aberration (PGA) analyses

DISCUSSION

- No significant results (< 0.05) have been found in the proportion of the genome with a copy number aberration for autosomal chromosomes, both together and individually; non-synonymous SNV burden per sample; proportions of non-synonymous SNVs on a gene-level after adjusting the p-values. However, this may be due to the small sample sizes, as indicated in Table 1 and 2 above, leading to low statistical power. Increasing the sample sizes may yield a different result.

- Some significant results have been obtained in RNA-seq data and copy number aberrations. However, further analyses using different methods are needed validate these results, especially for RNA-seq. It has been planned that FPKM-UQ values will not be used in the next method. Moreover, some samples in several genes have been identified to contain zero as their FPKM-UQ values so this will be accounted for in our next analyses. Also, it is worthwhile to remember that some adjusted p-values, which are insignificant for certain genes, may be due to multiple testing burden.

CONCLUSION

Bioinformatics studies to investigate sex differences in pediatric cancers were carried out on various types of data available through TARGET. Confounding variables, such as age, must be accounted for and remaining datasets will be analysed. Studying molecular sex differences is important because it can provide insight into why women's and men's tumour are different (8). This knowledge can then be transferred to the clinic to improve the efficacy of treatment (8). Therefore, this research has the potential to improve precision medicine.

ACKNOWLEDGEMENTS

The analyses were based on data that was produced by the Therapeutically Applicable Research to Generate Effective Treatments (https://ocg.cancer.gov/programs/target) initiative, phs000218. The data utilized in this project can be found on https://portal.gdc.cancer.gov/projects. Furthermore I would like to acknowledge and thank the patients and their families for making the datasets available and making this research possible.

REFERENCES

1. Mauvais-Jarvis F, Merz NB, Barnes PJ, Brinton RD, Carrero JJ, DeMeo DL, De Vries GJ, Epperson CN, Govindan R, Klein SL, Lonardo A. Sex and gender: modifiers of health, disease, and medicine. The Lancet. 2020 Aug 22;396(10250):565-82.
2. Golden LC, Voskuhl R. The importance of studying sex differences in disease: The example of multiple sclerosis. Journal of neuroscience research. 2017 Jan 2;95(1-2):633-43.
3. Siegiel R, Miller K, Jemal A. Cancer statistics, 2017. CA Cancer J Clin. 2017;67(1):7-30.
4. Zhu Y, Shao X, Wang X, Liu L, Liang H. Sex disparities in cancer. Cancer letters. 2019 Dec 1;466:35-8.
5. Wisnivesky JP, Halm EA. Sex differences in lung cancer survival: do tumors behave differently in elderly women?. Journal of clinical oncology. 2007 May 1;25(13):1705-12.
6. OuYang PY, Zhang LN, Lan XW, Xie C, Zhang WW, Wang QX, Su Z, Tang J, Xie FY. The significant survival advantage of female sex in nasopharyngeal carcinoma: a propensity-matched analysis. British journal of cancer. 2015 Apr;112(9):1554-61.
7. Li CH, Haider S, Shiah YJ, Thai K, Boutros PC. Sex differences in cancer driver genes and biomarkers. Cancer research. 2018 Oct 1;78(19):5527-37.
8. Li CH, Prokopec SD, Sun RX, Yousif F, Schmitz N, Boutros PC. Sex differences in oncogenic mutational processes. Nature communications. 2020 Aug 28;11(1):1-24.
9. Ma X, Liu Y, Liu Y, Alexandrov LB, Edmonson MN, Gawad C, Zhou X, Li Y, Rusch MC, Easton J, Huether R. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. Nature. 2018 Mar;555(7696):371-6.
10. P'ng C, Green J, Chong LC, Waggott D, Prokopec SD, Shamsi M, Nguyen F, Mak DY, Lam F, Albuquerque MA, Wu Y. BPG: Seamless, automated and interactive visualization of scientific data. BMC bioinformatics. 2019 Dec;20(1):1-5.