

## ASSIGNMENT 1

NAME: BENSALEK Rihab

DISEASE NAME: Colorectal Cancer

Gene symbol	Gene name (NCBI)	Description	Function	Protein name	PDB structure availability	PDB ID
<b>APC</b>	Adenomatous polyposis coli	Tumour suppressor; cytoplasmic protein in WNT signalling at cell junctions Localisation: 5q22.2 Exon count: 20	Tumour suppressor	APC protein Amino acids: 2843	Yes	5B6G
<b>TP53</b>	Tumour protein p53	Transcription factor; nuclear Localisation: 17p13.1 Exon count: 13	Tumour suppressor	Cellular tumour antigen p53 Amino acids: 393	Yes	6SI2
<b>KRAS</b>	KRAS proto-oncogene, GTPase	Small GTPase; cytoplasmic membrane-associated Localisation: 12p12.1 Exon count: 7	Oncogene	GTPase KRAS Amino acids: 189	Yes	8T71
<b>PIK3CA</b>	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha	Kinase in PI3K/AKT pathway Localisation: 3q26.32 Exon count: 22	Oncogene	PI3K Amino acids: 1068	Yes	8EXL
<b>SMAD4</b>	SMAD family member 4	Transcription factor; TGF- $\beta$ signalling Localisation: 18q21.2 Exon count: 14	Tumour suppressor	SMAD4 Amino acids: 552	Yes	6YIC
<b>BRAF</b>	B-raf proto-oncogene, serine/threonine kinase	Ser/Thr kinase; MAPK pathway Localisation: 7q34 Exon count: 24	Oncogene	Serine/ threonin protein kinase B-raf Amino acids: 766	Yes	8C7X

<b>FBXW7</b>	F-box and WD repeat domain containing 7	Ubiquitin ligase component Localisation: 4q31.3 Exon count: 20	Tumour suppressor-like	FBXW7 Protein Amino acids: 707	Yes	7T1Z/ 7T1Y
<b>MLH1</b>	MutL homolog 1	DNA mismatch repair protein Localisation: 3p22.2 Exon count: 24	DNA repair	DNA mismatch repair protein MLH1 Amino acids: 756	Yes	6WBA/ 6WBC
<b>MSH2</b>	MutS homolog 2	DNA mismatch repair Localisation: 2p21-p16.3 Exon count: 43	DNA repair	MSH2 Amino acids: 934	Yes	8R7E/ 8R7C
<b>MUTYH</b>	MutY DNA glycosylase	Base excision repair Localisation: 1p34.1 Exon count: 22	DNA repair	Adenine DNA glycosylase Amino acids: 546	Yes	8FAY
<b>CEACAM5</b>	Carcinoembryonic antigen-related cell adhesion molecule 5	Cell surface glycoprotein Localisation: 19q13.2 Exon count: 10	Oncogene	CEACAM5 Amino acids: 702	Yes	8BWO
<b>MMP2</b>	Matrix metalloproteinase 2	ECM remodelling enzyme Localisation: 16q12.2 Exon count: 17	ECM protease	MMP-2 (72 kDa type IV collagenase) Amino acids: 660	Yes	7XJO
<b>TIMP2</b>	Tissue inhibitor of metalloproteinases 2	ECM regulator Localisation: 17q25.3 Exon count: 5	ECM regulator	Metalloproteinase inhibitor 2 Amino acids: 220	No	No
<b>CSF2RA</b>	Colony stimulating factor 2 receptor alpha	Cytokine receptor Localisation: pseudo-autosomal region (PAR) of X, Y Exon count: 23	Immune receptor	CSF2RA Amino acids: 400	No	No
<b>CD86</b>	CD86 molecule	Co-stimulatory immune protein Localisation: 3q13.33 Exon count: 8	Immune regulator	T-lymphocyte activation antigen CD86 Amino acids: 329	Yes	8HXB
<b>GREM1</b>	Gremlin 1	BMP antagonist Localisation: 15q13.3 Exon count: 3	Secreted regulator	Gremlin-1 Amino acids: 184	Yes	8B7H

<b>CLSTN3</b>	Calsyntenin 3	Cell adhesion protein Localisation: 12p13.31 Exon count: 20	Cell adhesion	CLSTN3 protein Amino acids: 956	No	No
<b>ADPGK</b>	ADP-dependent glucokinase	Metabolic enzyme Localisation: 15q24.1 Exon count: 10	Metabolism	ADP-dependent glucokinase Amino acids: 497	No	No
<b>COL3A1</b>	Collagen type III alpha 1 chain	ECM structural protein Localisation: 2q32.2 Exon count: 51	Structural ECM	Collagen III Amino acids: 1466	Yes	7WWR/ 7WWS/ 7XAN
<b>GUCA2A</b>	Guanylate cyclase activator 2A	Intestinal peptide Localisation: 1p34.2 Exon count: 3	Secreted regulator	Guanylin Amino acids: 115	No	No