Identification of alternative splicing alterations in small cell lung cancer

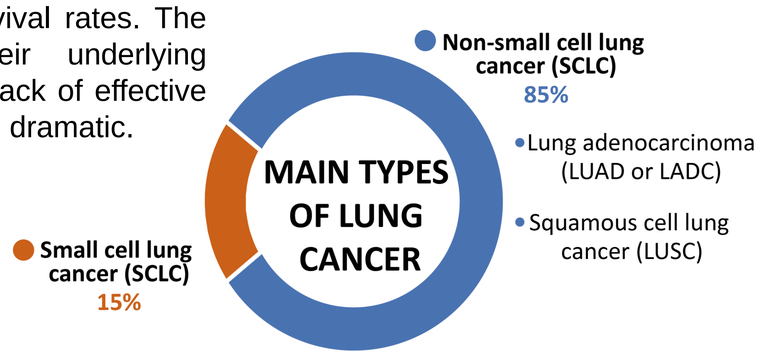
Marina Reixachs[1], Jun Yokota[2] and Eduardo Eyras[1,3]

[1]Universitat Pompeu Fabra (UPF) [2]Cancer Genome Biology, Institut de Medicina Predictiva i Personalitzada del Càncer (IMPPC) [3]Institució Catalana de Recerca i Estudis Avançats (ICREA)

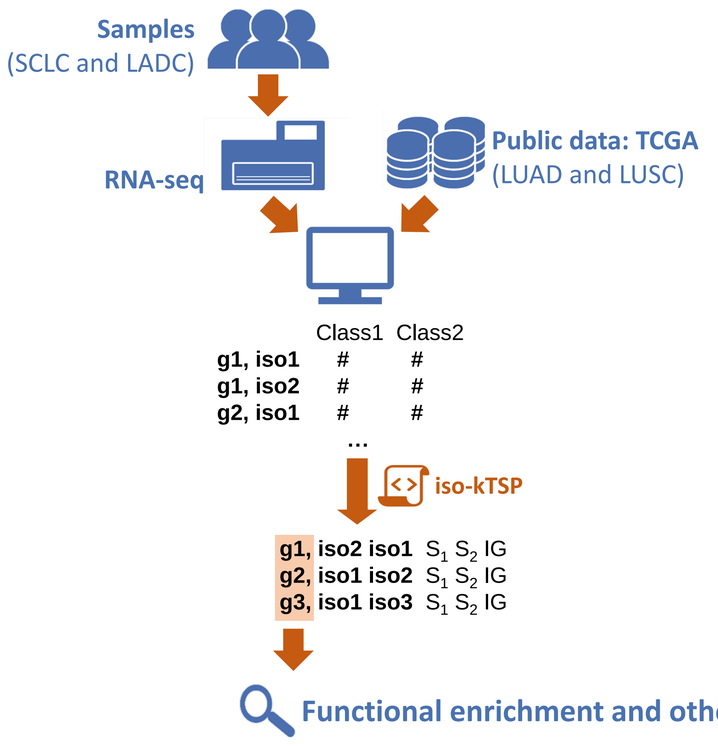
INTRODUCTION

Small cell lung cancer is the most aggressive type of lung cancer, showing poor survival rates. The lack of knowledge about their underlying tumorigenic mechanisms and the lack of effective treatments make the situation more dramatic.

Our main goal is to obtain a specific splicing signature for SCLC that may provide novel molecular targets for prognosis and therapy.



METHODOLOGY

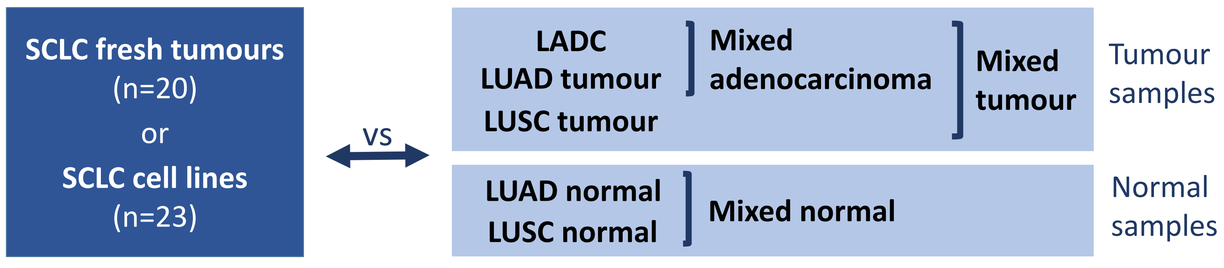


The iso-kTSP algorithm

It is a comparison-based approach able to identify transcript isoform changes across multiple samples from two different classes. The algorithm provides a score (S_1) for each possible pair of isoforms in each gene based on the consistency of the change in the relative ordering in a ranking of isoform expression values between the two classes. Additionally, a second score (S_2) and information gain (IG) values are provided. Finally, the isoform-pair with highest score value is selected for each gene in the final ranking.

Figure 1. We applied the iso-kTSP algorithm to RNA-seq data from different lung cancers, expressed in transcripts per million (TPM). The obtained list of aberrantly spliced genes was used to perform functional enrichment analyses based on Gene Ontology terms (GOrilla) and further information extraction.

We performed comparisons of SCLC samples from fresh tumours or cell lines against other eight different data sets. Analyses were performed using balanced groups of samples, with an equal number of samples from each class. In order to avoid bias due to sample selection, three random input files for each comparison were generated.



In order to delimit the values of S_1 score and IG resulting from analyses between similar samples, we run the algorithm comparing sets of normal samples, from LUAD and LUSC, which should not present differences.

ACKNOWLEDGEMENTS

We would like to thank Dr Jun Yokota and Dr Takashi Kohno for providing us data for SCLC and LADC samples. I also want to specially thank my supervisor, Dr Eduardo Eyras, for being patient and helpful when solving my doubts. Finally, I want to mention all the colleagues and friends that have contributed to this work and did not hesitate in helping me or providing me useful advice.

RESULTS

We identified **510** unique genes that show differential transcript expression in SCLC compared to other lung cancers or normal tissues. Results from different analyses were consistent and presented high score and IG values. S_1 values were greater than 0,85, many of them reaching the maximum value of 1. Results from analyses comparing similar samples showed lower scores (0,3-0,6) that confirmed the reliability of the obtained values for SCLC analyses.

.

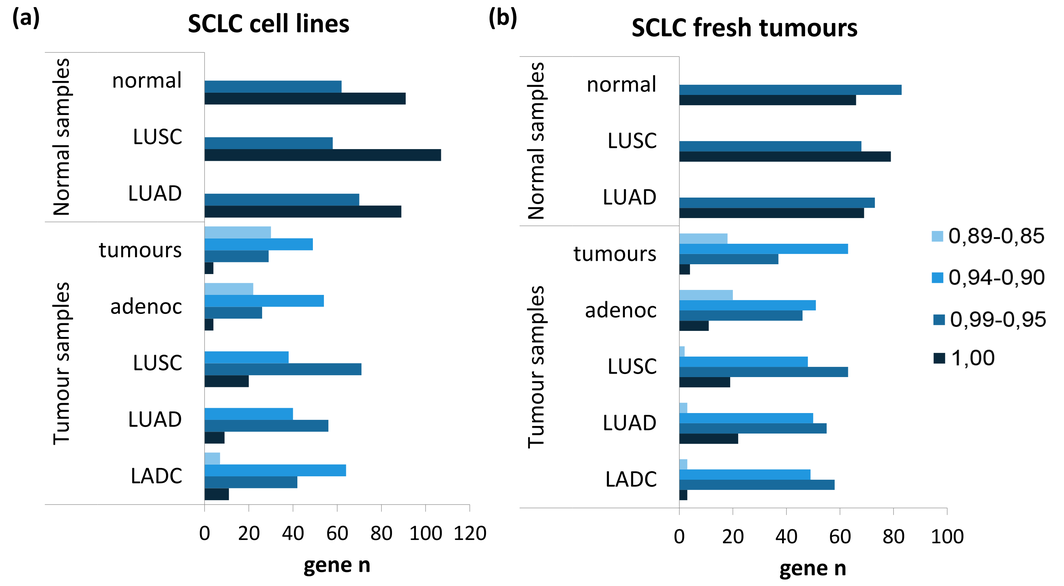


Figure 2. Number of genes presenting isoform switches with an average score value within the indicated ranges for every comparison between SCLC cell lines (a) or fresh tumours (b) against different data sets.

We also obtained a list of **49** genes for fresh tumours and **46** genes for cell lines that appeared recurrently as a result of at least 7 out of the total 8 performed analyses.

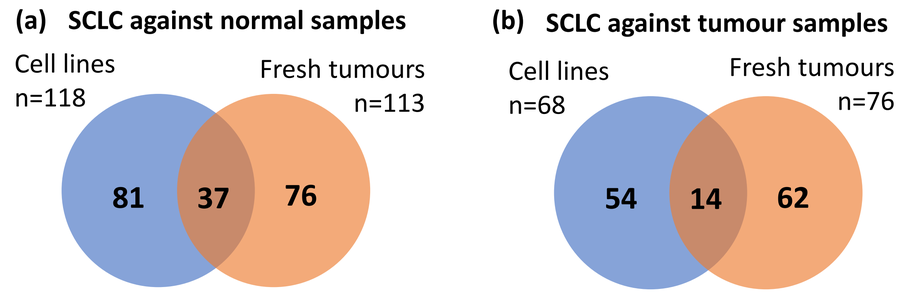
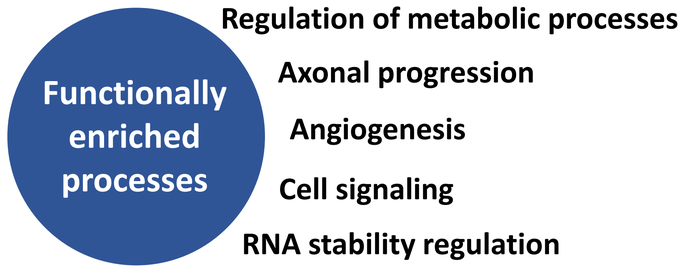


Figure 3. We considered gene lists from analyses of SCLC against tumour and normal samples, separately, to identify genes that appeared alternatively spliced in SCLC fresh tumours that were also present in cell lines. Diagrams show that both SCLC sample types share 37 genes from analyses against normal samples (a) and 14 from analyses against tumour samples (b).



CONCLUSIONS

- **A set of disctinct transcripts are expressed in SCLC compared to other samples.**
- **The identified genes may have functional relevance shedding light to the study of new molecular mechanisms for SCLC.**
- **Identification of these alterations in alternative splicing can benefit the development of prognostic and therapeutic targets of SCLC tumours.**

REFERENCES

- Sebestyén E, Zawisza M, Eyras E. Recurrent alternative splicing isoform switches in tumor samples provide novel signatures of cancer. Nucleic Acids Res. 2015 Jan 10;43(3):1345–56.
- Iwakawa R, Kohno T, Totoki Y, Shibata T, Tsuchihara K, Mimaki S, et al. Expression and clinical significance of genes frequently mutated in small cell lung cancers defined by whole exome/RNA sequencing. Carcinogenesis. 2015;36:616–21.
- Kohno T, Ichikawa H, Totoki Y, Yasuda K, Hiramoto M, Nammo T, et al. KIF5B-RET fusions in lung adenocarcinoma. Nat Med. 2012;18:375–7.
- TCGA: The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511(7511):543–50.
- TCGA: The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489:519–25.
- Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10.