Findings

As part of my university dissertation in Bioinformatics, I carried out a differential methylation and gene expression analysis comparing Chronic Lymphocytic Leukaemia (CLL) patients with healthy controls.

The volcano plot highlights differentially methylated CpG sites between the two groups, while the heatmap visualises DMLs (Differentially Methylated Loci) linked to key genomic regions, showing clear separation between healthy and CLL samples.

A box plot illustrates differential gene expression for selected genes, and the Kaplan–Meier survival curve demonstrates the clinical relevance of one such gene by stratifying patients based on expression levels. Together, these visualisations provide insight into the epigenetic landscape of CLL.

Teeshyna Jhummun
BSc Applied Biomedical Science
Graduate

Table 2: Top 30 DMLs showing greatest loss in methylation difference between CLL and healthy samples.

| Probe ID	Gene	Position	Mean Healthy	Mean CLL	Diff	Pvalue	pFDR
CG19753067	SLC25A43	Body	0.8622	0.1584	-0.7038	3.84E-06	0.0007
CG24462702	CD40LG	1stExon	0.9153	0.2116	-0.7037	9.24E-06	0.0008
CG21302055	CD40LG	TSS200	0.8484	0.1574	-0.6910	5.62E-06	0.0007
CG20287862	CD40LG	TSS1500	0.8581	0.1878	-0.6703	4.95E-06	0.0007
CG10700483	CD40LG	Body	0.8452	0.1802	-0.6650	8.69E-06	0.0008
CG25437005	PHE	1stExon;5'UTR	0.9067	0.2517	-0.6550	1.60E-05	0.0011
CG04595053	OGT	TSS1500	0.8090	0.1835	-0.6255	3.84E-06	0.0007
CG23907260	CD40LG	TSS200	0.7491	0.1366	-0.6124	8.69E-06	0.0008
CG19244546	XKRX	Body	0.8442	0.2393	-0.6049	5.28E-06	0.0007
CG00292305	CD40LG	TSS1500	0.8125	0.2196	-0.5929	5.28E-06	0.0007
CG27492220	PAGE2B	TSS200	0.8182	0.2353	-0.5829	5.28E-06	0.0007
CG13011976	PAGE2B	TSS1500	0.7333	0.1533	-0.5800	4.09E-06	0.0007
CG27207932	CD40LG	TSS200	0.7988	0.2258	-0.5730	5.28E-06	0.0007
CG18983126	PAGE2B	TSS200	0.8656	0.2975	-0.5681	6.78E-06	0.0007
CG21943052	BGN	TSS200	0.7581	0.1958	-0.5624	7.22E-06	0.0007
CG11944101	CD40LG	1stExon	0.8298	0.2681	-0.5618	1.18E-05	0.0009
CG20205061	MSL3	Body;3'UTR	0.8565	0.3119	-0.5446	6.92E-06	0.0007
CG12274524	PIM2	Body	0.8114	0.2673	-0.5440	5.62E-06	0.0007
CG08843314	CXCR3	5'UTR; Body	0.6765	0.1341	-0.5424	5.98E-06	0.0007
CG26239772	PAGE2B	TSS1500	0.7649	0.2226	-0.5423	7.68E-06	0.0008
CG03166428	BGN	TSS200	0.7607	0.2311	-0.5296	4.65E-06	0.0007
CG07601068	BCOR	5'UTR	0.7684	0.2441	-0.5243	6.10E-06	0.0007
CG16166036	WWC3	Body	0.8611	0.3467	-0.5144	6.78E-06	0.0007
CG24548265	ZNF275	Body	0.7219	0.2085	-0.5134	5.28E-06	0.0007
CG04929865	BGN	TSS1500	0.7396	0.2332	-0.5064	1.18E-05	0.0009
CG15325759	PHEX	TSS200	0.8863	0.3841	-0.5022	0.0002	0.0055
CG17964575	TCEANC	Body	0.8091	0.3129	-0.4962	5.13E-05	0.0021
CG07919695	CD40LG	TSS1500	0.8396	0.3439	-0.4958	5.28E-06	0.0007
CG01110765	BCOR	5'UTR	0.7718	0.2818	-0.4900	2.42E-05	0.0014
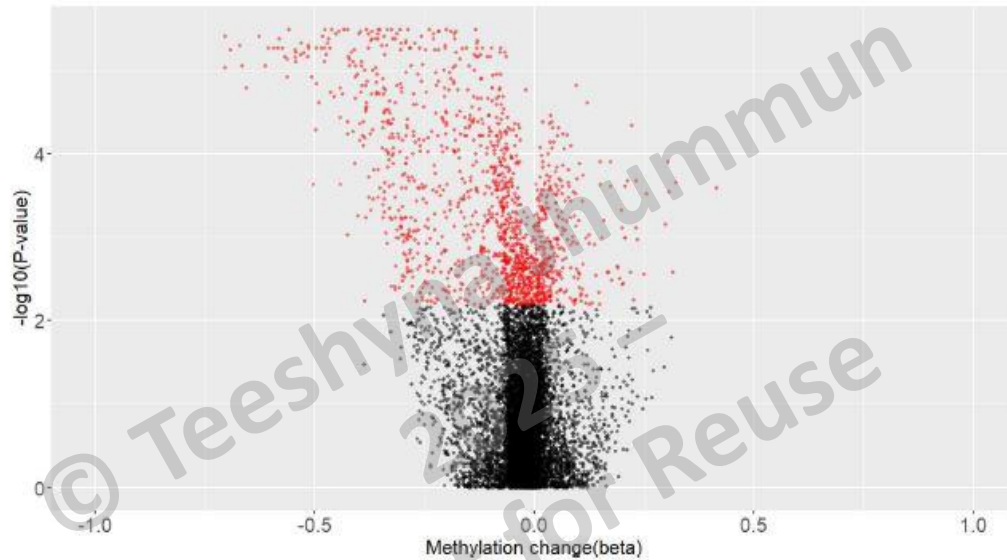
Figure 3: Volcano plot of differentially methylated CpG sites between CLL patients and normal samples. The beta values range from 0 (unmethylated) and 1 (methylated). Methylation difference is shown on the x axis. Log10 - pvalue is shown on the y-axis. CpG sites that were considered significantly differently methylated are coloured in red (pFDR <0.05) and those that are not significant are coloured in black.
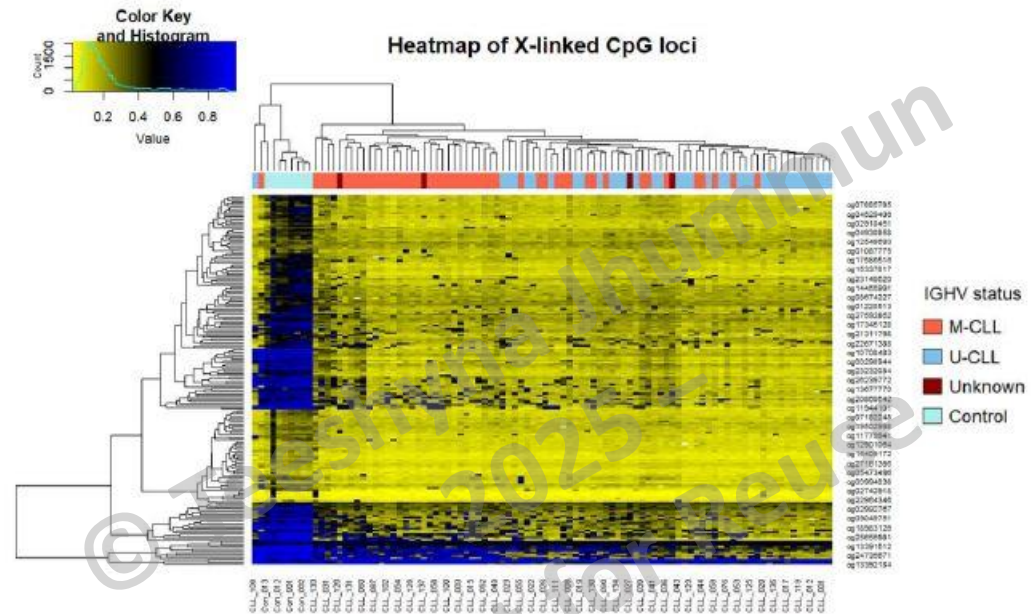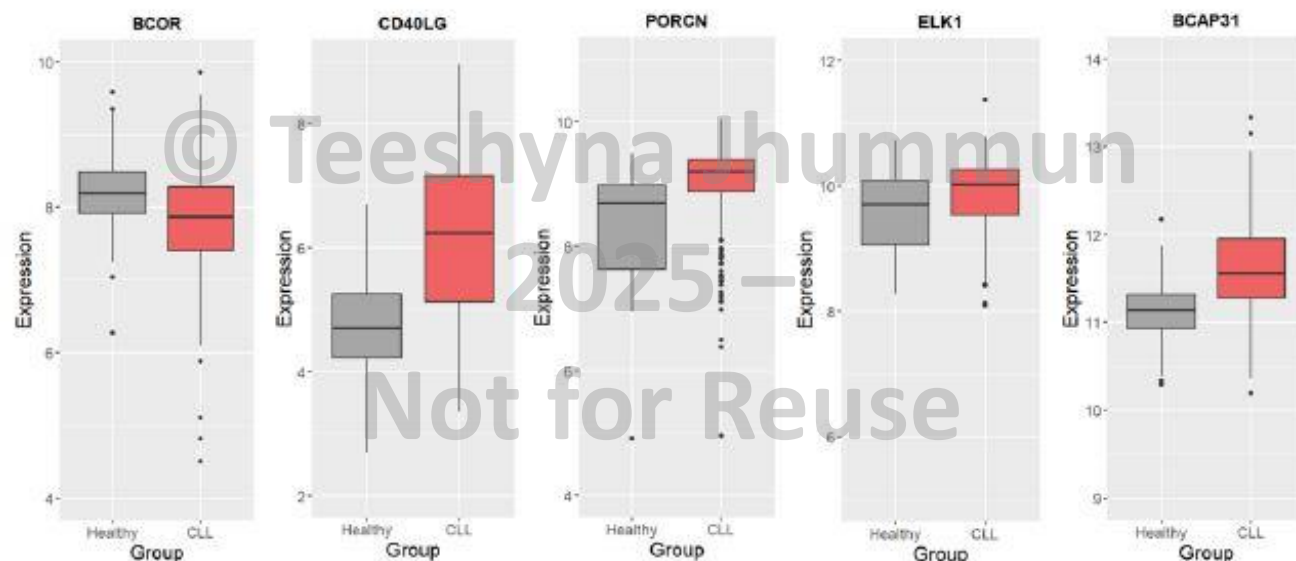

Figure 4: Heatmap of X-linked DMLs between healthy and CLL patients. Each row represents an individual CpG site, and each column represent a different patient sample. The colour scale indicates methylation level, from low (yellow) to high (blue). Legends represent hierarchical cluster analysis of 1,456 CpGs differentially methylated in U-CLL versus M-CLL.

Box-plot : showing expression of key genes against healthy and CLL patients
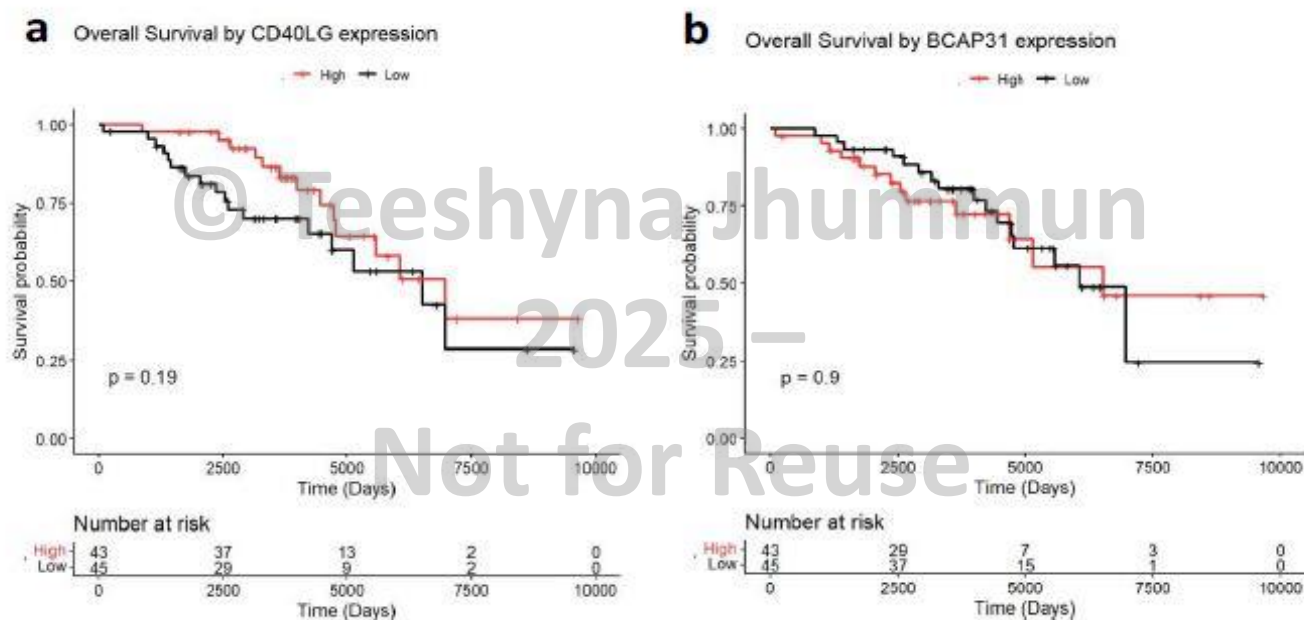
Kaplan-Meier – Survival Analysis