

Cristóvão Lanna^{1,2*}, Luís Felipe Ribeiro Pinto², Mariana Boroni¹

¹Laboratório de Bioinformática e Biologia Computacional – LBBC/INCA. ²Programa de Carcinogênese Molecular – PCM/INCA
Instituto Nacional de Câncer – INCA - Rio de Janeiro - RJ – Brasil. *Email: cristovaolanna@gmail.com

INTRODUCTION

Colorectal cancer (CRC) is the third most prevalent carcinoma in the world, with increasing occurrences due to changes in the population's lifestyle¹. The molecular basis of CRC progression is well known and the vast amount of tumor data available has allowed the classification of CRC tumors into molecular subtypes. However, many classification systems have been developed independently and they do not always match. Recently, the Colorectal Cancer Subtyping Consortium (CRCSC), a collaborative initiative involving several independent groups, has identified four consensus molecular subtypes (CMS) based on gene expression data from more than 4,000 primary tumor samples² (Fig. 1). Since cell lines are frequently used as *in vitro* tumor models³, this study aims to classify CRC cell lines into their respective consensus molecular subtypes.

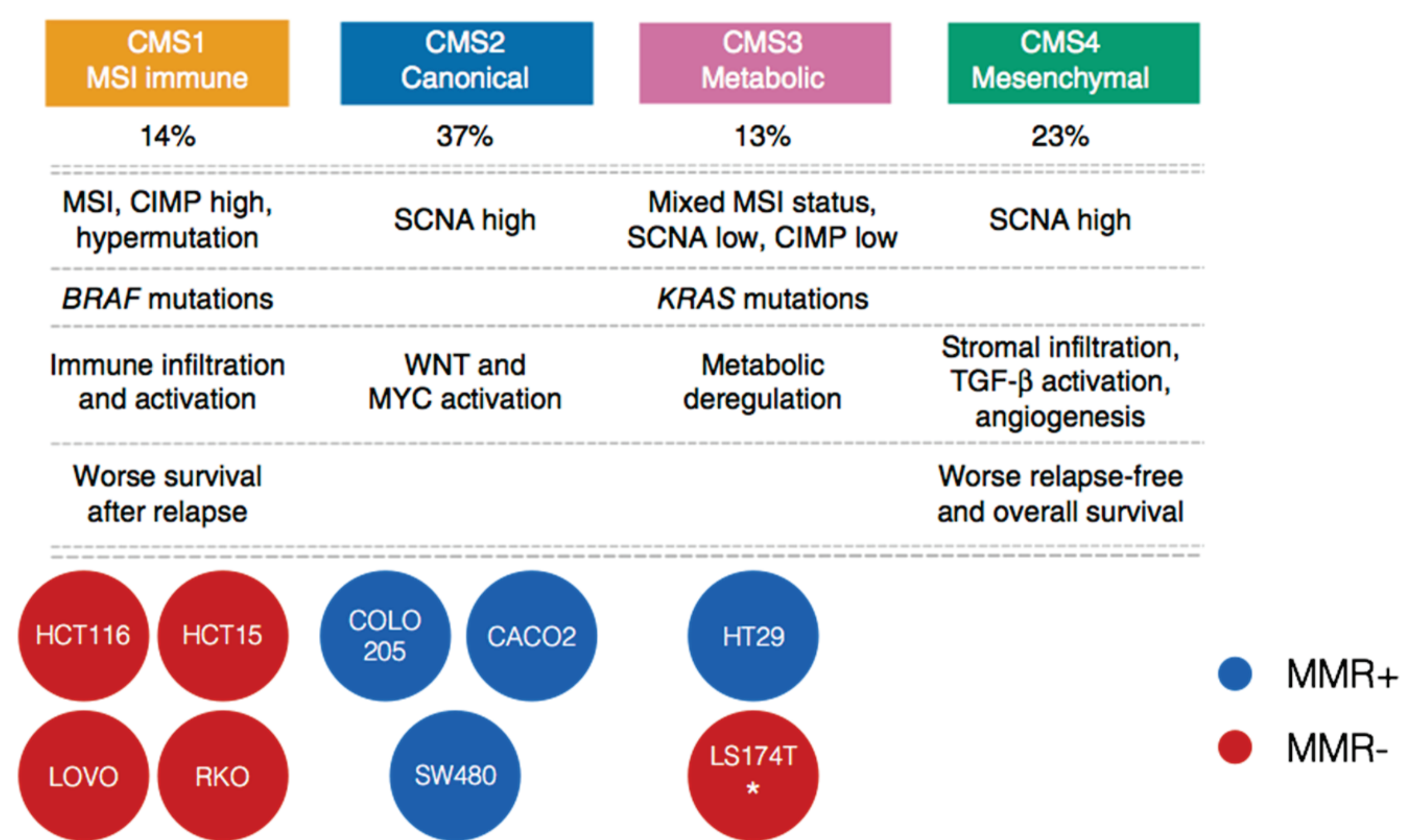
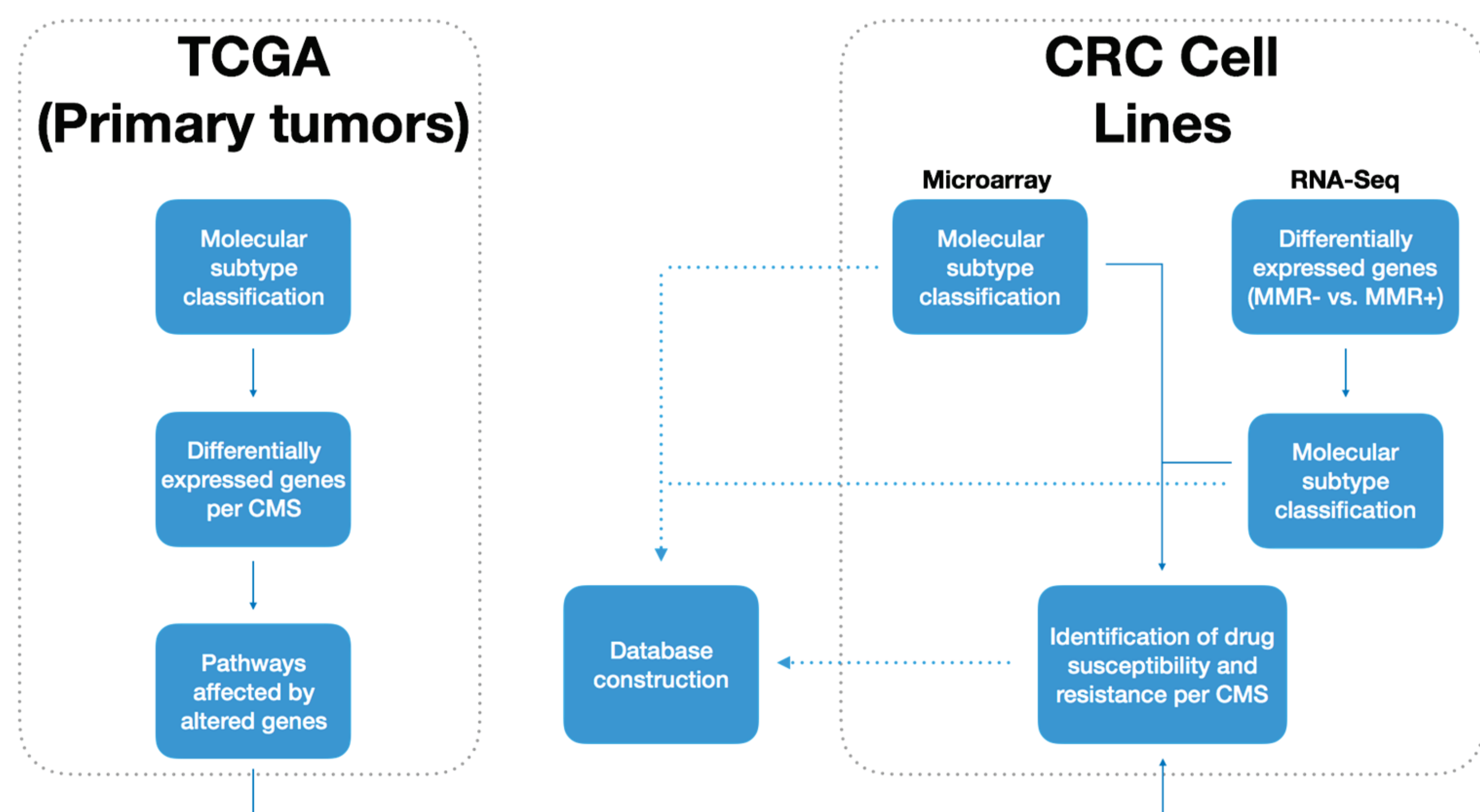


Figure 1 – Consensus molecular subtypes (CMS) characteristics and previous classification of cell lines. MMR – mismatch repair. Adapted from Guinney et al. (2015)².

METHODS



REFERENCES

- 1– BRENNER, H. et al. Colorectal cancer. *The Lancet*, v. 383, n.9927, p. 1490–1502, 2014.
- 2– GUINNEY, J. et al. The consensus molecular subtypes of colorectal cancer. *Nature Medicine*, v. 21, n. 11, p. 1350–1356, 2015.
- 3– MOURADOV, D. et al. Colorectal Cancer Cell Lines Are Representative Models of the Main Molecular Subtypes of Primary Cancer. *Cancer Research*, v. 74, n. 12, p. 3238–3247, 2014

RESULTS

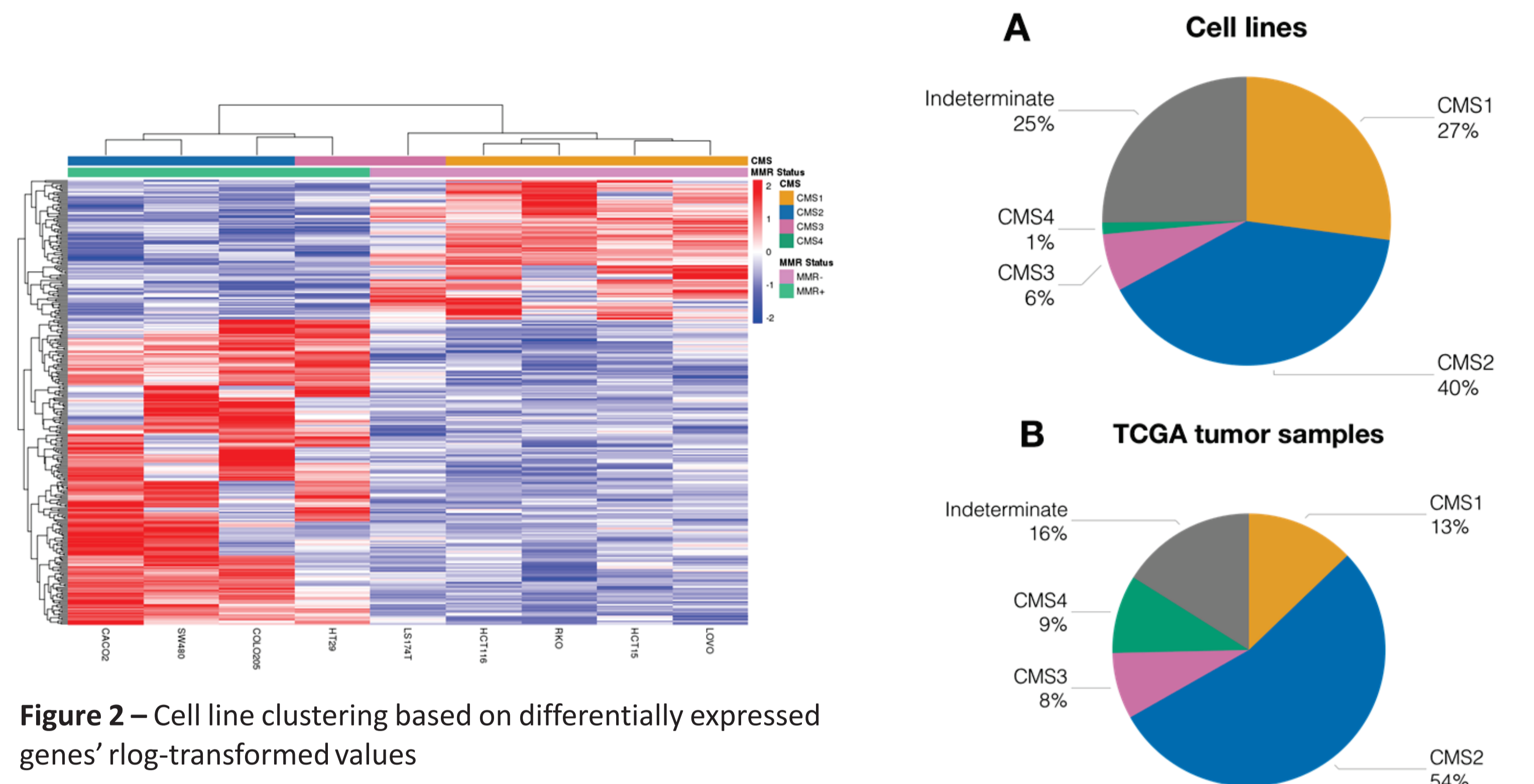


Figure 2 – Cell line clustering based on differentially expressed genes' log-transformed values

Figure 3 – CMS classification for cell lines (A) and TCGA primary tumor samples (B).

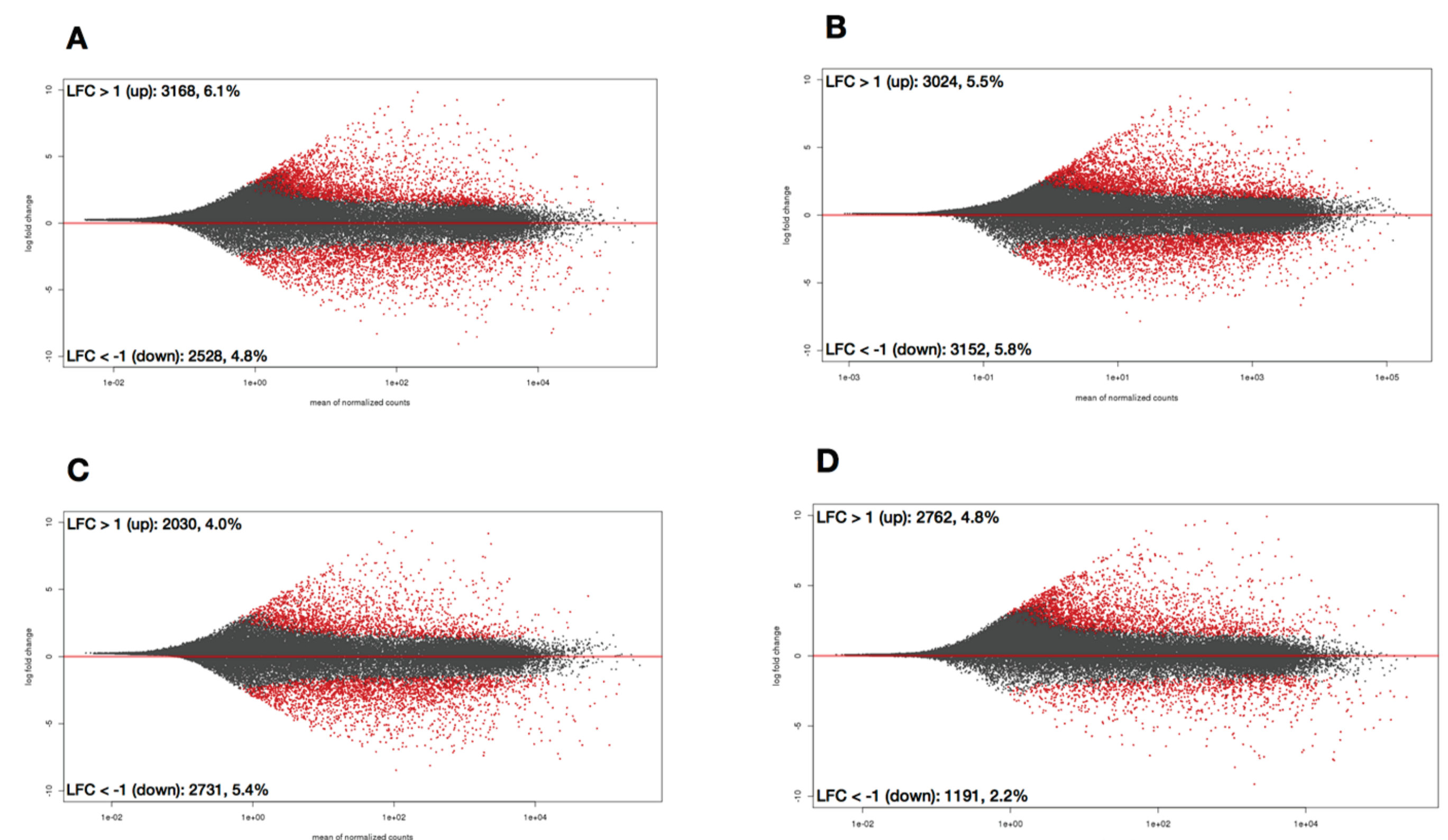


Figure 4 – Differentially expressed genes (in red) in TCGA samples when analyzed through DESeq2 (padj < 0.01). LFC – log₂ fold change. A – CMS1; B – CMS2; C – CMS3; D – CMS4.

NEXT STEPS

Representative cell lines of each subtype may be used to validate the molecular features that predict resistance/sensitivity to agents that target these aberrations. This will improve our comprehension of the mechanisms of tumorigenesis and foster the development of new therapies targeted to selectively interfere with one or more of these processes. We also plan to organize and make all data generated from these analyses available to the scientific community in the form of an online database (Fig. 3).

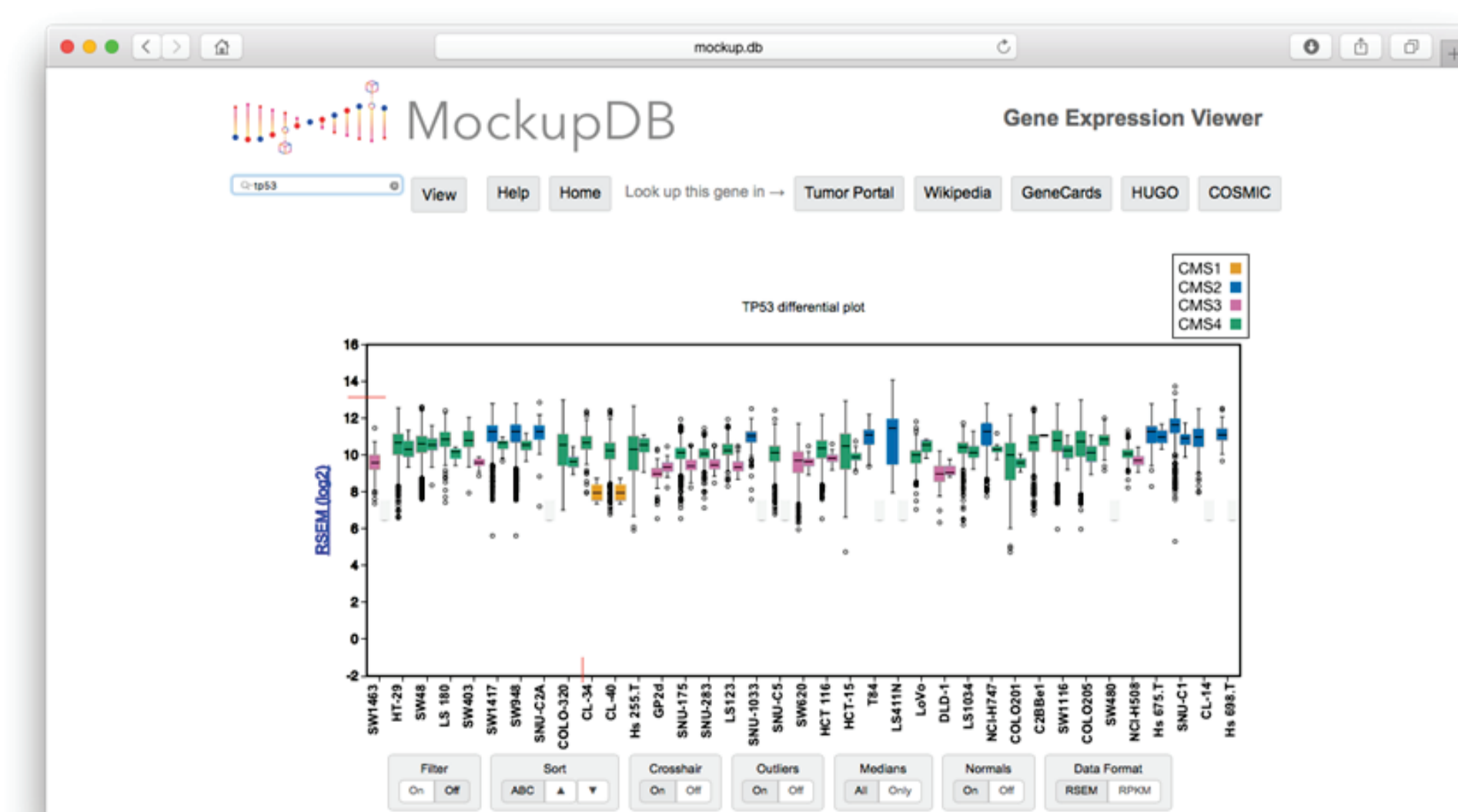


Figure 5 – Mockup of a database containing cell lines classification and expression data.