Chapter 4

A Mechanistic Analysis of Spontaneous Cancer Remission Phenomenon —Identification of Genomic Basis and Effector Biomolecules for Therapeutic Applicability: Melanoma as a case study.

CHAPTER 4: A MECHANISTIC ANALYSIS OF SPONTANEOUS CANCER REMISSION PHENOMENON —— IDENTIFICATION OF GENOMIC BASIS AND EFFECTOR BIOMOLECULES FOR THERAPEUTIC APPLICABILITY: MELANOMA AS A CASE STUDY.

4. Outline

Based on our previous two chapters' studies about the well-documented phenomenon of the permanent spontaneous tumour regression process, here we are going to investigate the host tissue-initiated causative factors behind this permanent regression process so that it can be therapeutically replicated on patients. For this, in this chapter we developed a systems biological formulation of the regression process with experimental verification and identified the relevant candidate biomolecules for therapeutic utility. We devised a cellular kinetics-based quantitative model of tumour extinction in terms of the temporal behavior of three main tumour-lysis entities: DNA blockade factor, Cytotoxic T-lymphocyte, and Interleukin-2. As a case study, we analysed the time-wise biopsy and microarrays of spontaneously regressing melanoma and fibrosarcoma tumours in mammalian/human hosts. We analysed the differentially-expressed genes (DEGs), signaling pathways, and bioinformatics framework of regression. Additionally, prospective biomolecules that could cause complete tumour regression were investigated. We identified 176 upregulated and 116 downregulated DEGs, and enrichment analysis showed that the most significant were downregulated cell-division genes: TOP2A-KIF20A-KIF23-CDK1-CCNB1. Moreover, Topoisomerase-IIA inhibition might actuate spontaneous regression, with collateral confirmation provided from survival and genomic analysis of melanoma patients. Candidate molecules such as Dexrazoxane/Mitoxantrone, with Interleukin-2 and Antitumour-lymphocytes, may potentially replicate permanent tumour regression process

of melanoma. To conclude, episodic permanent tumour regression is a unique biological reversal process of malignant progression, and signaling pathway understanding, with candidate biomolecules, may plausibly therapeutically replicate the regression process on tumours clinically.

4.1 Introduction

Combination therapy or multimodal treatment is an incisive approach used to treat different types of cancer [87]. It is a combination of at least two of these three modalities: chemotherapy, radiotherapy, and immunotherapy [88]. Radiation therapy or surgery are typically local therapies for cancer that has spread locally, while chemotherapy and immunotherapy are systemic therapies, that address the fundamental problem of cancer spread in the metastatic phase [89]. Regarding systemic therapy, chemotherapy is often used as a first modality, it is inexpensive, and generally eliminates the majority of the tumour cells. Typically, chemotherapy drugs work by impeding DNA replication capacity, so that the cells cannot replicate. Since malignant cells reproduce faster than normal cells (which are usually in stable non-dividing conditions), the drugs preferentially damage malignant cells considerably more than normal cells. Nevertheless, chemotherapy and radiotherapy also somewhat adversely affect the other normal surrounding tissue, particularly by impeding the protective leucocyte system, for example by damaging the haematopoietic cells in the marrow [90, 91]. Hence as a remedial course of action, the modality of immunotherapy or biological therapy has been introduced, which helps the immunological agent to seek out the malignant cells and eliminate them, while also avoiding action on normal cells. Immunotherapy is indeed an antitumour therapy that can enhance the immune system to counter various types of cancer [92]. Furthermore, there are a few categories of immunotherapy that use checkpoint inhibitors, cytokines, or immunomodulators [93]. In immunotherapy, both immune cells (innate and adaptive cells) impair the tumour cell population. Innate immune cells include

monocytes, dendritic cells, and natural killer cells, whereas adaptive cells include T-cells and B-cells [94].

Though the outcome of malignancy is often lethal, some malignant focus can undergo extinction naturally by the biochemical and immunological environment of the host tissue. This elimination is a well-documented process, termed spontaneous regression or remission of cancer, where there is no further recurrence of malignancy. Here, the host tissue may establish internally-produced chemotherapy-like and immunotherapy-like milieu, i.e. DNA alkylation is enabled or replication is impaired, while the leucocytes and interleukin-2 is activated to attack the malignant cells. Though the leucocytes damage the tumour cells, they do not assail normal tissue cells. According to a study of numerous clinical cases of spontaneous cancer remission by Everson and Cole, the immune process is one of the main factors in the spontaneous regression of tumours [95]. Indeed, spontaneous regression is an internally-generated "endogenous" process, while tumour remission by treatment is externally or "exogenously" produced intervention. The phenomenon of spontaneous cancer regression is a well recorded phenomenon [96]. However, the mechanism of spontaneous regression is poorly understood. Nevertheless, it would be more advantageous if the mechanism of spontaneous regression could be comprehended so that the physician could replicate the regression process on a patient's tumour clinically so that the tumour becomes eliminated. Formulating the mechanistic analysis and insights into the tumour regression process is the main aim of our endeavor here.

However, the researcher generally encounters a malignant growth in its progression phase, the reversal process of complete spontaneous regression of tumours happen subclinically in human populations at a 22% - 46% rate, as per the Wisconsin and Scandinavian Cancer Screening Registries, which have long-term monitored the populations of 2.95 million and 0.34 million people respectively [10, 11] . In fact, in post-mortem studies,

it has been noticed that about half of subjects have malignant lesions in prostate and uterine cervix, with evidence of complete containment. Moreover, many malignant neuroblastoma patients have full regression from larger sized tumours[95]. To underscore, a PubMed search shows over 14,000 titles of papers dealing with spontaneous cancer regression, encompassing virtually all types of malignancies, e.g. carcinomas, sarcomas, melanomas, lymphomas, etc ("Sponteneous regression of cancer or sponteneous remission of cancer" - Search Results - PubMed, 2022). Indeed, an incisive model has been developed to elucidate the energetics or biothermodynamic framework of spontaneous tumour regression [15, 51].

For instance, melanoma cases can undergo fully effective spontaneous regression (PubMed has 585 cases studied in detail; "Spontaneous cancer regression of melanoma" - Search Results - PubMed, 2022), and this melanoma regression occurs appreciably at 10 - 35% [34]. Indeed, analysis of 10,098 melanoma regression patients showed that these patients can have incisive clinical correlates [35] . Understanding melanoma regression is critically needed, as it is a preeminent malignant tumour whose incidence is increasing rapidly. [34]. Actually, spontaneous cancer regression covers virtually all types of malignancy, grades and classes such as sarcomas, carcinomas, lymphomas, melanomas and so on [96]. In our study, we have investigated the phenomenon of spontaneous regression of malignant melanoma tumour as a case study, so as to delineate the therapeutic application. Here, we endeavour to develop a general quantitative methodology of permanent tumour regression while protecting normal tissue, whether the extinction of malignant cells is induced endogenously (by host tissue) or exogenously (by chemotherapy and immunotherapy). An important aspect here is that we show a malignant lesion can undergo permanent regression by an optimized synchronization of (1) antitumour immunological and cytotoxic effectors based on the tumour load or tumour cell population, and (2) normal tissue protection as formalized by minimization of tissue toxicity. Furthermore, we established the signaling pathways that

enable spontaneous cancer regression with the help of microarray gene expression data. In addition, we have performed network pharmacology and pathway enrichment analysis to find out the most significant genes and thereby identify therapeutic agents that can replicate the spontaneous regression process. Collateral corroboration of the projected efficacy of the candidate therapeutic molecules are provided from human survival analysis and tumour histopathological assessment.

4.2 Materials and Methods

4.2.1 Formulation of Genomic and Pharmacological Analysis

Our system biology model provides a route for complete tumour elimination with normal host tissue being protected. Here we validate our theoretical model using experimental immunohistochemical findings of the melanoma tumour regression process.

4.2.2 Microarray investigation of preclinical study

We obtained the microarray information of the process of spontaneous tumour regression of melanoma from ArrayExpress database (https://www.ebi.ac.uk/arrayexpress): Experiment (E-MEXP-1152 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MEXP-1152/) [37]. In this study, Libechov minipigs (MeLiM) that were bearing malignant cutaneous melanoblastoma tumour [37], was used. Here, in a majority of the pigs, the tumour progresses and becomes fatal, but in a minority of the animals, the tumour grows up to a certain time and then spontaneously regresses and heals, and these minority animals stay healthy. In this investigation, time-dependent gene expression profiling of 6 such minority pigs, whose tumour initially grew and then regressed permanently, was assessed at five different time points, $t_0 - t_4$ (at 3-weekly intervals) during the full course of regression (3 months). We accessed the raw data information (E-MEXP-1152.raw.1.zip) to validate our theoretical mathematical model of spontaneous cancer regression using this experiment.

4.2.3 Identification of differentially expressed genes (DEGs) of tumour regression

Subsequently, we utilized R-platform and its Bioconductor packages statistical facility **(http://**www.bioconductor.org/)[82], to analyze the time-dependent gene profiling of the microarray information of the aforesaid spontaneously regressing melanoma tumour at the 5 different time points. We used the GCRAM algorithm for background correction and normalization. The differentially expressed genes (DEGs) were filtered by applying two cut-off criteria: $-2>FC>+2$ and $p\text{-}value<0.05$ after the unpaired *t*-test.

4.2.4 Biological Signaling Pathway analysis

We analyzed the aforesaid filtered DEGs, namely 70 genes at time point t_1, 322 genes at time point t_2, 1147 genes at time point t_3, and 1349 genes at t_4, using the Ingenuity Pathway Analysis platform (IPA). Of these, 61, 292, 1049, and 1227 genes for time points t_1, t_2, t_3, and t_4, were mapped by IPA. We undertook IPA core analysis on the identified genes for each time point and performed a comparison analysis to find out the signaling pathways and biological functional pathway analysis over the four time-points (t_1-t_4) for the melanoma regression process.

4.2.5 Gene Ontology and Pathway Enrichment Analysis

Subsequently, we performed gene ontology (GO) analysis of identified differentially expressed genes (DEGs) by using ClueGo, DAVID and FunRich platforms. We also examined pathway enrichment analysis with significant cellular component (CC), biological process (BP), and molecular function(MF), utilizing the ClueGo platform from Cytoscape [97]. The functionally related GO term analysis was adjusted with a kappa score greater than 0.4 and a *p*-value < 0.05 [98].

4.2.6 Protein-Protein Interaction Network Construction and Hub-Gene Identification

We then built the protein-protein interaction (PPI) networks by string database on the Cytoscape platform [99]. The Molecular Complex Detection (MCODE), a technique from Cytoscape was used to screen modules of the PPI network with degree cut-off = 2, node score cut-off = 0.2, *k*-score = 2, and maximum depth = 100. The top 10 hub genes were identified by the CytoNCA method in the Cytoscape scheme [100].

4.2.7 Identification of Candidate Drugs from Hub Gene Interaction

Thence, we utilized the CyTargetLinker technique from Cytoscape[101], to identify the candidate pharmacological molecules that can target the dysregulated genes that we identified from the analysis of our spontaneous melanoma tumour regression process. We constructed *Homo sapiens* drug-target interactions link set from the pharmacological library DrugBank [102] and from the DGIdb system (https://www.dgidb.org/search_interactions).

4.2.8 Molecular Processes Study and Drug-Ligand Interaction

We conducted a molecular interaction study using our proteins of interest (e.g., protein TOP2A (PDB ID- 4FM9; Chain A) with different ligands. For this we used the method of [103]. We used the autodock 1.5.7 tool for investigating the molecular inhibiting interaction using the Lamarckian genetic algorithm. Before docking, all the heteroatoms and water molecules were removed from the protein crystal structure. Then 2D SDF files of different ligands were accessed from the PubChem system, and Prix platform did further energy minimization. Ligand and macromolecule preparation were conducted by adding polar hydrogen bond atoms, Kollman charges, solvation parameters, and grid box formation. The grid point spacing was 0.500 Angstroms, and the coordinates of the central Grid point were 44.552, 35.073, and 9.781 Å with grid box size 39, 50, and 29Å in the *x, y,* and *z*

directions. Finally, both protein and ligand structures were saved as. pdbqt format for docking analysis.

4.2.9 Corroboration on Human Subjects: Normal Controls and Melanoma Tumour Patients

We then used the Human Protein Atlas (THPA), an open-access bank and platform that maps all the human proteins in cells, tissues, and organs by integrating various omics technologies, which helps us to probe the human proteome[104]. We used this THPA platform to investigate the overall survival rate of melanoma patients having high-versus-low hub gene expression with a significant *p*-value. Next, we used the Gene Expression Profiling Interactive Analysis (GEPIA; http://gepia.cancer-pku.cn/index.html), an interactive analysis and visualization tool, for further verification of the functionalities of hub genes and for finding the difference in gene expression level in the melanoma tissue (cutaneous skin melanoma, SKCM) vis-à-vis the normal skin tissue[105]. The resultant data were then mapped in the TCGA database and GTEx database using the GEPIA facility for Box Plots.

4.3 Results

4.3.1 Formulation of Genomic and Pharmacological Analysis

Here we corroborated the experimental findings from the preclinical and clinical study of tumour regression with our theoretical model and explored the signaling pathways related to spontaneous tumour regression. We also provided the network pharmacology analysis corresponding to spontaneous melanoma regression.

4.3.2 Microarray Investigation and Analysis

Raw data file (E-MEXP-1152.raw.1.zip) of melanoma microarray study of spontaneous cancer regression in pigs was obtained from the ArrayExpress database with experiment no. E-MEXP-1152 (E-MEXP-1152

(https://www.ebi.ac.uk/arrayexpress/experiments/E-MEXP-1152/)[37]. This zipped archive file contains 25 cel files having $n = 6$ tumours at t_0 and t_1 time points, $n = 5$ tumours at t_2 and t_3, and $n = 3$ tumours at t_4 (where n = no. of excised tumours). The tumour biopsies were taken at these five different time points (in days after birth: $t_0 = 8$ days, $t_1 = 28$ days (4 weeks); $t_2 = 49$ days (7 weeks), $t_3 = 70$ days (10 weeks), and $t_4 = 91$ days (13 weeks), i.e., each time point advancing by 3 weeks[37]. We analyzed this microarray data on the R platform; for normalization and statistical analysis Bioconductor tools were used. An unpaired t-test was performed to find out the $\log_2$ - transform and the p-value of the test was adjusted using the FDR algorithm. Differentially expressed genes were identified for all 4-time points (t_1-t_4) with respect to t_0, and the following two criteria were used namely the fold change (FC) should be greater than 2, and the p-value should be less than 0.05. Thereby, we extracted 70, 322, 1147, and 1349 differentially expressed genes (DEGs) from the microarray data expressions at these four different time points: t_1, t_2, t_3, and t_4 (the expressions were calibrated with respect to the expression at the first time observation done at $t_0 = 8$ days after birth). Our findings are given below, where R denotes the ratio of the number of downregulated genes to upregulated genes (expressed as a percentage):

(i) *At 4 weeks:*	54 genes upregulated,	16 genes downregulated,	$R = 29.6\%$
(ii) *At 7 weeks:*	189 gene upregulated,	133 genes downregulated,	$R = 70.4\%$
(iii) *At 10 weeks:*	537 genes upregulated,	610 genes downregulated,	$R = 113.6\%$
(iv) *At 13 weeks:*	569 genes upregulated,	780 genes downregulated.	$R = 137.1\%$

Note that the R value crosses the equipoised point of 100% in the third month, indicating that the number of downregulated genes > upregulated genes from that time onwards. We observe that as the tumour regression process proceeds, initially the number of upregulated

genes exceeds that of downregulated genes, but as the process further advances, the reverse situation occurs, i.e. number of downregulated genes surpasses that of upregulated genes.

4.3.3 IPA Signaling Pathway Analysis

After finding the aforesaid DEGs of the regressing melanoma tumour, we performed the Ingenuity Pathway Analysis (IPA core analysis) for each time point. IPA core analysis provided a graphical summary showing a ready perspective of the significant biological themes (Figure 4.1). This feature enables us to highlight some of the powerful entities identified in the core analysis and shows how they are related to each other by creating a more comprehensive and understandable framework of what is being processed in the analysis. In Figure 4.1, we show the graphical summary including the pertinent entities as canonical pathways, upstream regulators, diseases, and biological functions. Each entity selected for the network passes Fisher's exact test (i.e., p-value<0.05). The biological process and regulators also pass an absolute z-score cutoff of 2 or greater. Nodes are colored by their activity predictor in the analysis, where orange nodes are predicted to be activated, and blue nodes are predicted to be inhibited (Figure 4.1).

Now, we compare the aforesaid biological graphical network between time points t_1 and t_2, as the tumour regression process is underway across 3 weeks. At the t_1 time point (Figure 4.1, first panel), the tumour has barely started to regress, it has been in the tumour progression stage just before that time. Hence in the first panel, we can observe that most activated biological functions are involved in carcinogenesis, such as adhesion of cancer cells, damage/degeneration of connective tissue (for cancer cells to invade), migration of skin cancer cell lines (due to the melanoma skin tumour activity), etc. In contrast, at time point t_2 (Figure 4.1, second panel), the regression process is well proceeding; here we observe that the most highly activated entities are involved in cell cycle G2/M DNA damage checkpoint regulation (alteration of cell replication), immune response of phagocytes,

leukocyte Extravasation signaling, differentiation of B lymphocytes, and increased response of leukocytes. From the second panel, we can observe the functionalities of all three antitumour components for tumour regression:

(i) Retardation of cell replication as by DNA interference,

(ii) Increased activation of interleukin-2,

(iii) Enhanced activation of antitumour lymphocytes.

Regarding item (iii) above, we further noted the increased signaling activity of the extravasation of circulating leucocytes in the second panel. Indeed, one knows that during the immune response, the extravasation of circulating leucocyte through the vascular wall into the tumourous tissue, correlates with extravasation and actuation of lymphocytes[83]. Furthermore, we can infer that the first phase of melanoma regression is shown by the situation time point t_2. Thus, the time point t_2 is the reversal in the path of melanoma progression: at time t_1 the malignant activity in the tumour was definitive, but at t_2 the melanoma progression has halted, and the inverse process, tumour regression, is now occurring.

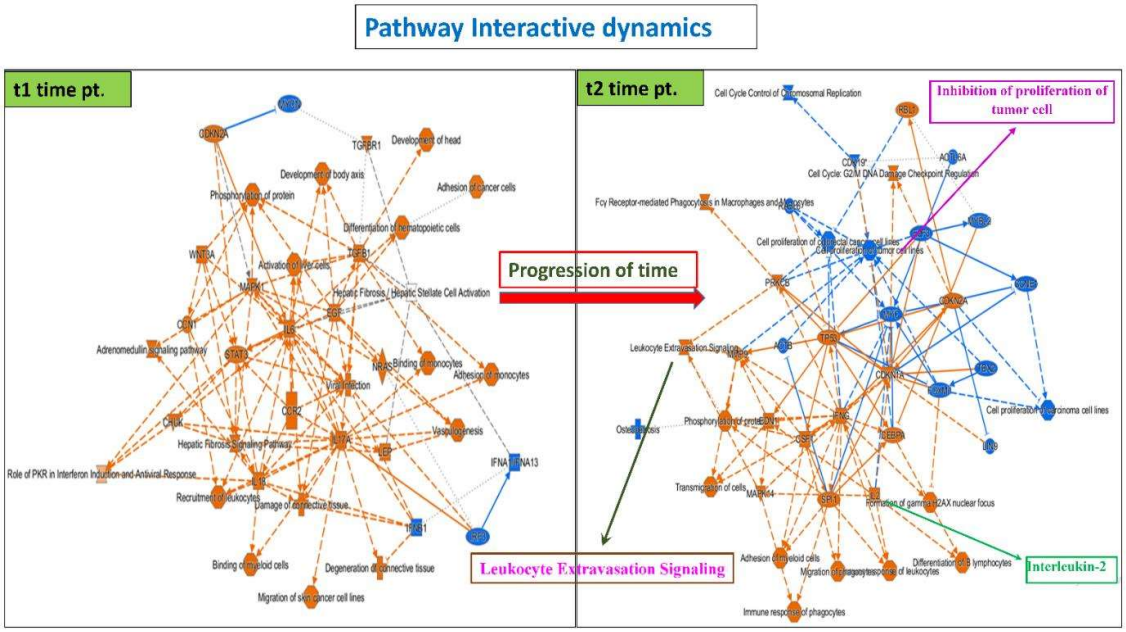


Figure 4.1 Graphical summary of comparison between two different time points: the first time point is when the tumour is progressing (left panel), the second time point is when it is regressing (right panel). The networks show a ready overview of the major biological entities such as canonical pathways, upstream regulators, diseases, and biological functions; here orange color nodes show the predicated activated entities, while blue color nodes show the predicated inhibited entities (the intensity of color is based on magnitude of z-score value in the analysis). In the regression phase (right panel), one can note that the inhibited region (blue) is in the upper right portion) and is mainly concerned with the first-order kinetics based decline of malignant cell population, namely by DNA interference and inhibition of tumour cells replication. The lower red colored region (right panel) show activation of the other

Thereafter, we have used IPA's interaction networks to provide a mechanistic understanding of how a group of molecules of the data set might work together, whether by physical interaction or by acting on each other via a less direct mechanism. These highly interconnected molecular networks are annotated with their likely biological function, which furnishes us with the framework of how they contribute to the biological process of tumour regression. We then ranked the networks in the order of their significance with respect to the probability of finding that the set of molecules in the given network occurs by random chance. We recapitulate that the regression phenomenon of this experiment has

been activated from t_2 time point. Therefore, we have considered the most significant networks as cancer, cell cycle, and hematological condition/hematological disease; note that we have taken the hematological network as oncological phenomena involves hematological processes, as white blood cell activation, hypoxia, local bleeds, etc. Using the statistical significance scoring, we found the involvement of 30 focus genes (Figure 4.2). We found that most of the downregulated genes at time point t_2 belong to cyclin-dependent kinases, cyclins and spindle formation/separation proteins, the names of the genes are given below. Downregulation of these genes enables retardation of cell division, so tumour progression is impeded, and tumour regression can occur.

Downregulated genes at the initial phase of spontaneous tumour regression:

Genes: CDC20, CDC27, CDC6, CCNB1, BUB1, BUB1B.

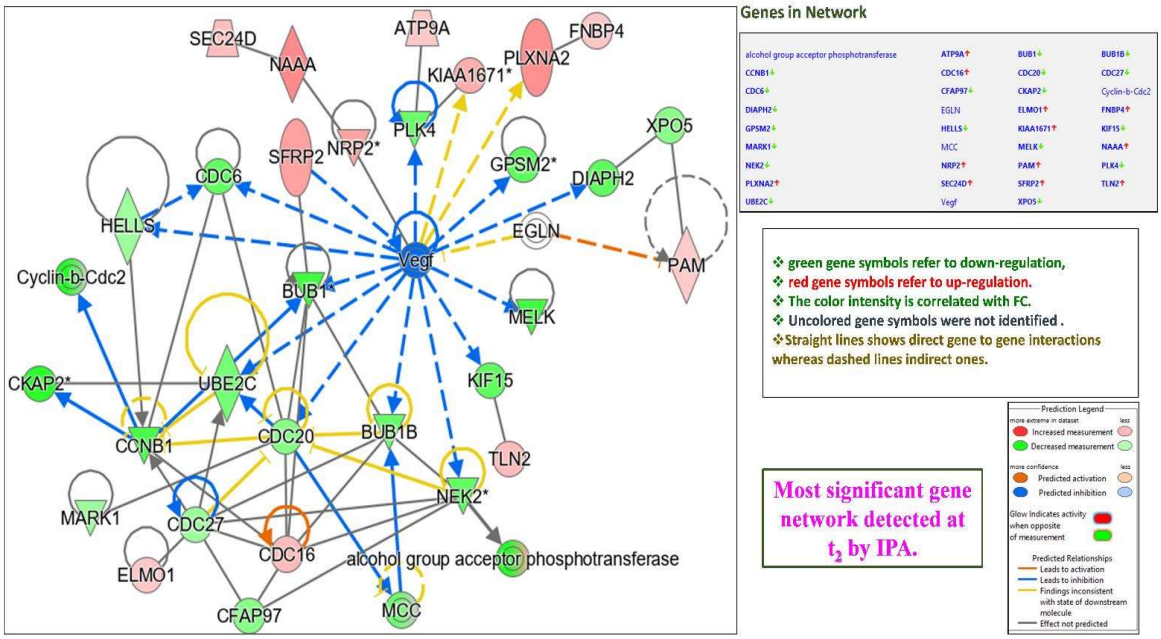


Figure 4.2 The most significant gene network detected by IPA Pathway analysis at time point t_2 (tumour regression process has started and is underway): A total of 30 focus genes were identified and mapped to top IPA functional classes, such as cancer, cell cycle, and hematological disease (significance score = 45). The green nodes are the downregulated biological diseases and functions, while the red nodes are the upregulated molecules; and the intensity of the focus nodes is based on the data measurement value in our dataset. The geometric symbols of triangle, oval circle and rhomboid respectively denote kinase, transmembrane receptor, chemical moiety and enzyme respectively.

4.3.4 Identification of Genes (DEGs) enabling first-order kinetics of tumour regression

To identify the genes responsible for the exponential decreasing curve of the tumour population, we have taken the common DEGs genes at time points t_2, t_3, and t_4, as the first sign of spontaneous regression is shown at time point t_2 (Figure 4.1). Therefore, we have collected the genes expressed at all these 3-time points (regression phase) but have a negative or very low expression at time t_1 when tumour progression occurred. Using the Venn diagram, a total of 292 overlapping DEGs were identified; out of these, 176 are upregulated genes (Figure 4.3a), and 116 are downregulated genes (Figure 4.3b). These

upregulated and downregulated genes are enumerated respectively in Table 4.1 and Table 4.2 with their fold-change values. The significance of these genes, both up and down-regulated, are elucidated later.

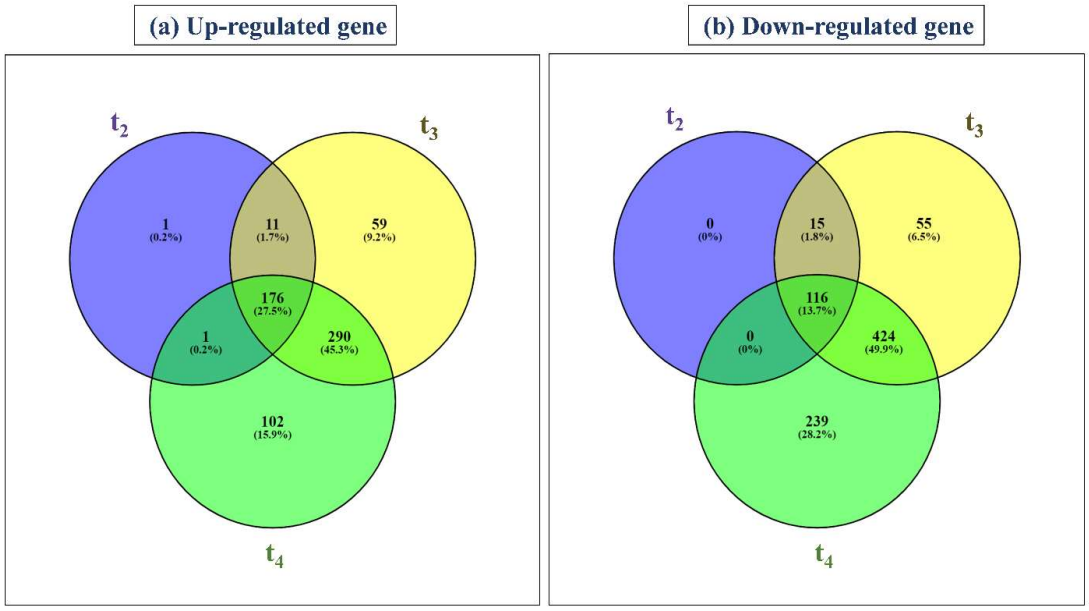


Figure 4.3 Venn diagram: Identification of differentially expressed genes (DEGs) responsible for first-order decreasing curve of tumour extinction, while the tumour undergoes regression across the duration of regression, i.e. along the three time points t_2, t_3, t_4. The common intersection zones indicated the genes that were effective at all the time points, throughout the regression process. Thereby, a total of **(a)** 176 upregulated and **(b)** 116 downregulated genes are obtained for further analysis.

Table 4.1 List of Upregulated Genes; These are expressed at all the three time points across the duration of the spontaneous tumour regression process (t_2, t_3, & t_4), each time point being at three-weekly intervals.

Gene name	log FC value (t_1)	log FC value (t_2)	log FC value (t_3)	log FC value (t_4)
IGHM	3.645186318	5.674816848	6.912221278	8.894142866
Q96LP2	5.096258616	4.442371928	6.063618336	4.833352905
IGLC1	2.073153915	4.053037664	4.888596515	6.687961122
ATP6V0D2	1.600578696	4.046134025	6.217235972	4.513096843
CKB	1.322759343	3.574536655	5.198463973	1.954223725
CALCR	0.560492854	3.557711183	7.120822129	2.655220061
SLC37A2	0.148433568	1.037410689	3.014766223	1.641033719
FN1	2.02799876	2.895633099	4.478628041	3.271786693
AMBN	0.925444244	2.67210484	4.034182413	5.876099107
ACP5	0.635513047	2.565404226	3.448297012	2.604655881
MYO1D	1.106633504	2.51308356	4.606384876	2.265977716
NP_940851	0.622407405	1.916956715	3.858546452	1.979009617
ITGB2	1.220964868	2.464950073	3.898100291	4.447100628
TVB1_HUMAN	1.069293704	2.395249702	3.226895716	6.522472048
EBI2	0.199826589	2.436796782	4.559328537	4.227503425
SPI1	1.178457482	2.42691874	3.975365659	3.949202446
SNX10	0.604214762	2.384256676	3.984937424	3.934485614
PSD4	0.923825254	2.33951757	3.84103499	3.87304396
REM1	1.985077361	2.297087573	3.959112204	2.444206418
ASAHL	0.629761132	2.268804525	4.312923649	3.571738856
EPAS1	1.75478446	2.242479913	3.423898705	2.751936475
FCER1G	1.104812099	2.224435259	3.617828903	3.974769549
CSF1R	1.120665607	2.20903005	3.682998727	2.949924417
PLXNA2	0.706829847	2.166753527	3.808959931	1.957080036
CCL5	0.95432399	2.166594618	3.02566777	5.575662535
CCR1	1.276697384	2.154739502	3.683050276	3.518325657
TYROBP	1.066511954	2.150859549	3.53415228	3.54948725
FMOD	1.864397501	2.143191473	3.782397274	2.645607072
DNM1	0.870186479	2.127124773	3.934906822	3.242519324
COL27A1	1.01765134	2.083164085	3.706622355	3.854343537
TCA_HUMAN	0.495087407	2.082731971	3.285685187	5.618101535
CYBB	0.265420269	2.081162505	3.163883026	4.146227952
ICAM2	0.924284277	2.039845207	3.367814486	3.727834151
Q8N4T3	-0.403550597	0.42158902	1.079602934	2.07701023
CUTL2	0.80103405	1.984192083	3.928426416	1.60183375
KIF23	0.827249425	1.98418174	3.664080664	3.065167151
TP53I11	1.860299174	1.983858373	3.973396981	2.007927284
CLU	1.302345567	1.969244026	3.641999023	2.712297393
IL18BP	1.131722025	1.921919296	2.815166316	2.255435323
MBP	-0.282328619	0.13797539	1.470800545	1.932379746

CBLB	1.04402778	1.906346585	3.490856484	2.970926701
S100A4	0.927305294	1.900636919	2.872591319	2.364316176
CTSZ	0.508066771	1.898388894	2.545786026	2.681814838
TCIRG1	0.668132981	1.892059562	3.234522977	2.329080921
KCNAB2	1.116473679	1.875784957	3.010688616	2.946611671
NP_942123	1.46197865	1.871558598	4.35450591	2.967055535
DPYSL3	0.976232094	1.862293842	3.059270956	2.548008247
SLC7A3	0.458227306	1.856933363	2.759540942	3.997383132
SIAT1	1.300152153	1.84461889	2.859339061	4.005865244
PAPOLA	0.518641812	1.834000793	2.64667217	1.493771538
SIGLEC5	0.156382684	1.807829476	4.546221744	1.87620082
NRP2	0.422436865	1.79691417	3.105828971	2.383216531
CRYAB	-0.481819998	1.790591444	3.479414016	3.394787735
RUNX1	1.510285934	1.787155626	3.419941559	2.791587884
MAPK9	1.491011301	1.773160817	3.239719598	3.611948454
BTK	0.672841279	1.76048106	3.398132444	3.812624961
CD3E	0.280740192	1.74730373	1.721663138	4.699617117
CLECSF5	1.291117218	1.746778609	4.714240907	5.111437442
DOK2	0.592284915	1.73921237	3.026577549	3.775726924
RH25_HUMAN	0.769257689	1.735465391	3.155261212	3.725736371
HEM1	0.025192843	1.110483577	3.020647371	3.946788299
NP_443170	0.859144369	1.716886296	3.60006391	1.700125007
ZNF205	0.620105475	1.711317459	2.693408719	2.961863162
IMP1_HUMAN	0.710506466	1.70001846	2.651311008	3.167535929
BIN2	0.632244194	1.697221451	3.592732892	3.127994563
VAV1	1.166124578	1.692861783	2.878627222	3.578314195
BLNK	0.648350119	1.686481903	3.421082413	3.946622562
GGT1	-0.051297	1.673304164	4.248247574	2.665194937
SLC15A3	0.413221813	1.666020167	3.058732194	3.587348835
CNOT2	0.775741411	1.652635694	3.026722902	3.910971278
PRKCB1	0.917693369	1.645674369	2.866621257	3.521856595
CPM	0.806031568	1.644519633	3.495407683	2.848497912
LCP1	0.663371266	1.609829315	3.137552862	3.198107717
SLC31A2	-0.113239669	0.257457652	1.484652689	3.02189655
CCR5	0.49541379	1.59067943	3.122720294	4.385600017
Q9BY89	0.45608131	1.302861249	3.505291462	2.901872211
C6orf103	0.221336169	1.579536701	2.657289703	2.644154274
Q8N200	1.052163084	1.576587043	3.102525929	2.762472414
FNDC1	1.616754759	1.561967888	4.154415178	2.700918397
PRELP	0.532056304	1.55078928	2.37677867	2.41412945
CORO1A	1.091158284	1.540712777	2.514373225	3.462906199
MAP1A	0.517919567	1.283912968	2.734221293	1.905218057
LAM5_HUMAN	0.286721662	1.37376004	2.555352681	2.417305013
ANPEP	0.285552037	1.535244876	3.821490634	2.719078164

SATB1	0.70065846	1.532222808	3.274660006	3.346215565
PTPRCAP	0.751595449	1.526106381	2.730323788	3.694183556
CNGB1	0.211417551	1.525533219	3.027187952	2.940759033
CXCR4	0.988550365	1.522656782	3.316792177	3.367955737
RASSF2	1.131108088	1.512810669	2.597018952	2.705891499
NP_054778	0.42950347	1.509832693	2.412068784	3.309159332
NP_061900	0.09001486	0.47151655	1.677216426	2.486484304
Q6NVV9	1.476384391	1.456814528	2.871872818	1.998218918
SEMA4D	0.058381741	1.44001061	3.160436406	2.347504392
Q8NHP8	0.68665869	1.435640289	3.255652518	3.318380563
ECE2	1.016765034	1.434572366	2.671541524	2.025870102
TNS	0.464326531	1.415420586	2.71219254	2.671581735
DISC1	0.400493279	1.414660967	3.179515503	3.167446229
Q8N2F5	0.692648761	1.412994476	3.162832858	2.868537883
SCTR	0.337190192	1.406395647	2.704272141	4.570756426
O75894	0.662831728	1.403824606	2.482149546	2.719200834
EMILIN1	0.976143978	1.395132572	2.456618724	1.340904165
GLIPR1	0.907591423	1.380544229	2.741004369	3.580088038
EFHD2	0.635297753	1.380369076	2.31335568	2.216821705
CENTG3	0.597908479	1.375408226	2.002300589	1.235852335
KRT8	0.865446092	1.374881437	2.825096017	2.047129115
ARL7	0.911665392	1.358146244	2.958856469	2.559224674
TFRC	0.321823228	1.35686249	2.475197639	1.475175583
SIDT1	0.607821522	1.354653326	2.88159093	2.571132457
ADAMTS19	0.45826903	1.347244028	2.668048887	3.91519201
CSF1	0.412473108	1.335036887	2.939334787	2.391011937
MYO1F	0.90234382	1.334582492	3.128122299	3.617198276
Q8N5I0	0.856714219	1.334037887	2.736801496	2.579511008
NP_056234	1.020691926	1.323631442	2.753064616	1.954705871
PFKFB3	1.32202806	1.323108554	2.23683459	1.75808628
THY1	0.611109587	1.32011618	3.362712522	2.792191586
Q8WUA8	0.477579058	1.288608014	2.467476245	1.890994835
ARHGDIB	0.468539164	1.31390178	1.935721435	2.718975116
RAC2	0.494809377	1.310443986	2.653321363	2.588489371
GALGT	-0.051317129	1.297837786	2.467374436	2.849554813
CLCN7	0.542319109	1.296818856	2.424005919	1.33314438
NFIB	1.76622848	1.283915724	3.282434201	2.210951789
ARSB	0.362718595	1.274539147	3.168571881	1.563024816
ARK5_HUMAN	1.143308169	1.27311259	2.691167775	2.25701816
IFI30	0.592087543	1.268896729	2.323116284	2.240364236
PLEKHF1	0.76788855	1.256534553	2.986116439	2.400325086
ITGA8	1.766213011	1.247540935	3.317546057	1.874212302
PARVB	0.449081494	1.24632244	3.295375971	2.006905437
PTK2B	0.690780316	1.245750153	2.262527125	3.480111001

C9	0.7886835	1.245679241	1.427885112	3.731337981
NP_056341	0.761962842	1.245552983	2.693713323	1.367536635
DOK3	0.276264308	1.242725918	3.245157674	3.57460434
RGS10	0.351565073	1.237001265	2.826145225	2.537838206
SOX18	1.093716153	1.235715117	2.256388178	1.508936926
BCAR3	-0.146286529	0.749077514	2.906178444	1.204559751
BOP1	0.767222878	1.223769753	1.981680127	2.060553807
RARRES1	0.473704159	1.223012904	2.403970079	2.848894468
FAM40B	0.729100261	1.222402424	3.560913096	1.821271308
INPP5D	0.427345953	1.219139245	2.649284634	3.053892347
BCL2L11	0.723092658	1.213056777	2.767850203	2.629435746
WAS	0.610466342	1.211510866	2.085108175	2.490853505
DSCR1L1	0.622708206	1.209648929	3.197511759	2.663350887
RASSF4	0.320379895	1.202890327	2.574144453	2.314470154
PDGFRB	0.876261121	1.201539137	2.236955757	1.321421566
TEC	-0.454845294	1.190385859	3.293744219	2.72362325
CRIP1	0.423545231	1.182977648	1.614944189	1.684624821
Q6ZTY0	0.485401047	1.176809366	2.433150079	3.167868955
MFRP	0.915716323	1.169115211	1.972857258	1.035708657
IQSEC1	0.571547939	1.165908905	1.683259017	2.227649612
CD99	0.546673922	1.152547119	1.9292463	1.571511479
SELM_HUMAN	0.918425128	1.150858941	2.348578121	1.252799879
TIMP2	0.56065451	0.926753597	1.775774181	1.100933186
TRGV9	0.82828407	1.146657978	2.632106066	2.527012624
Q641Q3	0.391985266	1.130648841	2.78416287	2.752097123
AKAP13	0.376565798	0.839437889	1.237010592	1.714725919
SYNE1	1.030834578	1.109263618	2.473159596	2.072580079
SMPDL3A	0.774293833	1.105790732	2.557872548	2.463698352
ADRBK2	1.382937392	1.103790948	2.083822214	1.850762415
SH3BGRL3	0.403422006	1.101877756	2.254240501	2.274815183
SNN	0.413333782	1.098245244	1.514380202	1.304290524
GOLGB1	0.127985793	0.202018949	0.608284301	0.631492179
MMP11	0.661382505	1.084764753	2.144389772	1.517219427
Q8N682	0.600761285	1.083750283	2.511046709	3.230866406
MER1_HUMAN	0.860136764	1.081443725	2.496883579	2.111252631
ELMO1	0.428636815	1.077037131	2.410250064	1.789387053
MYO1E	0.568906955	1.077026527	2.723237132	2.120974546
Q9NX17	0.439686834	1.069473889	1.784261477	1.99650758
NP_065850	-0.080421191	1.05435568	2.321424209	2.692925479
DYRK3	0.70177097	1.05339238	2.178118936	1.519090702
PTPRC	0.335729487	1.053136908	2.394687269	2.931935692
PPGB	0.766097203	1.053091263	1.586943548	1.365526292
LOXL1	0.878645363	1.050343166	2.055777565	1.027742967
ANGPTL2	1.111166349	1.034366051	2.517474568	1.444061844

B4GALT5	0.21180962	0.281003803	1.083286585	1.033460815
CST3	-0.037259532	1.027753505	1.938035558	1.932808252
SIAT8D	1.767421823	1.022529506	2.174964919	2.134273817
KLF13	0.76175411	1.011301234	1.61638118	1.37655903

Table 4.2 List of Downregulated Genes; These are expressed at all the three time points across the duration of spontaneous tumour regression process (t_2, t_3, t_4), each time point being at three-weekly intervals.

Gene name	log FC value (t1)	log FC value (t2)	log FC value (t3)	log FC value (t4)
Q9BWC9	-0.778302796	-1.015788359	-2.156635552	-2.172318848
NRXN3	-0.172774239	-1.021559342	-1.63262697	-1.260691859
C10orf78	-0.474428959	-1.028963041	-1.813235048	-1.463048137
HELLS	-0.565900211	-1.03684863	-2.072570951	-1.796169981
MARK1	0.121983081	-1.043934331	-3.048620321	-3.103759074
EZH2	-0.198661932	-1.051700236	-1.729870816	-1.035924852
RFX4	0.191423576	-1.056148857	-2.102000018	-3.16803969
ESPL1	0.117286224	-1.075303978	-1.858966765	-2.148887897
WDR43	-0.437154702	-1.092558032	-2.04231639	-1.934522982
TRIM59	-0.187159461	-1.10308949	-1.616462984	-1.606243748
C10orf64	-0.228922152	-1.107608712	-1.765242009	-1.025419962
SNX25	-0.511188785	-1.108837505	-2.403702914	-3.507362768
PCNA	-0.637342871	-1.264147839	-2.087509965	-1.431196098
CDH17	-0.64890014	-1.111276772	-1.169173545	-1.091522501
Q9H989	0.058231373	-1.117867969	-1.281723942	-1.766683832
UHRF1	0.009593079	-1.123359542	-1.966816103	-1.68538213
MCM2	-0.282739515	-1.12336248	-1.943996475	-1.953537052
C6orf162	-0.525544159	-1.127448655	-1.96529177	-2.00335052
MCM4	-0.46712429	-1.128008415	-1.716594249	-1.540671652
NASP	-0.104163735	-1.130654034	-1.629537238	-2.018830587
NP_054828	-0.186613658	-1.136862374	-2.265419211	-2.149591673
MCM6	-0.494403944	-1.143385085	-1.611812977	-1.419578549
TMOD3	-0.544254044	-1.14960657	-2.420386683	-2.265753336
HMGCR	-1.397291374	-1.154086866	-2.693516461	-2.276502022
NP_057532	-0.229518585	-1.166597687	-2.295911849	-2.62275489
VRK1	-0.643955977	-1.782286523	-1.880717184	-1.101528531
OSGEP	-0.908947313	-1.180674491	-1.244047966	-1.22603296
ACSL3	0.402089592	-0.680591441	-1.682605574	-4.008399462
USP1	-0.242726464	-0.536545049	-0.731859669	-0.543392156
PHLDA1	-1.105085757	-1.190708196	-2.367776978	-2.685928539
DZIP1	0.207023923	-0.394263638	-0.861843504	-1.351286042
Q86TG7	-0.272171508	-1.213775052	-3.241558241	-6.248721546
CHEK1	-0.510946563	-1.22006972	-1.519230504	-1.091561664
RFC3	-0.495075467	-1.236128608	-1.957867329	-1.535093594
KIAA1333	-0.538701727	-1.236741687	-1.284098396	-1.502435217
SLC23A3	-0.503640128	-0.598466331	-1.358070792	-1.358251691
IMMP2L	-0.885356789	-1.243643922	-1.433700295	-1.882127478
CDC45L	-0.288381579	-1.253833032	-1.803169332	-1.654916567
TIP1_HUMAN	0.499482393	-1.253918706	-3.141184234	-3.284823595
RNF2	-0.27126307	-0.912486333	-2.556648961	-2.709758054

KIAA1430	-0.425315855	-1.261129182	-2.174867101	-2.217561921
NP_079143	0.046623812	-1.28426114	-1.508320567	-1.790381437
DNA2L	-0.58948246	-1.299918763	-1.345274808	-1.380811844
NP_057185	-0.706455141	-0.710245404	-1.579592878	-1.44736332
RACGAP1	-0.325821861	-1.322622072	-1.655618556	-1.142388526
CCNB3	-0.531383464	-2.030643193	-3.001184581	-1.614929107
FANCD2	-0.41180977	-1.332417101	-1.9884613	-1.514389858
CTNNA3	-0.544326561	-1.363460351	-2.279848512	-3.408341485
TOM1L2	-0.332744666	-1.368805944	-2.24766144	-2.558502985
CHAF1B	-0.653256488	-1.372239053	-2.100016391	-2.246140044
ECT2	-0.340661145	-1.374566246	-2.247228371	-1.390251973
NP_055544	-0.185644559	-1.392432733	-2.403230328	-1.64500586
NP_955389	-0.39682472	-1.397590098	-1.554076782	-1.49213562
Q96JN1	-0.041015441	-1.407295567	-1.814990515	-1.103383991
Q86T96	0.413094555	-0.705971563	-2.357813803	-3.758540745
NSD1	-0.540688432	-1.439432963	-2.287907819	-1.443226599
CCNA2	-1.338388721	-2.449128131	-2.607963513	-1.924035845
NUSAP1	-0.221181625	-1.443067285	-2.176667858	-1.6842026
PTTG2	-0.257354664	-1.44754692	-2.444475008	-1.903216004
Q9BSJ6	-0.074416788	-1.4504619	-2.265600404	-3.008761333
UBE2C	-0.477136272	-1.452705783	-2.19109403	-2.070236083
C21orf45	-0.532513241	-1.472043087	-2.030784313	-1.204433365
Q8N340	-0.097964979	-1.478225761	-2.193210889	-1.981514726
TRPC5	-0.440699451	-1.495540047	-2.19451937	-1.865962344
TPX2	-0.382245463	-1.500688333	-1.975816611	-1.586811669
MAD2L1	-0.53582931	-1.501542322	-2.171823575	-1.586253097
SPAG5	-0.336161959	-1.503124725	-2.107132507	-1.211044004
CACNA2D1	-0.562127272	-1.518813166	-1.739286087	-3.503323131
PAFAH1B2	-0.996147511	-1.522887266	-1.564808493	-1.440002946
ANLN	-0.343856033	-1.530966988	-1.888831611	-2.019678674
BIRC5	-0.464688583	-1.532222919	-2.334150259	-1.817261363
Q9BSD3	-0.537995857	-1.541810026	-1.930032321	-1.391021242
Q14691	-0.648594663	-1.548965084	-2.389399732	-1.807298769
KPNA2	-0.379839008	-0.729526725	-1.16806851	-0.548919252
RAB6B	-0.078906723	-0.640049852	-1.799620831	-3.994266052
PRC1	-0.27257583	-1.569510232	-2.145663695	-1.896278896
PLK4	-0.400348021	-1.590777827	-2.071109452	-1.496781295
APOF	-0.162354168	-1.601162134	-2.309283281	-1.748733196
BUB1B	-0.205789146	-1.608319671	-2.296878874	-1.649891329
POLE2	-0.627601564	-1.659045419	-2.041985081	-1.49675969
KIF2C	-0.512553164	-1.670447432	-2.526961521	-1.901272827
Q6ZN37	-0.476543776	-1.677050225	-2.352222948	-1.398827531
CDC6	-0.699644219	-1.681445914	-2.984634691	-2.93949199
TMPO	-0.493279884	-0.842426331	-1.513722359	-0.946551112

KIAA0101	-0.49776834	-1.682476081	-2.454238336	-1.873329934
NEK2	-0.288711507	-1.684222677	-2.595551506	-2.144489555
Q7Z3Z3	-0.440417859	-1.692090132	-2.210455697	-1.404460941
CENPF	-0.054143053	-1.692208041	-1.991961637	-1.803660711
CDK11	-0.60824927	-1.705278134	-2.507175348	-1.958300035
DIAPH2	-0.667057166	-1.732465946	-2.522852484	-2.022152041
GJB7	-0.586619879	-1.748273763	-3.912091056	-4.27951304
CDCA1	-0.362911152	-1.760153411	-2.17885349	-1.428408537
Q9HCJ3	-0.607985525	-1.77161714	-1.735751124	-1.734709566
DLG7	-0.569755961	-1.786881599	-1.619228513	-1.358840686
KIF20A	-0.211683269	-1.807779605	-2.478033628	-2.293056648
CKAP2	-0.879823499	-2.281943625	-2.59677303	-2.439519126
TOP2A	-0.459517001	-1.810469734	-2.25435386	-1.672954085
Q969N5	-0.851239484	-1.825079181	-2.489520424	-2.105774916
NP_060230	-0.68684981	-1.8467457	-3.14040826	-2.848031493
KNTC2	-1.159244386	-1.998700486	-2.749066739	-1.351287417
POLQ	-0.639765042	-1.86094413	-2.355763656	-1.288694768
C18orf24	-0.355347838	-1.870578513	-2.456229027	-2.864675412
CCNB1	-0.636137771	-1.936436546	-2.738281085	-2.068221461
CLSPN	-0.753978486	-1.939985409	-2.749631716	-1.583439531
HMMR	-0.544442829	-1.974174448	-2.377854905	-1.686304565
NP_060962	-0.526570511	-2.008224315	-2.610980606	-2.567325696
MELK	-0.705602703	-2.021191036	-3.120674018	-2.13368805
KIF4A	-0.439975213	-2.02904303	-2.743355193	-1.951664135
HSD17B12	-0.626462474	-2.068943478	-2.326125246	-2.235404174
GSTM3	-0.354856623	-2.069707085	-3.468384665	-2.825538031
Q9BQ75	-0.988999746	-2.125614954	-2.566584083	-2.505791567
NP_006470	-0.866296535	-2.212584635	-3.289193264	-2.476289109
KIF11	-0.896395986	-2.28949233	-2.7775018	-1.959737739
DEPDC1B	-0.591050835	-2.348739089	-3.306392516	-1.958524078
E2F7	-0.943263438	-2.41215193	-3.059196198	-2.026485912
CDC2	-1.09722215	-2.415204476	-3.692431714	-2.26071039

4.3.5 Protein-Protein Networking analysis

We performed Gene ontology (GO) and Pathway enrichment analysis utilizing multiple databases, like KEGG pathway [(https://www.genome.jp/kegg/: release 102.0 [106]], DAVID database[107], FunRich[98], and ClueGo[108], with $p<0.05$ as the cut-off criterion. The Gene Ontology (GO) of DEGs classified the differentially expressed genes

in terms of three main aspects of the genes: Cellular components (CC), the Biological processes (BP) in which it participates, and its Molecular function (MF). Our results follow:

Cellular component:

The upregulated genes were enriched in the Plasma membrane (Figure 4.4a) and maintained in the Tertiary granule, as shown in figure 4b. In contrast, the downregulated genes were involved in Nucleoplasm, Microtubila, Chromosome, Nucleus, and Microtubule cytoskeleton (Figure 4.4c, 4.4d).

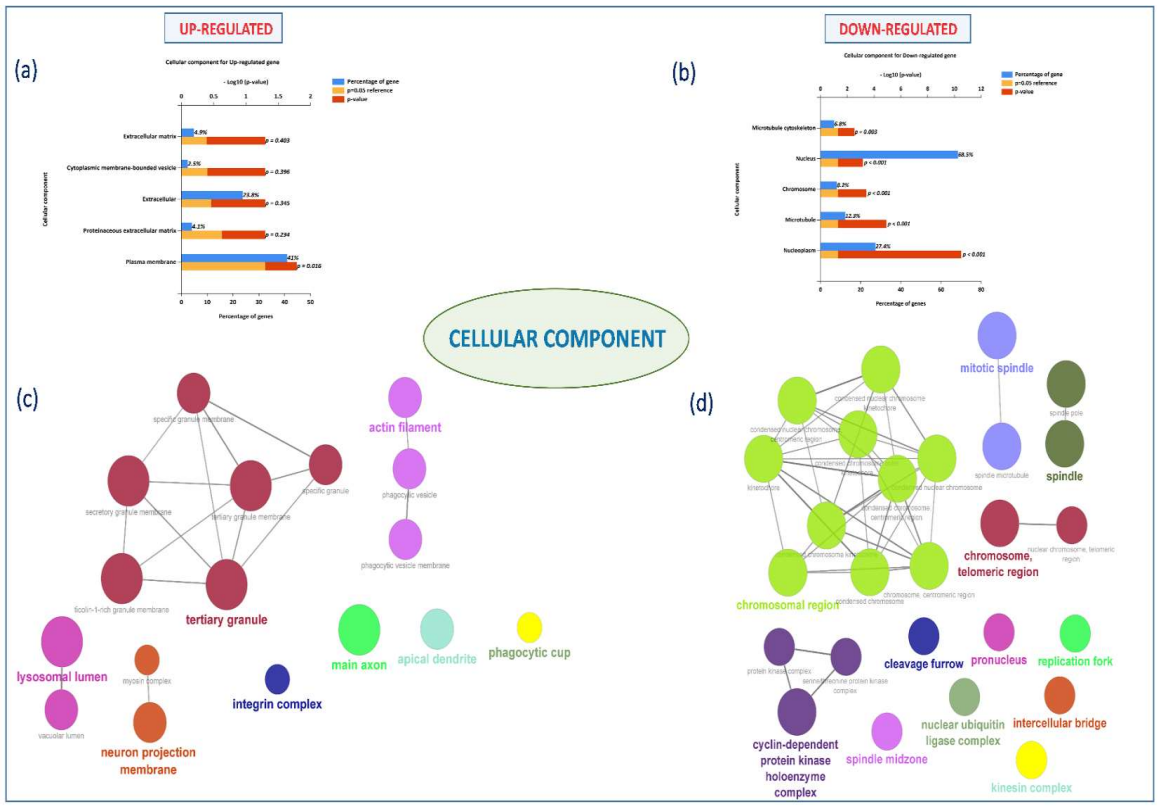


Figure 4.4 Gene ontology (GO) analysis and significant enrichment of "Cellular Component" responsible for exponential decreasing curve in melanoma regression model; **(a, b)** GO analysis for upregulated differentially expressed genes (DEGs); **(c, d)** GO analysis for downregulated DEGs.

Molecular functioning:

The upregulated genes were mainly engaged in Virus receptor activity and Metallopeptidase activity (Figure 4.5a, b). In contrast, the downregulated genes were related to Kinase binding and also Cyclin-dependent protein serine/threonine kinase activity (Figure 4.5c, d).

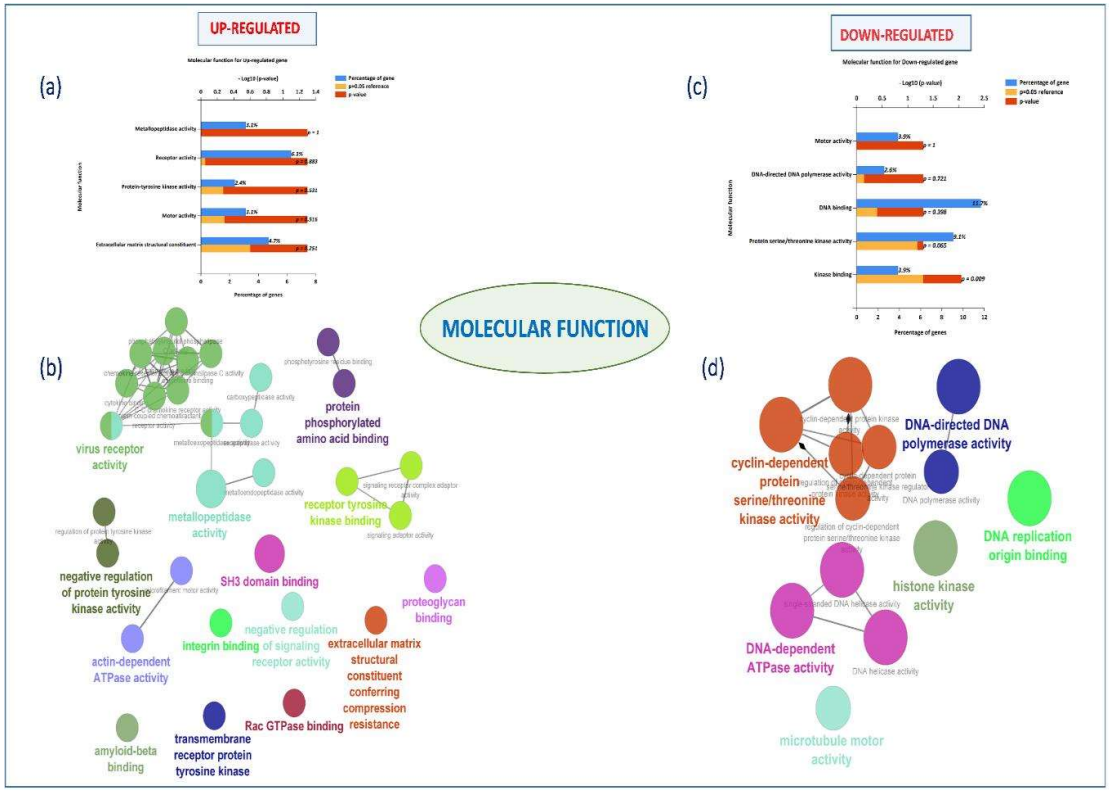


Figure 4.5 Gene Ontology analysis and significant enrichment of "Molecular Function" responsible for exponential decreasing curve in melanoma regression model; **(a, b)** GO analysis for upregulated DEGs; **(c, d)** GO analysis for downregulated DEGs.

Biological processing:

The upregulated genes were involved in Cell growth or maintenance, and Leukocyte migration activity (Figure 4.6a, b, c), whereas the downregulated genes were involved in regulation of Mitotic cell cycle (Figure 4.7a, b, c). The GO term analysis showed that most of the DEGs genes were enriched in Immune cell activation, Leukocyte migration, Serine Kinase activity, Cell cycle, Cell growth, and Mitotic cell cycle regulation.

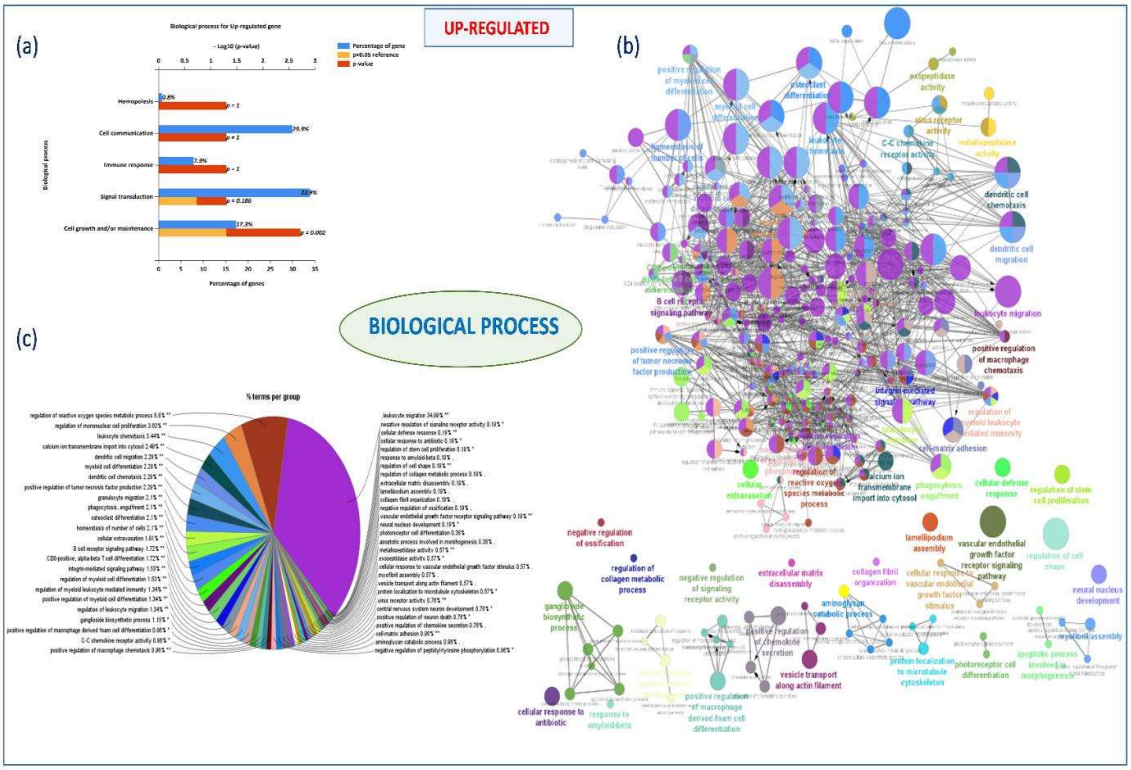


Figure 4.6 Gene Ontology analysis and significant enrichment of "Biological Process" responsible for exponential decreasing curve in melanoma regression model; **(a, b)** GO analysis for upregulated DEGs; **(c)** pi-chart representation of GO analysis.

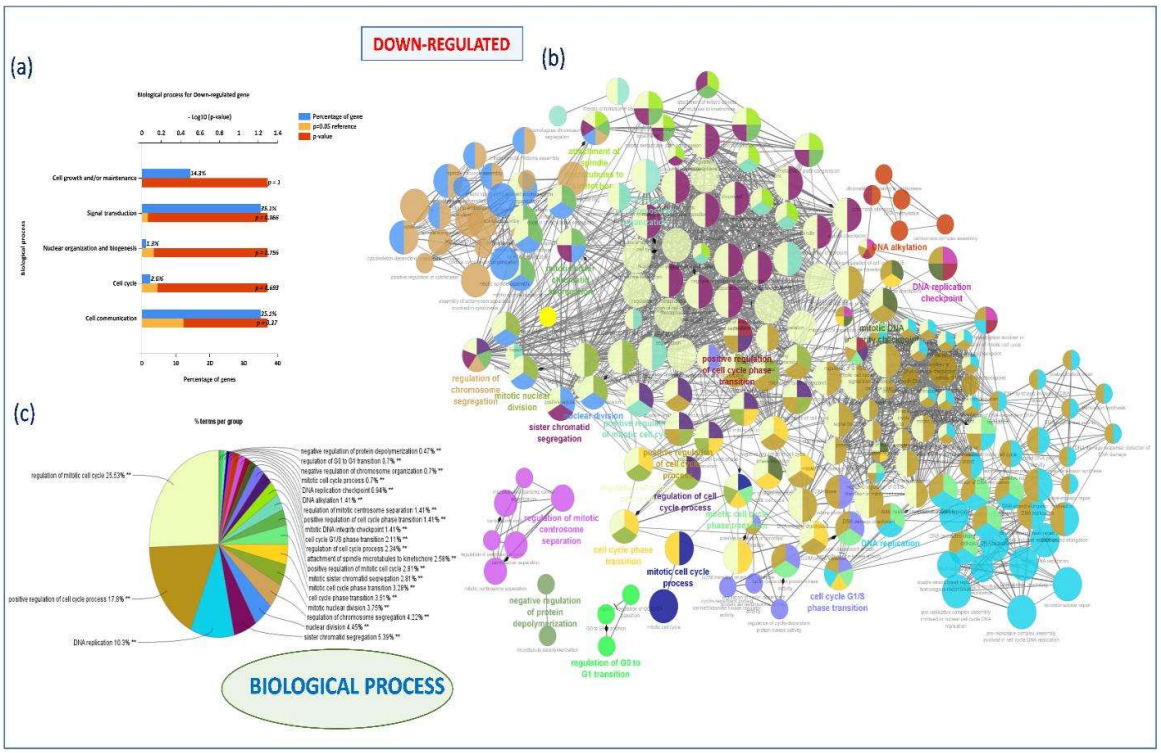


Figure 4.7 Gene Ontology analysis and significant enrichment of "Biological process" responsible for exponential decreasing curve in melanoma regression model; **(a, b)** GO analysis for *downregulated* DEGs; **(c)** pi-chart representation of GO analysis.

4.3.6 Functional Enrichment analysis

We found from the enrichment data that two classes of genes are responsible for tumour regression. We also did find that the first class of genes activated the cellular immune system and leucocyte regulation, thus enabling these immune effector cells to damage and eliminate the tumour cells. On the other hand, the second class of genes were observed to be involved in blocking the cell cycle regulation or spindle formation, so the malignant cells would not be able to multiply. Thus these two classes of genes operate complimentarily for achieving full tumour regression, one class attacks the tumour cells, and the other class makes the tumour cells non-functional, without being able to multiply. Furthermore, enrichment analysis with different databases showed that the most significant

GO terms for tumour regression belong to downregulated genes. Accordingly, biological pathway enrichment was performed which showed that the downregulated genes were enriched in the following entities (Figure 4.8):

Cell cycle, Mitotic phase, DNA replication, Checkpoints of Cell cycle,

Mitotic M-M/G1 phases, and G2/M DNA damage checkpoint.

To highlight, this finding signifies that one can control the tumour cell population and actuate the tumour regression process by intervening in the cell cycle regulation process and mitotic spindle formation process.

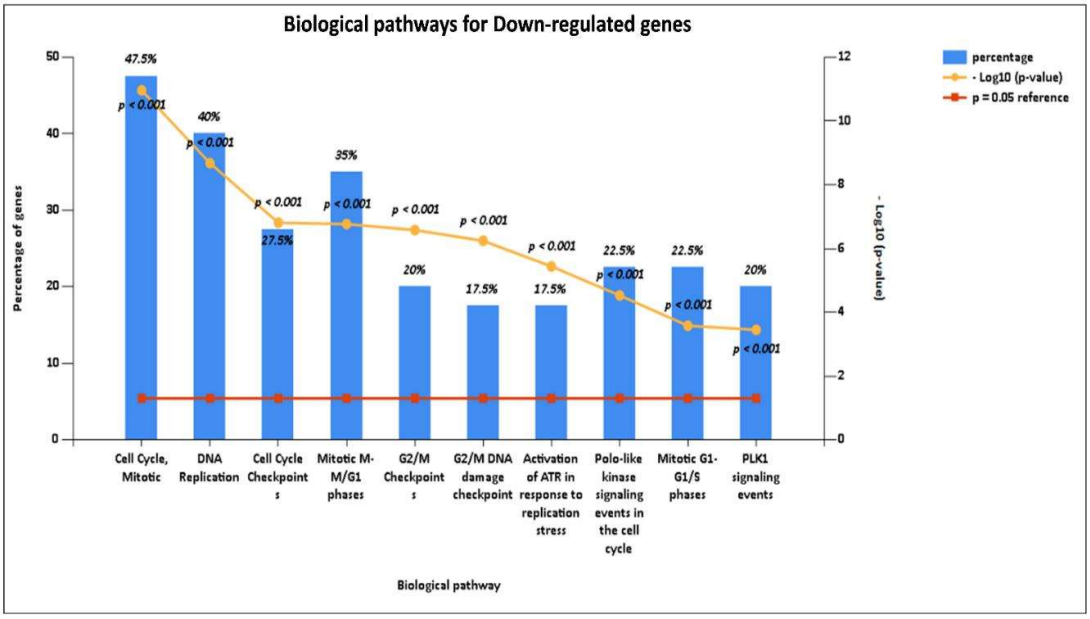


Figure 4.8 Gene ontology analysis and significant enrichment of biological pathway responsible for exponential decreasing curve in the melanoma regression model for downregulated differentially expressed genes. The blue bars denote the percentage of genes, while the yellow line is the level of statistical significance of the genes [$-\log_{10}$(p-value)]. Note that there is very significant downregulation of the genes associated with tumour cell multiplication, such as Mitosis function, DNA replication, Cell cycle checkpoint functionality, Kinase signaling, etc. Such inhibition of malignant cell reproduction produces rapid decline in tumour cell population.

4.3.7 Protein-Protein interaction analysis and Sub-network analysis

We undertook the protein-protein interaction (PPI) analysis by utilizing STRING database facility in Cytoscape platform to retrieve the interacting genes. We mapped a total 176 upregulated genes (DEGs) into Cytoscape; out of these, 164 DEG genes of humans were recognized using Cytoscape, there being 428 edges (Figure 4.9). Similarly, out of 116 downregulated DEGs, 115 genes were mapped into Cytoscape, having 1797 edges (Figure 4.10). Thereafter, we selected the top three significant clusters within the Protein-Protein interaction network using the MCODE facility in the Cytoscape platform (Module 1, MCODE score = 54.4; Module 2, MCODE score = 8.444; Module 3, MCODE score = 6). Our MCODE analysis shows that Module 1 consists of 61 nodes and 1632 edges (Figure 4.11a), Module 2 consists of 10 nodes and 38 edges (Figure 4.11b), while Module 3 consist of 16 nodes and 45 edges (Figure 4.11c). We have also performed the biological pathway enrichment analysis for the most significant item, Module 1, and these were mainly associated with the following entities (Figure 4.11d):

Mitotic M-M/G1 phases, G2/M checkpoints, Mitotic process, Cell cycle,
DNA Replication, G2/M DNA damage checkpoints.

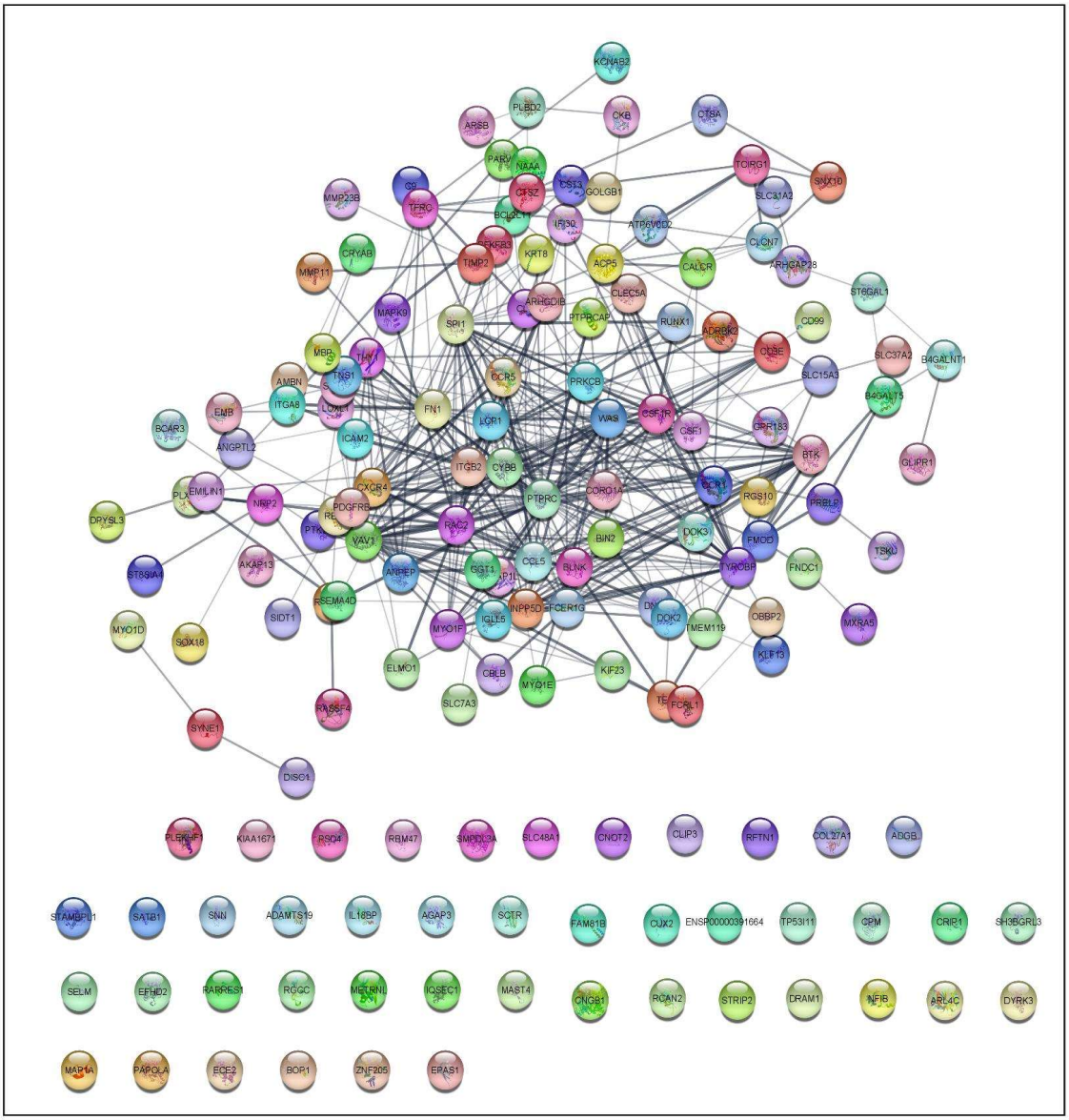


Figure 4.9 Relationship of 164 Human Upregulated DEGs (differentially expressed genes) and their 428 linking edges during spontaneous tumour regression, as obtained by Protein-Protein Interaction analysis

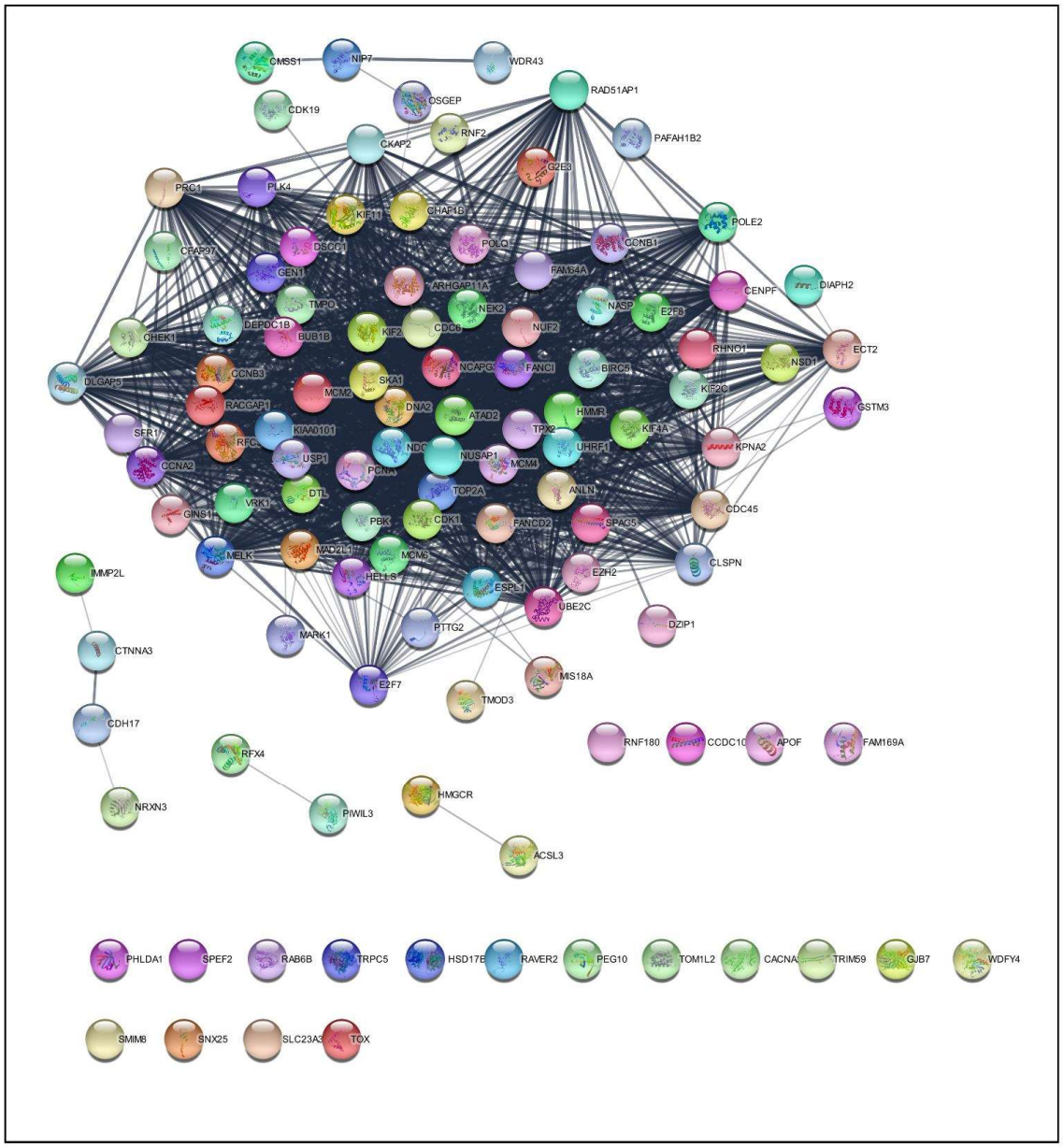


Figure 4.10 Relationship of 115 Human Downregulated DEGs (differentially expressed genes) and their 1,797 linking edges during spontaneous tumour regression, as obtained by Protein-Protein Interaction analysis.

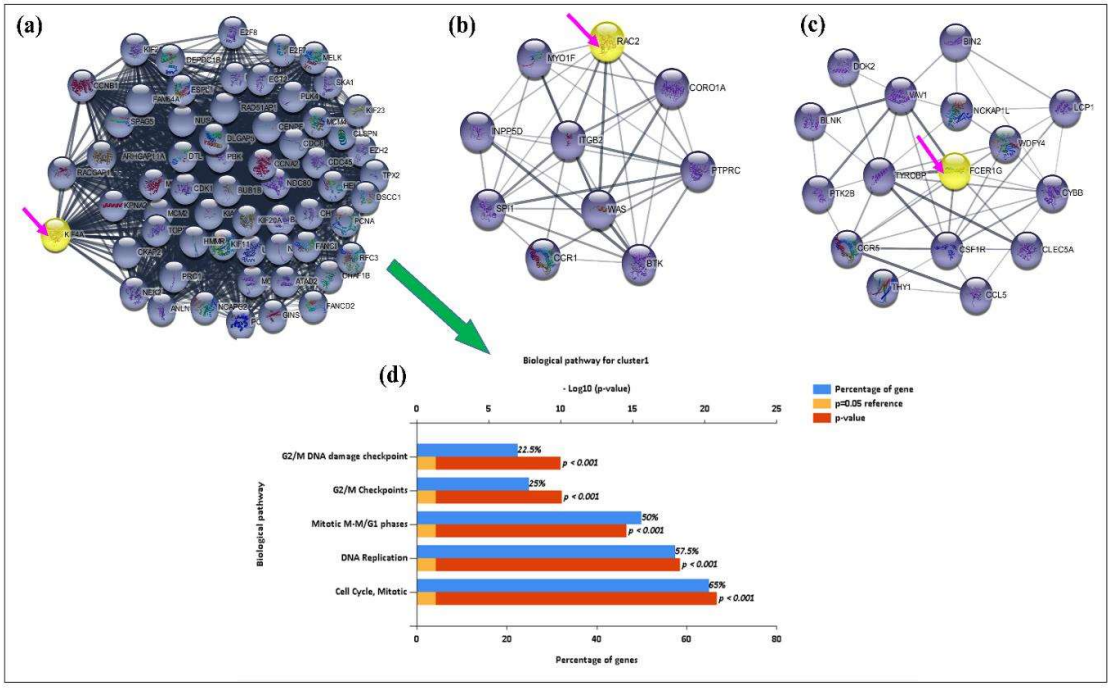


Figure 4.11 Sub-network of top 3 modules with seed node representation (in yellow color) from Protein-Protein Interaction (PPI) network: **(a)** Module-1, having KIF4A gene as a seed node; **(b)** Module-2 having RAC2 gene as a seed node; **(c)** Module13 having FCER1G gene as a seed node; **(d)** Biological pathway enrichment analysis of Module 1.

4.3.8 Identification of Hub Genes:

We demarcated the effect of gene functioning by utilizing the CytoNCA facility in the Cytoscape procedure, having different types of centrality analysis to identify the essential proteins in the biological network[100]. Utilizing the CytoNCA method, we used the three centrality measures to assess the degree of influence of a gene's function Degree [109], Betweenness [110], and Closeness [111]. Hence, from the PPI network with high degree of connectivity (Table 4.3), we identified the following configuration of genes, that can be ranked using the effective centrality score, which is the product of the aforesaid three centrality measures:

Top 10 hub genes:

CDK1, CCNA2, KIF23, TOP2A, CCNB1, CHEK1, KIF11, BUB1B, RAD51AP1, KIF20A.

Using Cytoscape, we constructed the PPI network to show the interconnection between the top 10 hub genes (Figure 4.12a), having 10 nodes and 45 edges. We utilized the STRING database to indicate the gene co-expression analysis of the ten hub genes, which shows that these genes may actively interact with each other (Figure 4.12b). These findings suggest that the functioning of these ten hub genes might play a crucial role in the spontaneous regression of malignant melanoma tumours, manifesting in the first-order kinetics of the tumour decline. For example, considering the first hub gene above (CDK1), we noted that the overexpression of CDK1 indeed has a gross tumourigenic potential in melanoma [112] and this CDK1 will interact with the other aforesaid hub genes (CCNB1, CHEK1, BUB1B and CCNA2) to activate the cell cycle, to make the tumour grow intensively and become invasive. In other words, under-expression of these genes (as CDK1) or pharmacological inhibition of the effect of these genes may induce melanoma regression.

Table 4.3 Top 10 hub genes identification: CytoNCA analysis of PPI network

Display name	Degree	Betweenness	Closeness	Effective Score
TOP2A	69	1980.160771	0.016653687	2275.411484
KIF20A	66	1846.282413	0.016644713	2028.235524
KIF23	69	1227.289816	0.016646707	1409.693009
CDK1	72	1174.577395	0.016636744	1406.962366
CCNB1	68	1141.440679	0.01663774	1291.387552
RAD51AP1	66	994.5498801	0.016621824	1091.061357
BUB1B	66	422.550881	0.016622818	463.583088
CHEK1	67	393.7200322	0.016630773	438.7071859
CCNA2	69	335.982852	0.016632763	385.5942962
KIF11	67	212.022604	0.016623812	236.1497951

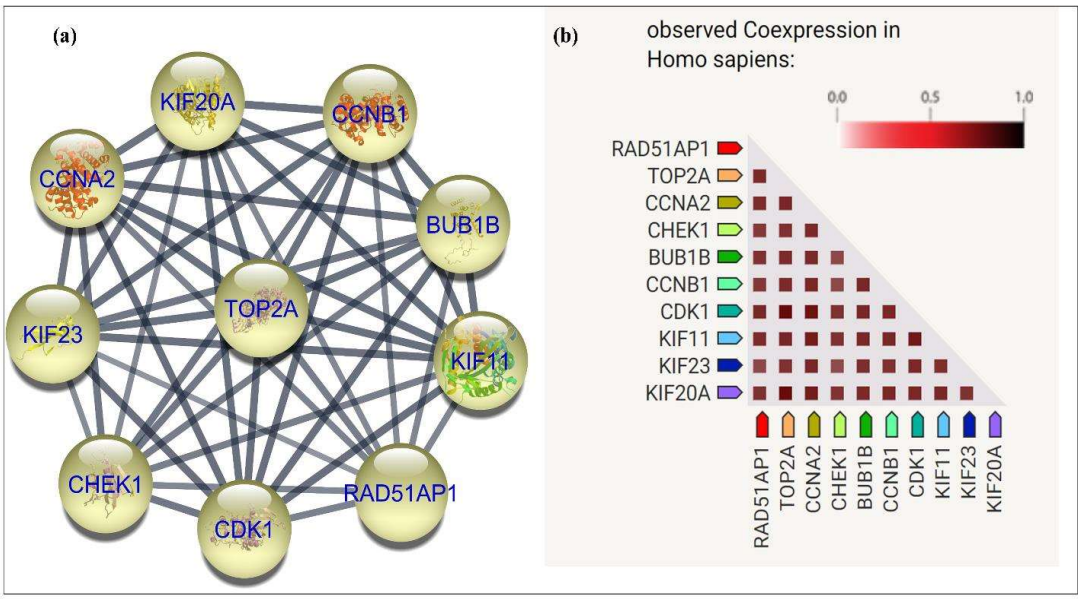


Figure 4.12 (a) Construction of Protein-Protein Interaction network of the top 10 hub genes, having 10 nodes and 45 edges; **(b)** The co-expression analysis of the 10 hub genes using the STRING platform, the color intensity of the triangle-matrix shows the level of confidence between two functionally connected proteins.

4.3.9. Identification of Candidate Drugs

We have earlier identified the hub genes and their functionality which are responsible for the first order decline of the malignant cells, enabling complete spontaneous tumour regression endogenously. The apparent question is that how can one induce a similarly efficient regression process exogenously, i.e. what could be candidate drug molecules that could function likewise to mimic the expression of the genes, so that if these candidate drugs are administered therapeutically to a progressing malignant melanoma tumour, then the tumour would undergo complete regression. Accordingly, we identified 34 drugs by using the top 10 hub genes by further probing drug-gene interactions. Out of these 10 hub genes, we found that incisive targets of these drugs include genes: CDK1, CCNA2, TOP2A, and CHEK1 (Table 4.4).

Table 4.4 Identification of candidate pharmacological molecules.

Serial No.	Gene	Candidate pharmacological molecule	Interaction types
1	CDK1	SERTRALINE	N/A
2	CDK1	RUCAPARIB	N/A
3	CDK1	FENOFIBRATE	N/A
4	CDK1	CINNARIZINE	N/A
5	CDK1	CLOFIBRATE	N/A
6	CDK1	CLOTRIMAZOLE	N/A
7	CCNA2	TAMOXIFEN	N/A
8	TOP2A	DIGITOXIN	N/A
9	TOP2A	DOXORUBICIN HYDROCHLORIDE	inhibitor
10	TOP2A	ETOPOSIDE PHOSPHATE	inhibitor
11	TOP2A	VINCRISTINE	N/A
12	TOP2A	IDARUBICIN	inhibitor
13	TOP2A	EPIRUBICIN	inhibitor
14	TOP2A	VALRUBICIN	inhibitor
15	TOP2A	TENIPOSIDE	inhibitor
16	TOP2A	AMSACRINE	inhibitor
17	TOP2A	DEXRAZOXANE	inhibitor
18	TOP2A	HYDROQUINONE	N/A
19	TOP2A	DAUNORUBICIN CITRATE	inhibitor
20	TOP2A	IDARUBICIN HYDROCHLORIDE	inhibitor
21	TOP2A	PACLITAXEL	N/A
22	TOP2A	FLUOROURACIL	N/A
23	TOP2A	DAUNORUBICIN HYDROCHLORIDE	inhibitor
24	TOP2A	ETOPOSIDE	inhibitor
25	TOP2A	DAUNORUBICIN	inhibitor
26	TOP2A	MITOXANTRONE	inhibitor
27	TOP2A	MITOXANTRONE HYDROCHLORIDE	inhibitor
28	TOP2A	DOXORUBICIN	inhibitor
29	TOP2A	PODOFILOX	inhibitor
30	CHEK1	PALBOCICLIB	N/A
31	CHEK1	CISPLATIN	N/A
32	CHEK1	OLAPARIB	N/A
33	CHEK1	GEMCITABINE	N/A
34	CHEK1	ETOPOSIDE	N/A

After comparing three different databases (IPA, DGIdb, and Cytoscape), we found that only one gene, TOP2A is common in all databases among these four genes. Also, from the cBioPortal database, the alteration information of these four genes shows that most of the mutation (8%) was alone in the TOP2A gene (Figure 4.13a). The increased expression of Topoisomerase II alpha (TOP2A) makes melanoma cancer cells resistant to chemotherapy[113]. According to an analysis, TOP2A inhibition induces cancer cell death by damaging DNA with a low toxicity effect to normal cells[114]. Thus, TOP2A functionality should be a potential biomarker for melanoma treatment. Hence we are focus more on the drugs related to the TOP2A gene. From three databases (IPA, DGIdb, and Cytoscape) we arrived at a total of 9 drugs that are common in all databases (Figure 4.13B), and all these drugs modulate Topoisomerase II alpha, these candidate drugs can be categorized as per their functionalities as:

Anthracycline derivatives: Idarubicin, Epirubicin, Valrubicin, Daunorubicin, Doxorubicin

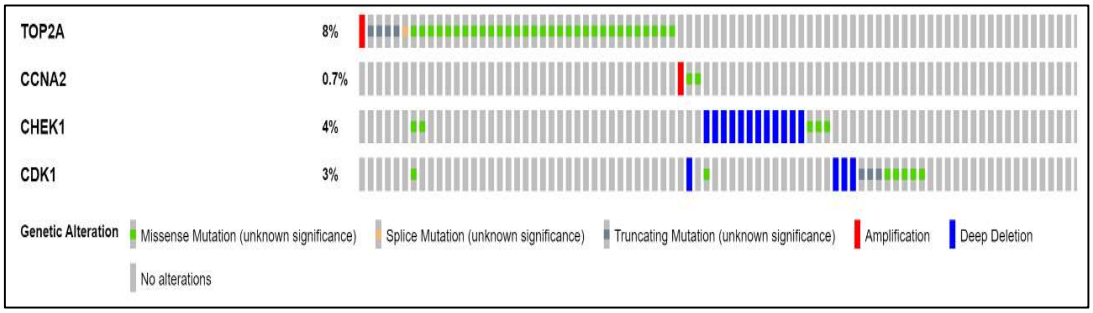
Podophyllin derivatives: Teniposide, Etoposide,

Adduction derivative: Mitoxantrone

Chelation derivative: Dexrazoxane.

Some of these drugs are repurposed here, for instance, Dexrazoxane is a cardioprotective drug and anti-malarial agent, while Mitoxantrone is drug also used in neuroinflammatory disease, such as multiple sclerosis.

(a)



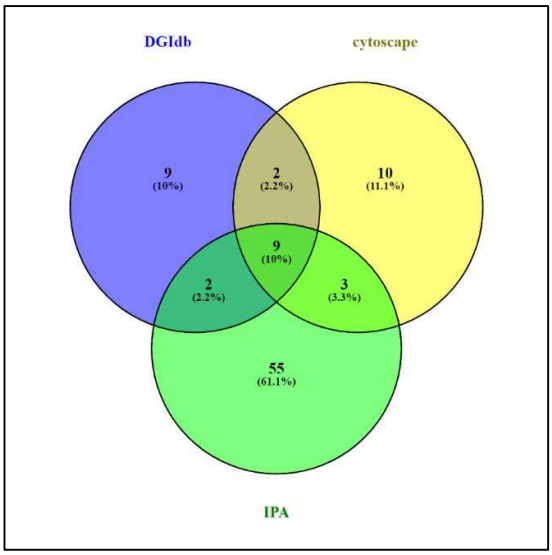


Figure 4.13 ***(a)*** An oncoprint summary across a set of melanoma tumours (skin cutaneous melanoma TCGA Pan-Cancer data) shows the genetic alteration connected with four main genes in a total of 442 patients; **(b)** Common drugs from three databases (DGIdb, Cytoscape, and IPA).

4.3.10 Molecular docking for TOP2A

We have used a molecular in-silico docking study to assess the affinity of the TOP2A gene with respect to the aforesaid nine drugs (Idarubicin, Epirubicin, Valrubicin, Teniposide, Dexrazoxane, Etoposide, Daunorubicin, Mitoxantrone, Doxorubicin), utilizing AutoDock analysis. Here we have taken a total of 12 amino-acid residues (GLY721, LYS723, GLN726, GLN773, ASN779, ASN851, GLY852, ARG929, LEU722, LEU771,

ALA853, and TRP931) for TOP2A [114]. Those nine drugs were docked with TOP2A receptors, for acting as inhibitors. Four of these nine drugs are found to have excellent inhibitory action on the TOP2A receptor based on its binding energy. (Table 4.5). The highest-ranking drug is the podophyllin-derivative Tenoposide which uniquely has a sulphur-based thiophene-derived 5-membered ring. The docking study of these four ligands (two drugs from podophyllin derivative and two from anthracycline derivatives) with their receptors are shown in Figure 4.14. Detailed information regarding the interacting amino acid residues and the respective binding energies is in Table 4.6.

Table 4.5 List of all the Nine Drugs with their binding energy and inhibitory constant (Ki) value, exhibiting inhibitory action on TOP2A gene (PDB ID: 4FM9).

Sl. No.	Drug	Total binding energy(Kcal/mol)	Calculated inhibitory constant (nanomolar)
1	Teniposide	-9.95	50.92
2	Etoposide	-9.51	107.27
3	Epirubicin	-8.77	372.72
4	Doxorubicin	-8.64	461.78
5	Daunorubicin	-8.57	520.54
6	Dexrazoxane	-6.83	9.78
7	Idarubicin	-6.35	22.22
8	Mitoxantrone	-4.29	721.98
9	Valrubicin	-3.32	3.69

Table 4.6 The binding energy, inhibitory constant (Ki) and interacting amino acid residues of the candidate drugs Teniposide, Etoposide, Epirubicin, and Doxorubicin against TOP2A receptor (PDB ID: 4FM9).

Drug	Total binding energy (Kcal/mol)	VDW +H bond + dissolution energy (kcal/mol)	Calculated inhibitory constant (Ki-Nanomolar)	Interacting amino acid residues
Teniposide	-9.95	-12.63	50.92	LYS723 GLN726 GLU854 TRP931 GLY796 LYS798
Etoposide	-9.51	-12.49	107.27	LYS723 ASN770 ASN851 PHE775 GLY777 LEU783
Epirubicin	-8.77	-12.05	372.72	ASN779 GLU854 LUE771 ARG929 GLY852 PHE775
Doxorubicin	-8.64	-11.92	461.78	ASN779 GLU854 GLY721 ALA853 GLY852 ASN851

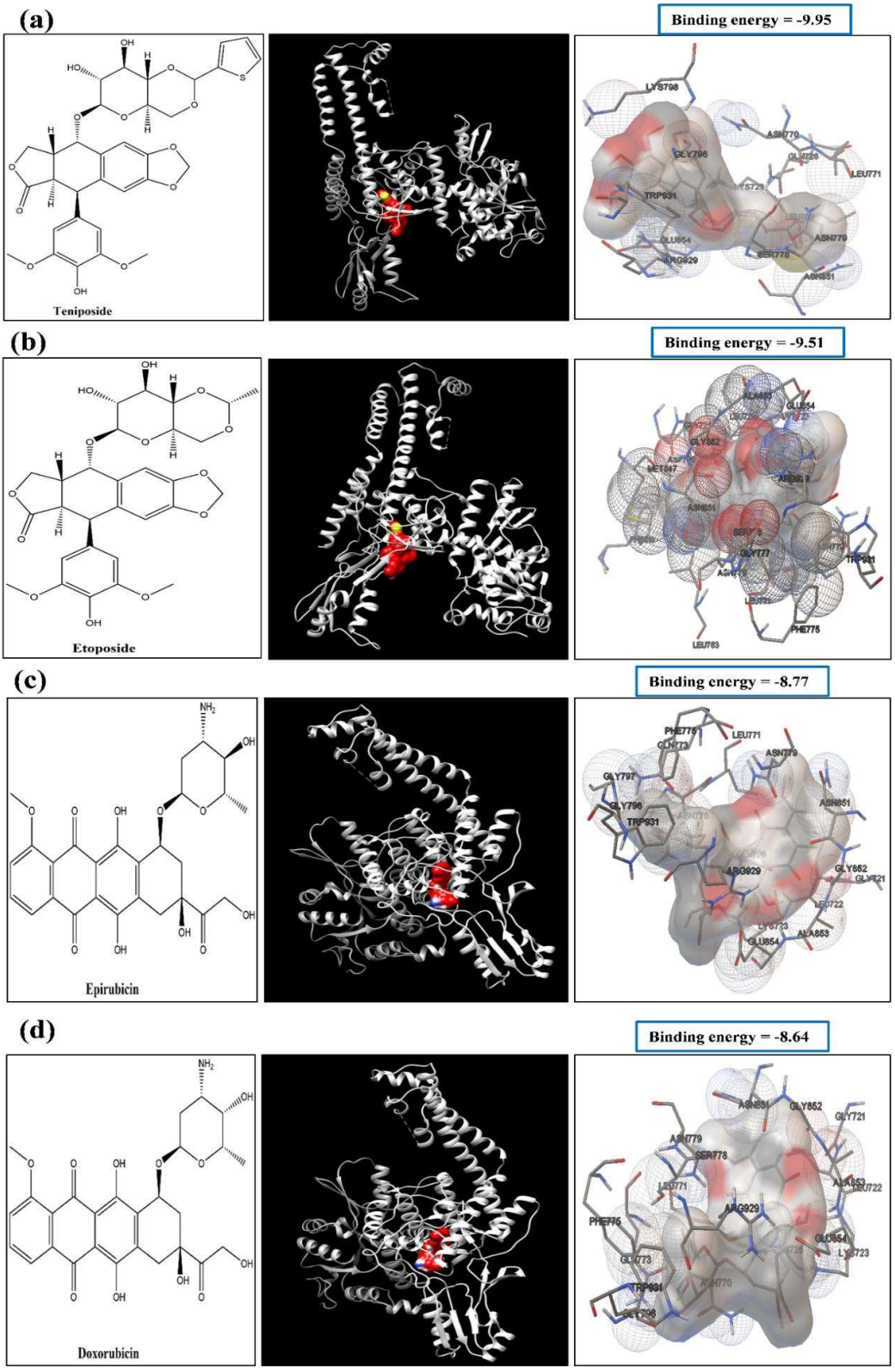


Figure 4.14 Computed structural comparison and binding features respectively of **(a)** Teniposide, **(b)** Etoposide, **(c)** Epirubicin, and **(d)** Doxorubicin against TOP2A receptor (PDB ID: 4FM9) (Visualization using UCSF Chimera and AutoDock).

4.3.11 Corroboration from Human Subjects and Clinical Patients

We have identified 34 drugs by using the top 10 hub genes to explore the drug-gene interaction, which has the possible involvement in the exponential decreasing curve of spontaneous cancer regression (Table 4.4). Regarding these 10 hub genes (Table 4.3), the promising targets of these drugs includes genes: CDK1, CCNA2, TOP2A, and CHEK1 (Table 4.4). These four hub genes are now further investigated to narrow down our search. Here we have used two techniques for garnering more information about the potential genes, these methods respectively utilize (i) GEPIA platform("GEPIA (Gene Expression Profiling Interactive Analysis)," 2022), to find the difference in the gene expression levels between melanoma cancer tissue and normal tissue in human subjects, and (ii) THPA platform("The Human Protein Atlas," 2022), for the overall survival analysis of human melanoma patients with different levels of the gene expression. Our findings from these two procedures are respectively shown in Figures 4.15 and 4.16. From Figure 4.15 (based on analysis of 461 melanoma patients and 558 normal subjects are 461 and 558, respectively), we observed that all four genes are significantly expressed in melanoma tissue in humans, and among these genes, the TOP2A gene has the highest mean absolute value (near around 4.5). Furthermore, we note from Figure 4.16 that human melanoma tumour patients with low expression levels of TOP2A gene have the lowest 3 years' survival rate (at 53%), when compared to genes CDK1, CCNA2, and CHEK1(survival at 59%, 60%, and 60% respectively). This analysis shows that (i) our approach of the focal significance of TOP2 gene is also applicable to the scenario of human melanoma patients (ii) among these four genes, investigating TOP2A gene may have high relevance to understanding the possible

mechanism behind spontaneous cancer regression, as well as to development of novel therapeutic molecules in the human context.

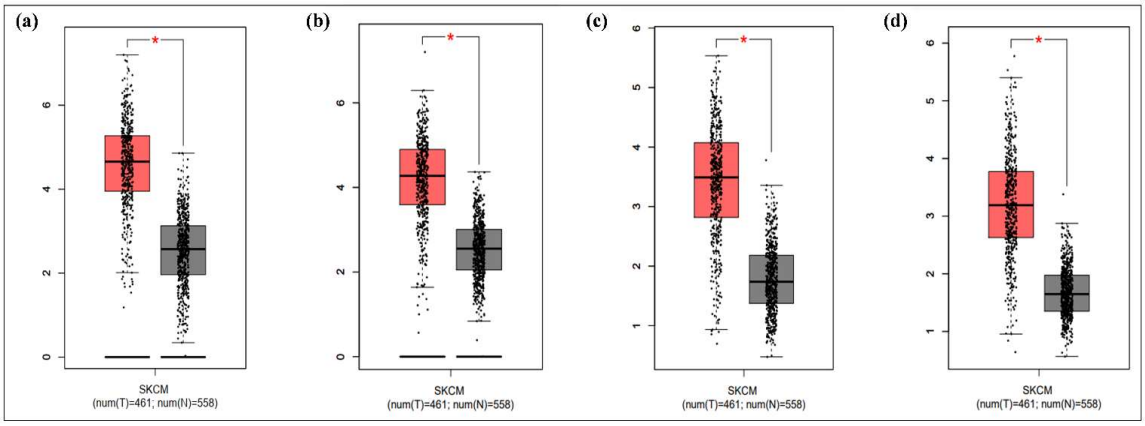


Figure 4.15 Expression level of four hub genes in human subjects and clinical patients. The four panels of the heat maps show the expression levels of the four hub genes [**(a)** TOP2A, **(b)** CDK1, **(c)** CCNA2, and **(d)** CHEK1) in melanoma tissue (skin cutaneous melanoma) vis-à-vis normal skin tissue, based on TCGA data analyzed using GEPIA platform. The red and grey box plots in each panel represents melanoma tissues and normal skin tissues respectively.

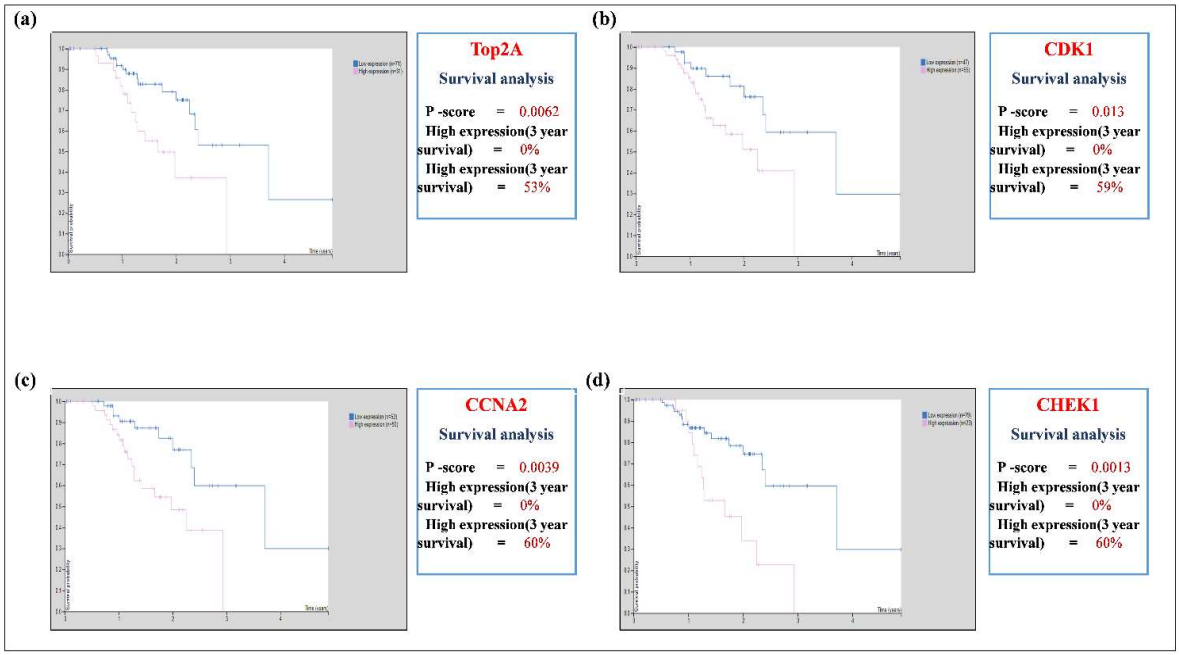


Figure 4.16 Correspondence of Hub Genes with Survival Rates in Clinical Patients of Malignant Melanoma tumour. The Human Protein atlas was used to evaluate the overall survival rates of high and low expression of the genes: **(a)** TOP2A, **(b)** CDK1, **(c)** CCNA2, and **(d)** CHEK1 in patients with melanoma. Blue graphs: low expression of gene; Red graphs: high expression of gene.

4.4 Discussion

In our present study, we have elucidated a systems biology approach using the three aforesaid entities that could be used therapeutically to induce tumour elimination in the clinical situation. We now elaborate the different implications that our investigations have revealed.

Associated genes and pathways:

Numerous studies have been performed to explore the mechanism behind the spontaneous cancer regression process during the past several decades. However, the studies need to be more comprehensive to systematically elucidate the exact factors behind this process. In contrast to prior studies that only focus on eliminating the cancer cells, our investigation is more focused on eradicating the malignant cells while protecting the normal tissue; these two aspects are implemented in the endogenous process of the spontaneous cancer remission phenomenon. As an exemplary case study, we have analyzed microarray data of spontaneous cancer regression in a well-documented mammalian melanoma system and formulated the involved genes and biological pathways that produce permanent tumour elimination from the experimental perspective. Finally, we identified the differentially expressed genes for five different time points (t_0 -t_1-t_2-t_3-t_4) at three weeks apart (total 3 months' duration), while the tumour changes from a progression phase to the regression phase and finally to the permanent eradication phase. We observed that the number of genes with altered levels of expression was maximum after midway, i.e. at time t_3, where the regression process was maximal. We found that the ratio of downregulated genes (with respect to the upregulated genes) increased while the regression process advanced in the following way: 29%-71% -113% -137%. The downregulation of the genes enabled the inhibition of:

Cell cycle, Mitotic phase, DNA replication, Checkpoints of Cell cycle,

Mitotic M-M/G1 phases, and G2/M DNA damage checkpoint.

IPA analyses were performed to find the interactive pathway network, which shows that the first sign of spontaneous cancer regression is at time t_2; here the regression has started 1.5 months after the initial start of progression (time t_0). For the identification of differentially expressed genes (DEGs) for exponential decreasing tumour regression (i.e. first-order kinetic process), we have taken the common DEG genes at time points t_2, t_3, and t_4, since the first sign of spontaneous regression was shown at the time point t_2. Finally, we identified 292 DEG genes for this regression (176 upregulated and 116 downregulated). Furthermore, biological pathway enrichment showed that (i) Downregulated genes were mainly enriched in cell cycle, mitosis, DNA replication, cell cycle checkpoints, Mitotic M- M/G1 phases, and G2/M DNA damage checkpoint, while (ii) Upregulated genes were enriched in cell growth or maintenance, leukocyte migration activity, immune cell activity, and leucocyte regulation. We undertook Protein-Protein Interaction Analysis (PPI) and found the three significant clusters of genes using MCODE, and identified the most effective module consisting of 61 nodes and 1632 edges the module being again enriched in the following processes: cell cycle, mitosis, DNA Replication, Mitotic M-M/G1 phases, G2/M checkpoints, and G2/M DNA damage checkpoints (Figure 4.11).

Nodal Hubs and Candidate Molecules for Inducing Complete Tumour Regression

CytoNCA, was used to find the 10 hub genes (*CDK1, CCNA2, KIF23, TOP2A, CCNB1, CHEK1, KIF11, BUB1B, RAD51AP1, KIF20A*) related to the exponential decreasing phenomenon of tumour regression process. From Table 4.3, the effective score of the TOP2A gene is the highest out of these 10 hub genes, which indicates that TOP2A may be

a potential prognosis biomarker for grading the degree of invasiveness of melanoma cancer. Out of the 10 hub genes, we have identified four potential gene targets of the drugs i.e the genes *TOP2A, CDK1, CCNA2,* and *CHEK1*. Targeting these identified candidate genes can serve as a potential source for gene therapy for permanent tumour regression in different types of cancer. For example, CHK1 gene activation may help to suppress tumour growth [115]. Similarly, CDK1 gene is a prognostic biomarker for the molecular pathology technique of Pan-cancer Analysis, and is also much significant in a broad gamut of oncology, and this gene can be used as a potential gene-targeted therapy (Yang et al., 2022). Likewise, from a gene targeting perspective, the chromosomal decatenation checkpoint is modulated by topoisomerases, and an important player in the process is the TOP2A gene which offers considerable prospects for gene therapy, aimed at personalized cancer treatment (Chen et al., 2015)

The Table 4 shows that most of the drugs were TOP2A inhibitors. Moreover, with the help of the GEPIA procedure and THPA platform, we validated the TOP2A functionality by showing that it has the highest expression rate and lowest survival rate compared to the other three genes in human melanoma patients. Subsequently, we performed molecular docking studies showing that two kinds of drug derivatives (podophyllin derivative and anthracycline derivatives) have high binding energy as TOP2A inhibitors. Indeed, Teniposide (podophyllotoxin derivative) has the highest binding energy (-9.95 kilocal/mole, Table 4.3), which could be the potential drug to target melanoma cancer cells.

Clinical Translation: Permanent Tumour Elimination by First-order Kinetics:

In chapter 3, we have demonstrated how the inputs of cytotoxic T-lymphocytes, DNA blockage, and interleukin-2 need to change time-wise for the melanoma tumour to undergo complete eradication, whether by the endogenous way (spontaneous regression) or by the exogenous way (therapy-induced regression). In the spontaneous regression mode, these inputs are themselves generated by the host tissue. Regarding therapy-induced regression, these three inputs (T-lymphocytes, chemotherapy drug, and interleukin-2) will be administered externally at the time-wise injection rate. Since our present study delineates that the most efficient chemotherapy drug for regression would be Teniposide, one can give an equivalent Teniposide dose (chemotherapeutically, dacarbazine and Teniposide are respectively monovalent and divalent, while their molarities is 182 and 657 respectively), hence the Teniposide dose would be 1.81 times the dacarbazine level given in Figure. 4.14a.

Thus, one can note the benefits of investigating spontaneous tumour regression and its intensive implications for more efficacious therapeutic regression of a tumour. Molecular biology-based analysis of spontaneous regression of different types of tumours would give an insight into what pharmacological agent would be maximally helpful for regressing that type of tumour in the therapeutic setting. Our study shows that most tumour cells decline by first-order kinetics actuated by very few hub genes. A significant factor of this decline is DNA interference, as by small molecule candidate antineoplastic agents that modulate the effect of those hub genes. This is an encouraging sign since antineoplastic chemotherapy agents are generally smaller molecules and comparatively much more affordable than highly sophisticated newer antitumour interventions such as monoclonal antibodies, biosimilars, or growth factor-modulating proteins. Of course, eliminating the small number of residual tumour cells may need interferon-2 and lymphocyte activation. However, the cost involved would not be prohibitive, as most of the tumour load has been eliminated.

Hence it would be desirable that microarray investigations of spontaneous regression cases of different tumours should be done in the future (we recollect there are already 30,000 studies on spontaneous cancer regression available in PubMed, as mentioned earlier). Thereby, for different tumour types, one could have the significant hub genes and the main chemotherapy agents for modulating the gene effects so that complete tumour regression could be induced from the clinical perspective.

4.5 Conclusion

In this study the noteworthy finding is the malignant cell extinction process occurring at the last stage of tumour regression. Here, a small negative biasing (about 1%-2% of original tumour load) eliminates the small number of residual tumour cells under the asymptotic tail of the exponentially declining curve of tumour cell population, thereby the tumour becomes permanently extinct. Enrichment Analysis shows that the most significant genes are those whose downregulation produces arrest of tumour cell multiplication through DNA blockage, cell cycle retardation, and mitotic activity diminution. Among these genes, downregulation of TOP2A gene was found pivotal for melanoma regression, and this gene is highly upregulated in melanoma tissues in clinical patients. We also elucidated two classes of drugs (podophyllin derivative and anthracycline derivative) that blocks the TOP2A receptors, and could be possible therapeutic agents in melanoma patients, thereby having the potential to duplicate the process of tumour regression in the clinical context.