



Table 1. List of Cancer Patient Cards. There are four cards per cancer type, each card representing a different patient sample. In total, 51 genes are associated with the cancers listed on the 32 cards.

Cancer Type	Sample	Genes					
Lung cancer	1	NF1	SETBP1	TP53			
	2	EGFR	MLL2				
	3	CTNNB1	KRAS	NF2	TP53		
	4	KT	MEN1	MLL3	TP53		
Breast cancer	1	BRCA1	TP53				
	2	CDH1	PIK3CA				
	3	ARID1B	TP53				
	4	FGFR2	GATA3				
Colorectal cancer	1	APC	TP53				
	2	KDM6A	KRAS	PIK3CA	SMAD4		
	3	APC	ATM				
	4	BRAF	CARD11	GNAS	РІКЗСА	SMAD4	TP53
Glioma	1	ac	IDH1	PIK3C5			
	2	CBL	TP53				
	3	ALK	ATM	BRCA1	TP53		
	4	HNF1A	PTEN				
Melanoma	1	BRAF	CREBBP	EP300			
	2	FGFR2	MLL3	NRAS	PTCH1		
	3	BRAF	APC	BCOR	JAK3	MLL2	
	4	BRAF	CDKN2A	MLL3	ATM	PAX5	
Hepatic cancer	1	ARID1A	ARID2	BRCA1	RB1	TP53	
	2	CTNNB1	MED12				
	3	HNF1A	TP53				
	4	ARID2	AXIN1	JAK2	TP53		
Pancreatic cancer	1	APC	GNAS	KRAS	RNF43		
	2	KRAS	TRAF7	TP53			
	3	CDKN2A	KRAS	MPL	TP53		
	4	KRAS	SMAD4	TP53			
Leukemia	1	MYD88	SETD2				
	2	BCOR	NOTCH1				
	3	BRAF	FAM123B	KRAS			
	4	NOTCH1	PIK3CA	TP53			

Analisi 2:

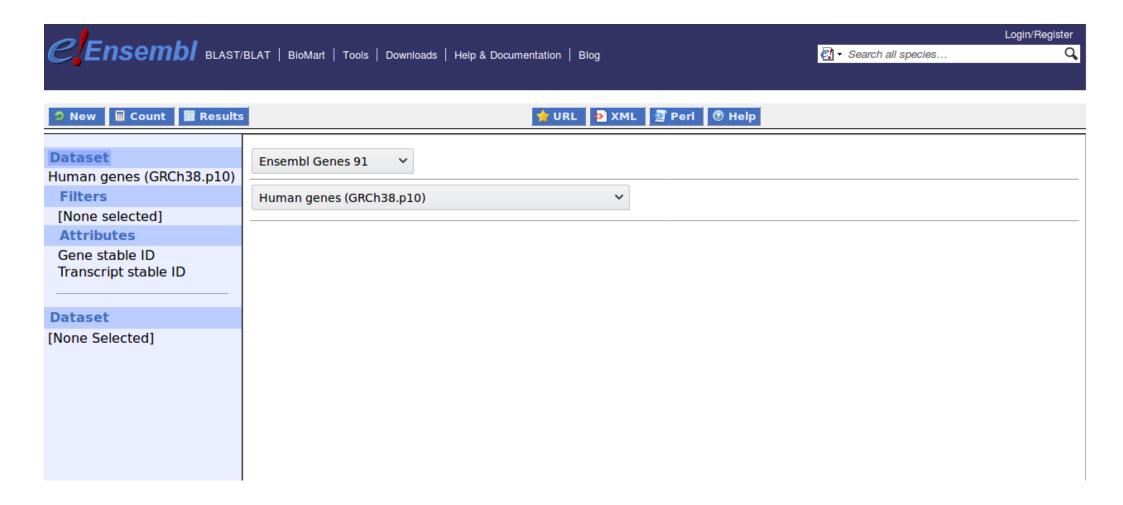
- Utilizza il portale di dati genomici *BioMart* per ottenere una tabella di coordinate genomiche relative ai geni tumorali elencati nella tabella fornita
- Il portale *BioMart* portale è accessibile al seguente sito: http://www.ensembl.org/biomart/martview/

Suggerimenti per svolgere l'analisi:

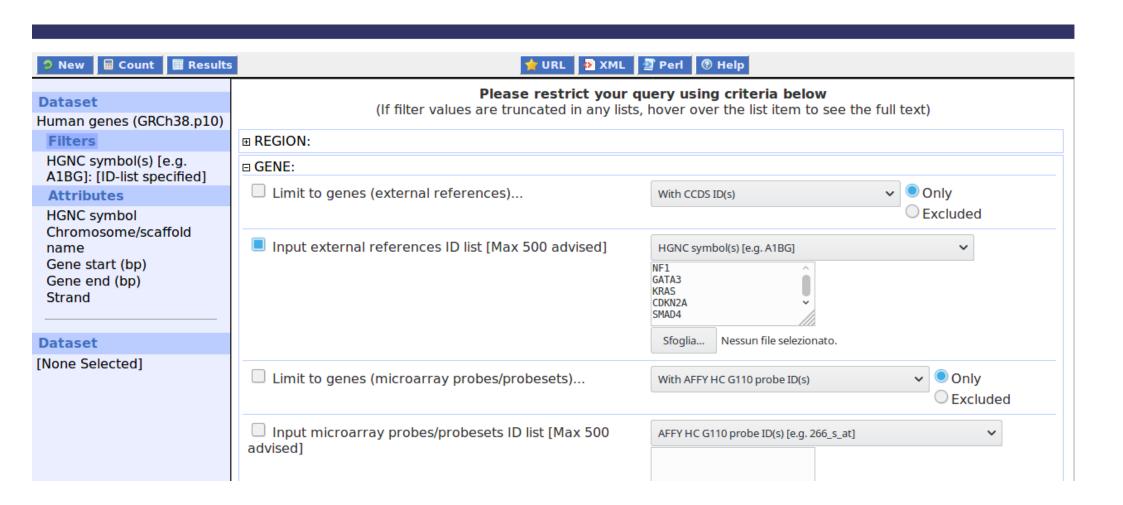
- Per coordinate genomiche di un gene si intende la sua localizzazione sul genoma: Nome del cromosoma su cui risiede il dato gene; posizione di inizio; posizione di fine.
- Segui la guida passo-passo nelle diapositive seguenti per vedere come ottenere queste informazioni per la lista di geni tumorali di nostro interesse

Guida passo-passo per ottenere le coordinate genomiche di una lista di geni con BioMart:

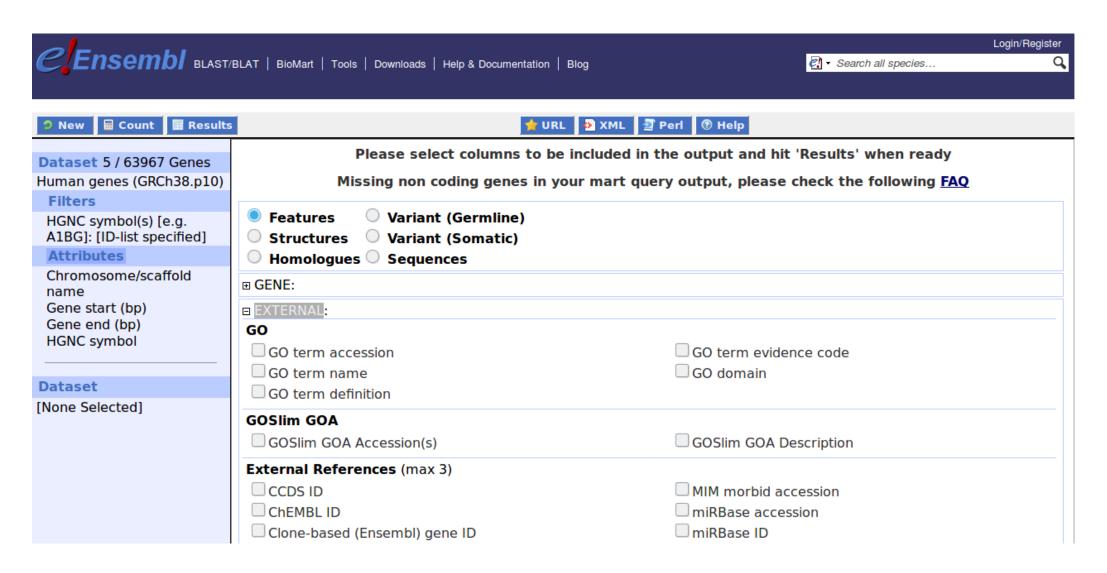
- 1. Apri una finestra di browser sul sito: http://www.ensembl.org/biomart/martview
- 2. Seleziona "Ensembl Genes 91" come database e "Human Genes" come dataset



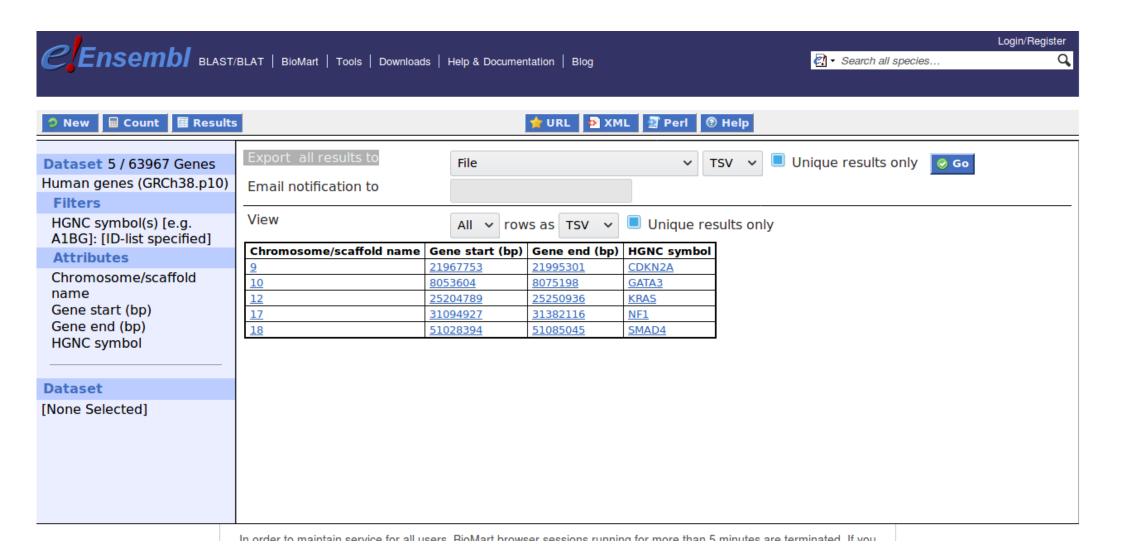
- 3. Seleziona la voce "Attributes" dal menu' di sinistra
- 4. Espandi la sezione "GENE" e seleziona "Input external references ID list" come metodo di filtro
- 5. Dal menu' a tendina corrispondente seleziona "HGNC symbols"
- 6. copia / incolla la lista di geni tumorali di interesse nell'apposito spazio (uno per riga), oppure Copiali in un file di testo (sempre uno per riga) e carica il file
- 7. Premi sul tasto counts dal menu' in alto per vedere quanti geni hai selezionato. Il numero Corrisponde a quello che ti aspettavi?



- 8. Seleziona la voce "Filters" dal menu' di sinistra
- 9. Espandi la sezione "GENE" e deseleziona eventuali voci selezionate. Poi seleziona:
- "Chromosome/scaffold name"
- "Gene start (bp)"
- "Gene end (bp)"
- 10. Espandi la sezione "EXTERNAL" e seleziona:
 - "HGNC symbol"



- 11. Seleziona la voce "Results" dal menu' in alto a sinistra
- 12. Come opzioni di output seleziona "Export all results to" file "TSV";
- 13. Spunta l'opzione "Unique results only" e poi premi il tasto "Go"
- 14. Apri un foglio di calcolo Excel e importaci il file scaricato da BioMart



Analisi 3:

- Crea un file BED di coordinate genomiche per i geni tumorali di nostro interesse e importala nel UCSC Genome Browser http://genome.ucsc.edu/facendo in modo che venga visualizzata come traccia di colore rosso, con nome e descrizione opportuni (vedi immagine qui sotto)

Suggerimenti per svolgere l'analisi:

- Se non ricordi come caricare un file BED nel UCSC Genome Browser, o come modificarne il colore / nome / descrizione, ripercorri i passaggi che lo illustrano dal tutorial #2 fatto in classe la volta scorsa
- Per quanto riguarda la creazione del file BED, utilizza il file excel appena creato nell'Analisi 2 eliminando il nome delle colonne e riarrangiandone il contenuto come mostrato qui sotto. Salva questo file come file di testo semplice separato da tabulazione.

