Modeling and Simulation studies of DNA repair genes MLH1, MSH2, MSH6, PMS1, PMS2 with reference to HE4 gene

Tanvi Katoch (131516)

Parveen Kaur Thakur (133805)

Under the supervision of

Dr. Tiratha Raj Singh

May – 2017

Submitted in partial fulfillment of the Degree of

Bachelor of Technology in Bioinformatics

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT

CERTIFICATE

This is to certify that project report entitled **"Modeling and Simulation studies of DNA repair genes MLH1, MSH2, MSH6, PMS1, PMS2 with reference to HE4 gene"**, submitted by **Tanvi Katoch** and **Parveen Kaur Thakur**, in partial fulfillment for the award of degree of Bachelor of Technology in Bioinformatics Engineering to Jaypee University of Information Technology, Waknaghat and Solan has been carried out under my supervision.

This work has not been submitted partially to any other university or institute for the award of this or any other degree or diploma.

Date:

Supervisor's Name: Dr. Tiratha Raj Singh

Designation:

Assistant Professor (Senior Grade)

Jaypee University of Information Technology

Solan

DECLARATION

I hereby declare that the work presented in this report entitled "Modeling and Simulation studies of DNA repair genes MLH1, MSH2, MSH6, PMS1, PMS2 with reference to HE4 gene" in partial fulfillment of the requirements for the award of degree of Bachelor of Technology in Bioinformatics submitted in the Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan – 173234, Himachal Pradesh is an authentic record of my own work carried out over a period of August 2016 to May 2017, under the supervision of Dr. Tiratha Raj Singh, Assistant Professor (Senior Grade), Department of Biotechnology and Bioinformatics.

The matter embodied in the report is not been submitted for the award of any other degree or diploma.

Tanvi Katoch (131516)

Parveen Kaur Thakur (133805)

ACKNOWLEDGEMENT

The completion of this undertaking could not have been possible without the guidance of our mentor Dr. Tiratha Raj Singh, Assistance Professor (Senior Grade) in Dept. of Biotechnology and Bioinformatics, who initiated as into the realm of research and supervised as with finite patience and without whose invaluable suggestions and unstinted cooperation, the present desertion would not have been possible. Also I would like to thank the officials of Jaypee University of Information Technology, Waknaghat for their help and cooperation.

ABSTRACT

The two most important components of our project are modeling and simulation of the genes involved in the endometrial and ovarian cancer. The bottom up approach was used starting from the data collection, followed by expanding the data by retrieving their interacting partners, thereby following the pathway targeting approach and using the gathered data to build the model, run the simulations so as to state the hypothesis which could be further validated in wet laboratory.

The curves obtained after running the simulations provide the insight about the attaining of the steady state under the constraints of time and concentration for all the genes under study. This hypothesis depicts the behavior of the genes in unit time and unit concentration and how the behavior deviates from the normal behavior when the expression values and reaction rates are changed in order to achieve the steady state.

TABLE OF FIGURES

TABLE OF CONTENTS

CHAPTER 1

Introduction

Carrying out experiments is part of day to day activity of modern biologists. But the challenge crops up in post-genomic era, wherein there is lots of omics data available and whose acquisition is not a problem but the real problem lies in processing of data, analyzing the data, generating useful knowledge and getting some useful insight from the research point of view, therefore one needs to move away from one gene to protein approach to start considering the system as a whole i.e. holistic approach by following the principles of System biology [1]. Mathematical modeling is an approach to unwind the complexity of the biological processes. Several attempts have been made in the past for simulating and modeling complex biological processes like cell signal transduction pathways, networks gene regulation and pathways regulating the metabolism. The generated pathway models do not only generate experimentally hypothesis which can be verified but it also provides significant insights depicting the behavior of complex biological systems. Latest studies state that the phenotypic variations inherited in the organisms usually operate at a level of gene expression [2].

Therefore, it is important to develop novel methods to mathematically represent and simulate biological systems. The main idea is to find a representation which is biologically significant for understanding the biological system under the observation. Due to the complexity of the pathway due to high number of interactions which involve many components, it is almost impossible to understand how actually cellular networks behave? So in order to understand the topological and dynamical insight of the networks, mathematical simulation and modeling studies are proved to be useful [3]. It is worth stressing that Cancer modeling is growing rapidly as one of the challenges involved among mathematicians who work in collaboration with the researchers who are working on biological related fields. The motivating factor being shifting of the cancer from position 7th to position 2nd in the list of the fatal diseases, being preceded by the heart diseases. In fact, according to the WHO estimation, presently cancer kills around 6 million people in a year, Endometrial Cancer is 6th most common cancer in women. According to American Cancer Society's (ACS), Ovarian Cancer is the seventh most common cancer in women [4].

Modeling & Simulation

System

A system is the aggregation of various components which interact among themselves and usually are bounded under the constraints of time and space.

Model

A model is a usually a simple way of representing a system at some point under the constraints of time and space usually to understand the real system.

Modeling

Biological system refers to group of organs that work as a single unit to perform a significant task. For Modeling the biological systems usually involve good knowledge of biological sciences, mathematics and information technology. After Modeling, simulating of biological systems is carried out, as in case of networks and enzymes in metabolic processes, gene regulation processes and signaling pathways, so as to analyze and build a hypothesis.

The aim is to generate significant models which operate on basis of a response generated by a system to stimuli generated by environmental and internal factors, like a model generated for a cancer cell so as to find changes in its signaling pathways, ion channel mutations to see effects on cardiomyocytes.

Simulation

Simulations refers to the way in which the model operates under the constraints of time and space in order to analyze the models, thus enabling perseverance of the interactions between the individual components of the system that would be difficult to understand because of limitation of time and space.

Study Schematic of Simulations

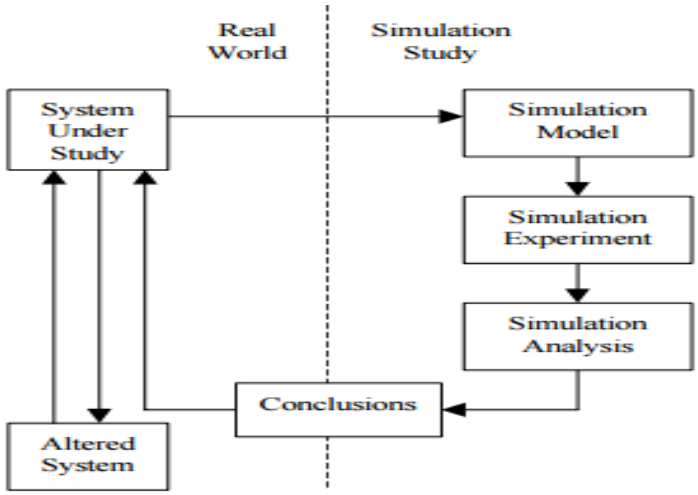


Fig 1: Schematic representation of simulation study [2]

Endometrial Cancer

Many common diseases of gynecological origin are associated with the abnormal proliferation in the endometrial layer of the uterus. These diseases include the endometrial cancer, adenomyosis and endometriosis. The Endometrial Cancer is the most common female origin malignancy that starts in the endometrial layer of the uterus. The Endometrial Cancer can originate from abnormal hormone estrogen level or improper bleeding, chemical side effects, radiation factors, viral particles inducing infections, inheritance of mutations down the generations. Tumors in endometrial layer of uterus include those developed in epithelial layer namely carcinomas and those which occur in stromal tissue are sarcomas. The former ones occur in more number as compared to latter ones [5].

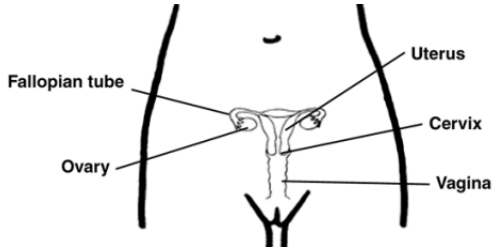


Fig 2: Female Reproductive System [3]

The uterus is called the womb and is usually organ which is hollow. It is the place of development of fetus and also it nourishes when the woman is pregnant. The main parts of uterus are as follows:

- The cervix is the end of the uterus i.e. the mouth of uterus that opens into the vagina.
- The body is the upper part of the uterus.

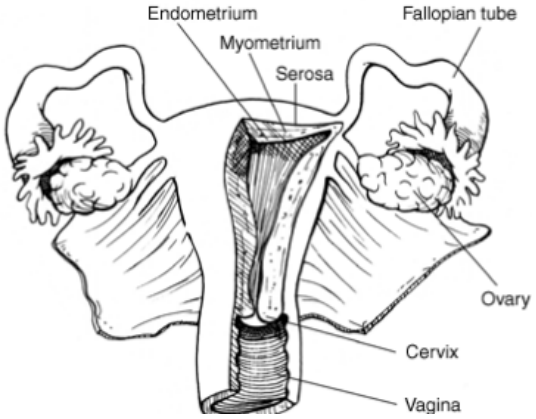


Fig 3: Layers of uterus [6]

The body of the uterus comprise of two significant layers i.e. endometrium and myometrium. Wherein former is the inner layer of the uterus and latter is the middle layer of the uterus. Nearly all cancers of the uterus occur in endometrial layer of uterus.

Several factors governing the chance of development of endometrial cancer include:

- Use of Birth control pills, and chances of developing polycystic ovarian syndrome
- Implantation of intrauterine devices
- Person's age
- Whether a person is diabetic or has chances of developing it through inheritance
- Genetic predispositions
- Chances of developing breast cancer or ovarian cancer
- If endometrial hyperplasia is diagnosed in the past
- Exposed to harmful radiations

Some of these factors are linked to a lower risk of developing endometrial cancer like pregnancy, birth control pills, and implantation intrauterine device.

Whereas many of them are linked to a higher risk of developing endometrial cancer as in case of genetic predispositions, previous diagnosis or exposure to harmful radiations.

Ovarian Cancer

Ovarian cancer refers to tumor formation in an ovary. It results in invasion abnormal cell growth which invades or spread to other parts of the body. Women who ovulate at abnormal level are at more risk of developing ovarian cancer. These women include those who are infertile, who start ovulating at a younger age or reach menopause at age later than it should be. Ovarian cancer remains asymptotical until it has spread to the pelvic and abdominal area. So when actually detected it is usually late, thus difficult to treat and is usually fatal. At initial stage the disease limits to the ovary and can be treated successfully. Surgery and chemotherapy are two common methods used to treat ovarian cancer. Ovarian cancer does not show any symptoms at the early stage. But at the later stage some few and nonspecific symptoms may appear that are often thought to be common benign conditions [6].

Ovarian cancer symptoms include: -

- Swelling in the abdomen
- Sudden loss in weight
- Pelvic area uneasiness
- Constipation
- Urinating frequently

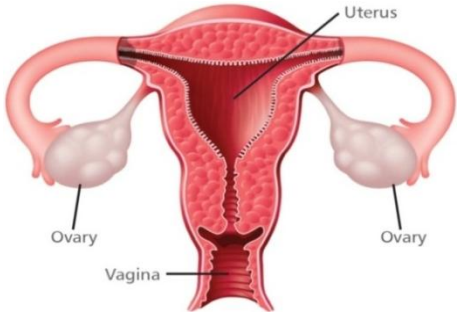


Fig 4: Ovaries showing tumour formation [6]

Types of ovarian cancer

The cell where the cancer begins usually is associated to the type of ovarian cancer developed in the body [7]. Ovarian cancer types include:

- Epithelial tumors: Tumors which occur and develop in the thin tissue layer covering the ovaries externally (~90%).
- Stromal tumors: Tumors which occur and develop in tissue that contains hormone-producing cells. These tumors are usually diagnosed at an earlier stage (~7%).
- Germ cell tumors: Tumors which occur and develop in tissue in the egg-producing cells. These are rare ovarian cancers usually occur in younger women.

Stages of ovarian cancer include:

- Stage I: When the cancer occurs only in one or both ovaries.
- Stage II: Cancer has invaded the other parts of the pelvic region.
- Stage III: Cancer has invaded the abdomen.
- Stage IV: Cancer is found outside the abdominal region.

Cancer risk and Polycystic ovary syndrome (PCOS):

Women suffering from polycystic ovary syndrome have more chances of developing the endometrial cancer. The main factor for this increased malignancy risk is regular exposure of the endometrium to estrogen that results due to an ovulation. Some women having PCOS undergo ovulation induction or receive progesterone resistance which is accompanied by anti regulation of gene expression which controls cell proliferation. Endometrial surveillance comprise of transvaginal ultrasound and endometrial biopsy so as to analyze the thickened endometrial layer, prolonged vaginal bleeding [8]. Remedies for abnormal vaginal bleeding i.e. endometrial hyperplasia consist of estrogen and progestin containing oral contraceptives etc. In order to treat obesity which is risk factor for developing endometrial disease, modifying the lifestyle with less calorie intake and exercises is a necessity [9].

Women having PCOS have increased chances of developing the ovarian cancer. Studies reveal that oral contraceptive use is effective against ovarian cancer. There is no mentioned association of PCOS and breast cancer, although the high prevalence of metabolic dysfunction from obesity is a common promoting factor for both conditions. Polycystic ovarian syndrome is known to increase ones chances of developing endometrial cancer. Irregular or no periods, causes the endometrium to pile up and get thickened. This thickening leads to development of endometrial cancer. Women with PCOS are about three times more likely to develop EC compared with women without it. More studies are required and it is yet to state about the underlying mechanisms in cancer progression and determination of the idol way of analyzing and preventing disease progression [10].

DNA Repair Pathways:

The collection of processes by which a cell recognizes, repairs and corrects damage to the DNA molecules that encode its genome.

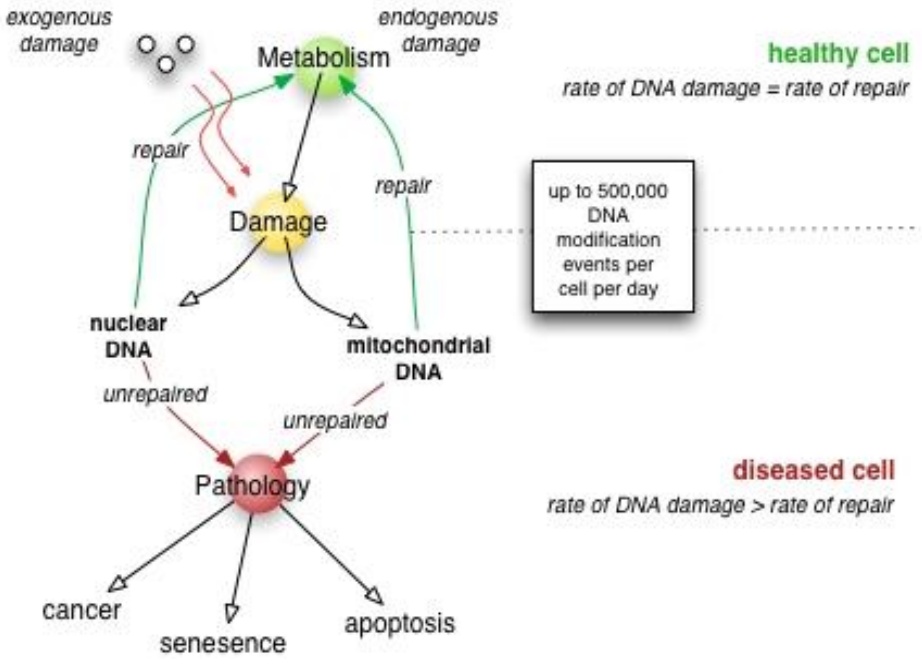


Fig 5: DNA Repair Mechanism [11]

At a minimum, most would agree that mammalian cells utilize five major DNA repair mechanisms: mismatch repair (MMR), base excision repair (BER), double-strand break repair, which includes both homologous recombination (HR) and non-homologous end joining (NHEJ) and nucleotide excision repair (NER).

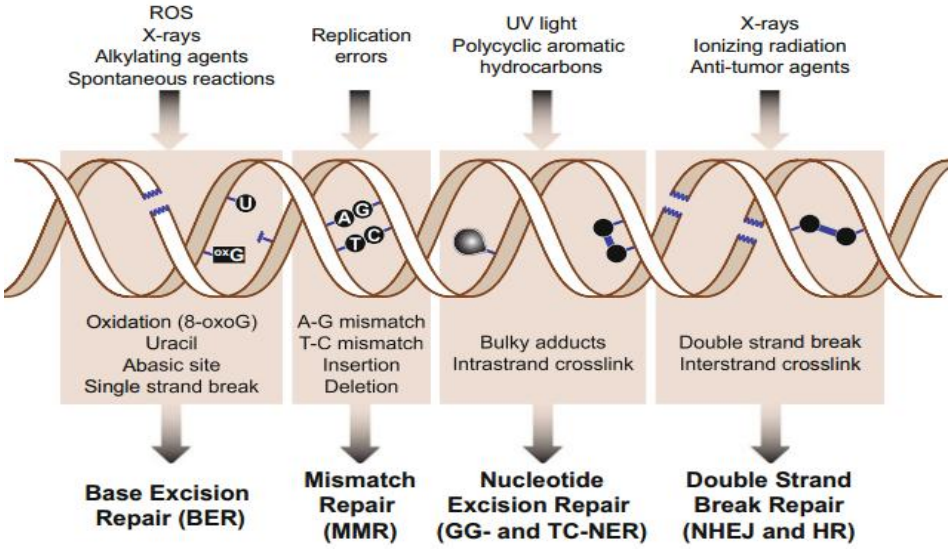


Fig 6: DNA Repair Pathways [11]

Genetics

The major risk factor for ovarian cancer is the mutations in BRCA2 or BRCA1 DNA MMR genes, which is presently 10% in the ovarian cancer cases. Family history colon cancer endometrial cancer or other gastrointestinal tumors may indicate the presence of Lynch syndrome also known as hereditary nonpolyposis colorectal cancer, which grants a higher risk for developing different cancers like ovarian cancer [12]. Mutations in mismatch repair gene including MLH1, MSH2, PMS1, PMS2, and MSH6 causes Lynch syndrome. Selective over expression of Human epididymal secretory protein E4 (HE4) indicates its role in ovarian cancer tumor genesis but very less is known about the HE4 gene role. Targeting selectively towards the HE4 protein indicate therapeutic benefits for the treatment of cancer. Studies show that HE4 can possibly serve as a major biomarker that can monitor and detect recurrence in women with ovarian cancer and endometrial cancer who are possibly are at high risk of recurrence, as the expression level data retrieved from the database as well as from the literature suggest that HE4 is over expressed in both ovarian and endometrial cancer [13].

CHAPTER 2

Material and Methods

Data collection

For collecting the data different databases including Human DNA Repair Genes, DNA Repair Database, DR-GAS and REPAIRtoire were searched along with literature review so as to perform the process of gene collection and pathway association.

- **Literature review**

Various research papers were consulted and a list of genes involved in the ovarian and endometrial cancer was made. This was a crucial step as we were following the bottom up approach and if our initial data would be incorrect our entire ongoing future project work would be affected (Refer Appendix I).

- **Gene Catalogue related to DNA Repair:**

Various DNA repair databases as listed below were searched and data was retrieved and analyzed so as move forward to the process of gene selection, i.e. selection of those genes which show a pivotal role in ovarian and endometrial cancer, ovarian cancer, and endometrial cancer respectively.

S.No	Databases	Number of Genes	Year of Publication
1	REPAIRtoire	304	2010
2	Human DNA Repair Genes	181	2007
3	DNA Repair Database	151	2014
4	DR-GAS	216	2014

- **Gene Selection and pathway association:**

After the analysis of the literature review and data retrieved from the databases the following list of genes were selected for further analysis based on their higher occurrence frequency and their role in endometrial and ovarian cancer [15]. Thereby these genes were mapped in accordance to the pathway they were involved in so as to move a step ahead in terms of analysis by targeting a particular pathway, followed by model building and simulations.

S.No	Gene Name	Pathway
1	MLH1	Mis-Match Repair
2	PMS1	Mis-Match Repair
3	MSH2	Mis-Match Repair
4	MSH6	Mis-Match Repair
5	PMS2	Mis-Match Repair
6	HE4	No pathway information

- **Model building**

- Model for the MMR pathway is built using CellDesigner 4.4 as the genes namely hMLH1, hMSH2, hMSH6, PMS1 and PMS2 are the mismatch repair genes which are directly involved in endometrial cancer.
- Based on the above collected data the model of MMR pathway was made using the cell designer and simulations were run but there was a need of model refinement and optimization of simulations, so that a hypothesis could be built which could be further tested in wet laboratories.

- **Simulation**

- Now simulations are carried out on the built model so as evaluate the behavior of generated model o hypothesis under the constraints of time and concentration. This is done so as to predict the behavior of the genes involved in Endometrial and Ovarian Cancer, and to see when they achieve equilibrium in terms of expression level under the constraints of time and concentration.

Software Used: CellDesigner (4.4)

CellDesigner is a structured diagram editor for drawing gene-regulatory and biochemical networks. It was developed by Kitano. Networks are drawn based on the pathways, with graphical notation system which are stored using the SBML (Systems Biology Markup Language), a standard for representing and building models of gene-regulatory networks and biochemical pathways. Models are able to link with analysis and other simulation packages through SBW. By using CellDesigner, you can browse and modify existing SBML models with references to existing databases, simulate and view the dynamics through an intuitive graphical interface [14].

Therefore the following steps were followed so as to achieve optimal simulation curves:

- **Literature review**

- Over-expression rates of MSH6, MSH2, PMS1 and MLH1 protein were 16.18%, 12.14%, 7.51% and 5.78% respectively.
- Variants in MLH3, MLH1, MSH2, PMS1, MSH6 and MSH3 contribute significantly to ovarian cancer susceptibility. The total loss rate of MMR protein was 29.89% (27/87).
- The observed association of PMS2 with ovarian cancer and endometrial cancer warrants confirmation in an independent study as the analysis is difficult due to presence of psuedogenes.

- **Ovarian Cancer data retrieval**

- Expression level data was collected from COSMIC database so as to see whether the gene involved in ovarian cancer is over expressed or under expressed and by what extent. This was done to point out the significance of each gene in ovarian cancer, which would definitely affect the model building and simulation process [16].

Gene Name	Expression Type	Expression Value (Over)	Expression Value (Under)
MLH1	Over, Under	2.675063	-2.113
PMS1	Over, Under	3.587018	-
PMS2	-	-	-
MSH2	Over	3.837351	-
MSH6	Over	4.119732	-
HE4	Over	3.285724	-

- **Endometrial Cancer data retrieval**

- Expression level data was collected from COSMIC database so as to see whether the gene involved in ovarian cancer is over expressed or under expressed and by what extent.
- This was done to point out the significance of each gene in ovarian cancer, which would definitely affect the model building and simulation process.

Genes	Expression Type	Over	Under	Expression Type	Over	Under
MLH1	Over, Under	2.65025	-2.11	Over	2.407714	-
PMS1	Over, Under	2.63633	-2.185	Over, Under	3.046785	-2.307
PMS2	-	-	-	-	-	-
MSH2	Over	3.48022	-	Over, Under	2.905782	-2.167
MSH6	Over	4.119732	-	Over, Under	2.800888	-2.0724
HE4	Over	3.285724	-	Over	3.563833	-

- **Retrieving interacting partners**

- Common interacting partners of the genes selected were collected from STRING database so as to analyze genes selected with other genes and expand our data so as to follow the bottom up approach and help reach a point where we could put forward a relevant hypothesis [17].

- **Following steps were taken:**
- Finding interacting partners of each selected genes

HE4	MLH	MSH2	PMS1	MSH6	PMS2
MSLN	BLM	BLM	FBXO18	BLM	BRIP1
CES5A	BRIP1	BRIP1	FAN1	BRIP1	EXO1
SPP1	EXO1	ERCC1	MSH3	EXO1	FAN1
LY6G5C	FAN1	EXO1	EXO1	MLH3	PCNA
EGFR	MLH3	MLH1	MSH4	MSH3	MLH3
IPO4	PCNA	MLH3	MSH5	MSH4	MSH3
CES2	MSH3	MSH3	TP53	MSH5	MSH4
MUC1	MSH4	MSH4	XRCC6	PCNA	MSH5
SLP1	MSH5	MSH5	-	XRCC6	-
CES1	-	PCNA	-	POLE	-
-	-	POLE	-	PRKDC	-
-	-	PRKDC	-	TP53	-

- Finding unique common interacting partners among all selected genes along with process involved.

S.No.	Gene	Process	Source
1.	BLM	DNA Replication	Uniprot
2.	BRIP1	DNA Replication	Uniprot
3.	EXO1	DNA Replication	Uniprot
4.	FAN1	DNA Replication	Uniprot
5.	MLH3	Mismatch Repair	Uniprot
6.	MSH3	Mismatch Repair	Uniprot
7.	MSH4	Mismatch Repair	Uniprot
8.	MSH5	Mismatch Repair	Uniprot
9.	PCNA	Mismatch Repair	Uniprot
10.	POLE	DNA Replication	Uniprot
11.	PRKDC	DNA Replication	Uniprot
12.	TP53	Cellular Tumor Antigen	Uniprot
13.	XRCC6	DNA Replication	Uniprot

- **Data for interacting partners in Ovarian Cancer**

- Expression level data was collected from COSMIC database for all unique common interacting partners, so as to see whether the gene involved in ovarian cancer is over expressed or under expressed and by what extent. This was done to point out the significance of each gene in ovarian cancer, which would definitely affect the model building and simulation process.

Genes	Expression Type	Over	Under
MSH3	Over, Under	2.645455	-1.653
POLE	Over	2.966059	-
EXO1	Over	3.081647	-
MLH3	Over, Under	3.113288	-2.22117
PCNA	Over	3.245296	-
BLM	Over	3.269947	-
XRCC6	Over, Under	3.5373	-2057534
MSH4	Over	3.608	-
TP53	Over, Under	3.730978	-2.67649
BRIP1	Over	6.22012	-
FAN1	CNV Data	-	-
MSH5	CNV Data	-	-
PRKDC	No Data	-	-

- **Data for interacting partners in Endometrial Cancer**

- Expression level data was collected from COSMIC database for all unique common interacting partners of the selected genes, so as to see whether the gene involved in ovarian cancer is over expressed or under expressed and by what extent.
- This was done to point out the significance of each gene in ovarian cancer, which would definitely affect the model building and simulation process.

Genes	Expression Type	Over	Under	Expression Type	Over	Under
BLM	Over	3.3304	-	Over	2.68865	-
BRIP1	Over	3.31778	-	Over	2.99018	-
EXO1	Over	3.21167	-	Over	3.15781	-
MLH3	Over, Under	4.30575	-3.0625	Over, Under	2.98736	-2.3832
MSH3	Over, Under	2.157	-2.032	Over, Under	2.82224	-2.1275
MSH4	Over	2.505	-	Over	3.90565	-
PCNA	Over	3.21421	-	Over	3.753	-
POLE	Over,	3.5132	-	Over	3.25958	-
TP53	Over	3.0.145	-2.2475	Over, Under	3.44993	-3.4493
XRCC6	Over, Under	2.204	-2.121	Over, Under	2.756965	-2.205
FAN1	CNV Data	-	-	-	-	-
MSH5	CNV Data	-	-	-	-	-
PRKDC	No Data	-	-	-	-	-

- **GO data**

- The Gene Ontology (GO) data was retrieved so as to point out the relevance of genes in terms of their function, component, process involved. Thus, bringing out the significant interrelations among the selected genes [18].

- **GO data : MLH1:**

GO Id	Go Name	Aspect	Reference	Source
GO:0005524	ATP Binding	Function	GO_REF:0000002	InterPro
GO:0005654	Nucleoplasm	Component	GO_REF:0000052	HPA
GO:0006298	Mismatch Repair	Process	GO_REF:0000002	InterPRo
GO:0030983	Mismatch DNA Binding	Function	GO_REF:0000002	InterPro

- **GO data : MSH2**

Go Name	Aspect	Reference	Source
DNA Binding	Function	GO_REF:0000002	InterPro
ATP Binding	Function	GO_REF:0000002	HPA
Mismatch Repair	Process	GO_REF:0000002	InterPRo
Mismatch DNA Binding	Function	GO_REF:0000002	InterPro
Mismatch Repair Complex	Component	GO_REF:0000002	InterPro

- **GO data : MSH6**

GO Id	Go Name	Aspect	Source
GO:0005524	ATP Binding	Function	InterPro
GO:0005654	Nucleoplasm	Component	HPA
GO:0005829	Cytosol	Component	HPA
GO:0005794	Golgi Apparatus	Component	HPA
GO:0006298	Mismatch Repair	Process	InterPRo
GO:0030983	Mismatch DNA Binding	Function	InterPro

- **GO data : PMS2**

GO Id	Go Name	Aspect	Source
GO:0000932	P-body	Component	HPA
GO:0005524	ATP Binding	Function	InterPro
GO:0005634	Nucleus	Component	HPA
GO:0005654	Cytosol	Component	HPA
GO:0005886	Plasma Membrane	Component	HPA
GO:0006298	Mismatch Repair Complex	Process	Interpro
GO:0015630	Microtubule Cytoskeleton	Component	HPA
GO:0030983	Mismatched DNA Binding	Function	InterPro

- **GO data : PMS1**

GO Id	Go Name	Aspect	Source
GO:0003677	DNA Binding	Function	Uniprot
GO:0005524	ATP Binding	Function	Uniprot
GO:0005634	Nucleus	Component	Uniprot
GO:0006298	Mismatch Repair	Process	Interpro
GO:0030983	Mismatched DNA Binding	Function	InterPro

- **GO data : HE4**

GO Id	Go Name	Aspect	Source
GO:0004866	Endopeptidase Indibitor Activity(EIA)	Function	PINC
GO:0004867	Serine-type EIA	Function	Uniprot
GO:0004869	Cysteine-type EIA	Fucntion	Uniprot
GO:0005576	Extracellular Region	Component	InterPro
GO:0006508	Proteolysis	Process	PINC
GO:0007283	Spermatogenesis	Process	PINC

- **Model refinement**

- Model for the MMR pathway was refined using CellDesigner 4.4, based on the data collected and analysis performed, as the genes namely hMLH1, hMSH2, hMSH6, PMS1 and PMS2 are the mismatch repair genes which are directly involved in endometrial cancer.

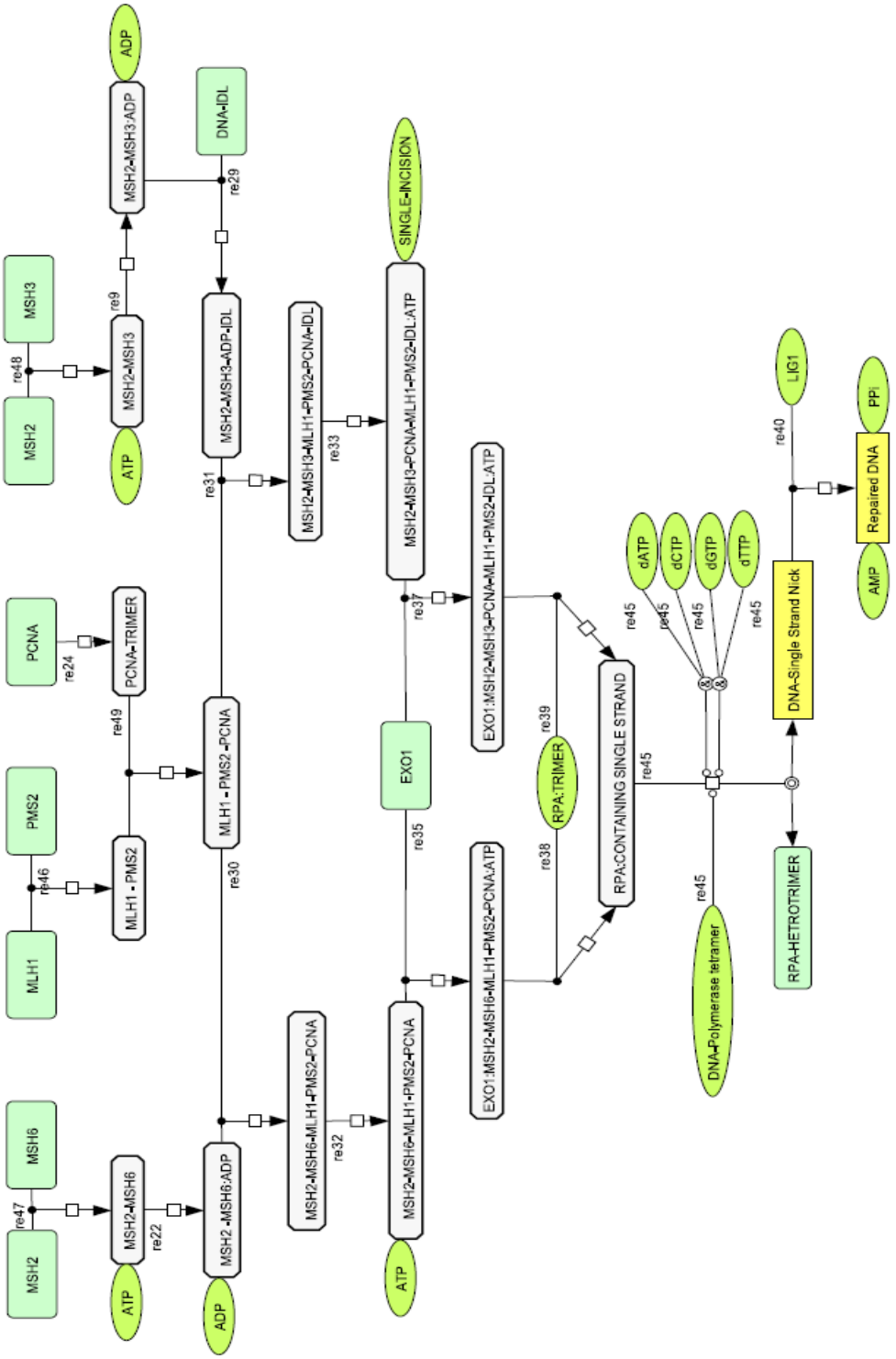
- **Optimization of simulation curves**

- Now simulations are carried out on the refined model so as evaluate the behavior of generated model hypothesis under the constraints of time and concentration. These simulations were optimized using expression value data, reaction rate data. As HE4 does not have any pathway data available as even KO number is not assigned to it.
- Therefore we could only point out its role in endometrial and ovarian cancer according to the expression level data available in the literature and the databases. Studies show that HE4 (Human Epididymis Protein 4) can possibly serve as a major biomarker that can monitor and detect recurrence in women with ovarian cancer and endometrial cancer who are possibly are at high risk of recurrence, as the expression level data retrieved from the database as well as from the literature suggest that HE4 is over expressed in both ovarian and endometrial cancer.

CHAPTER 3

Results

Pathway Modeling: Fig7: Pathway Model



Simulation Curves

The graphs given below represent the behavior of the genes namely MSH2, MSH6, PMS1, PMS2, MLH1 in endometrial and ovarian cancer under the constraints of time and concentration. We perform our simulation curves on both Ovarian Cancer and Endometrial Cancer.

- **For Ovarian Cancer**

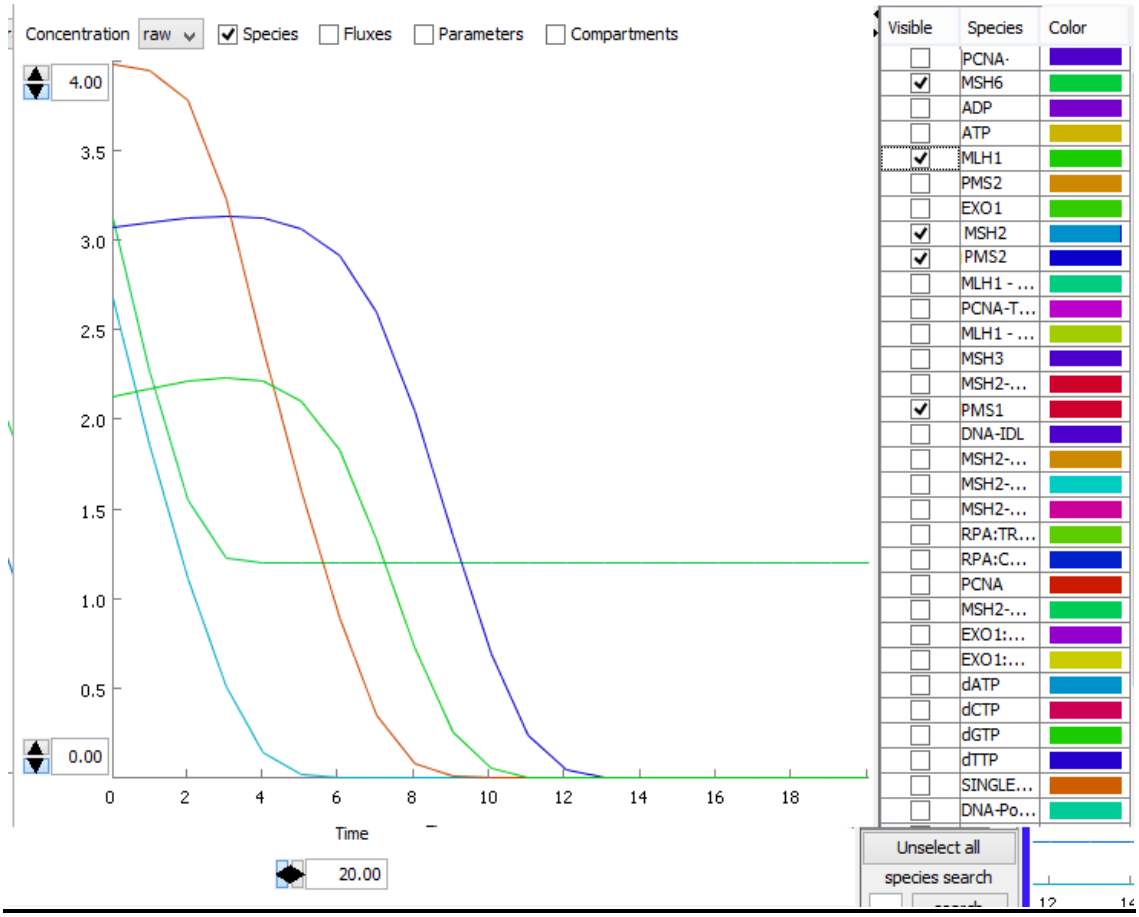


Fig8: Graph representing behavior of five genes involved in Ovarian Cancer

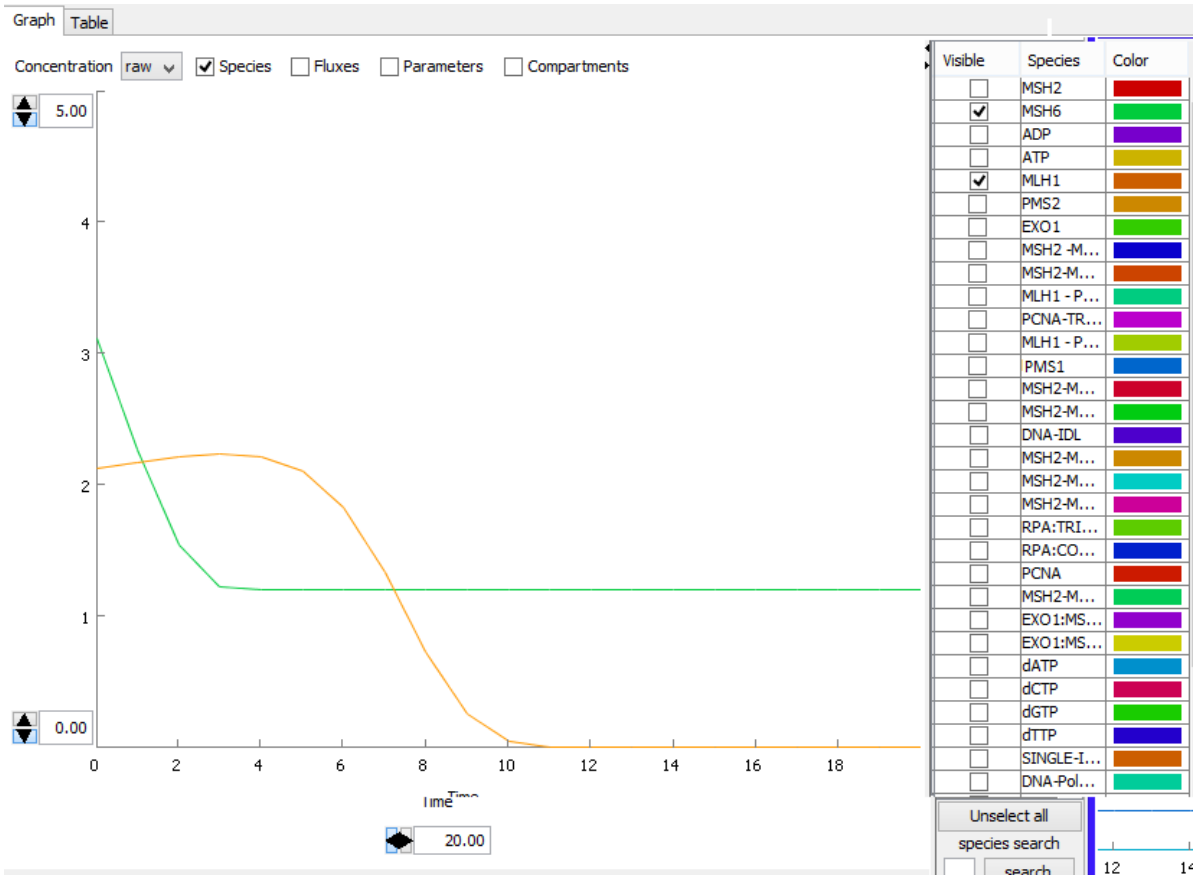


Fig9: Graph representing behavior of MSH6 and MLH1

- The above graph represents that MLH1 has the concentration of ~ 2.1 mM and time 0 ms, slightly increases to 2.2 mM up to 5.8 ms and then gradually decreases till 10 ms and achieves equilibrium at around 11 ms.
- In case of MSH6, the concentration is ~ 3.1 mM, then decreases till 3.5 ms and thereafter achieves equilibrium at ~ 1.2 mM.

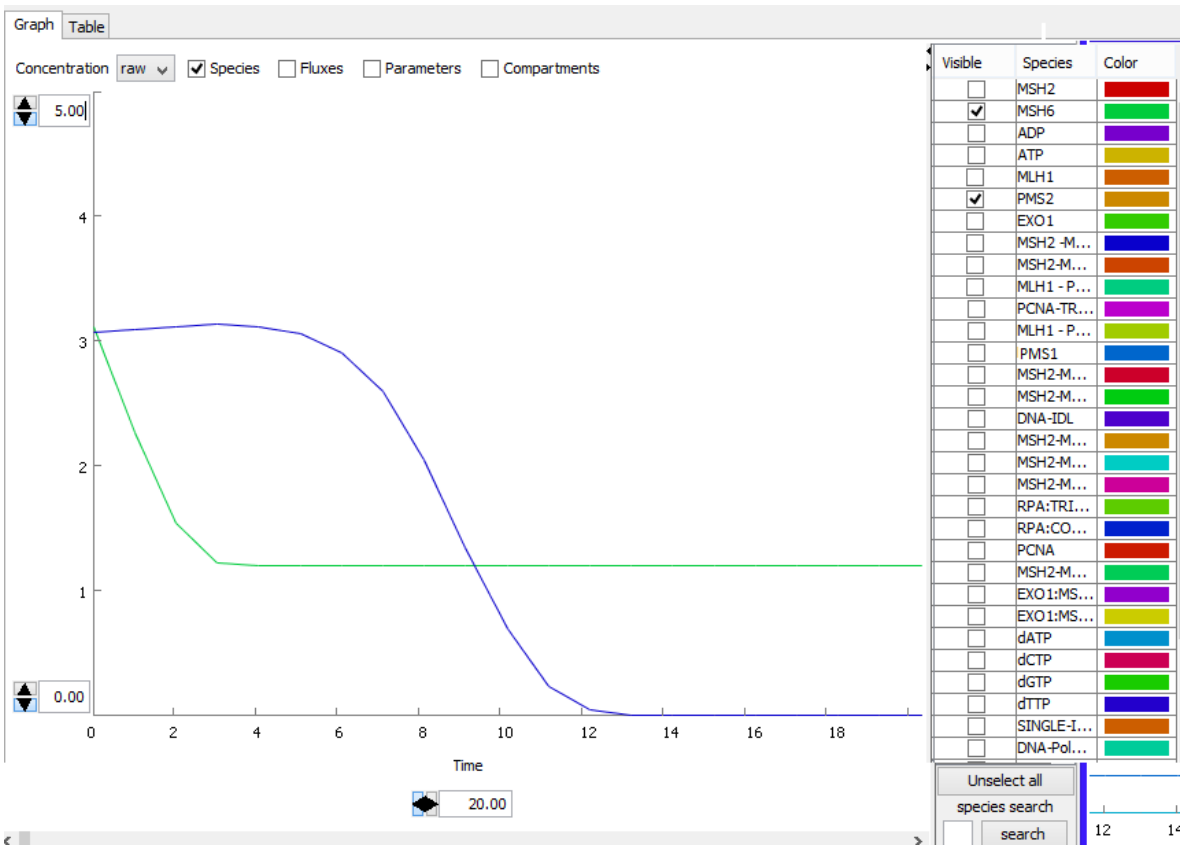


Fig10: Graph representing behavior of MSH6 and PMS2

- In case of MSH6, the concentration is ~ 3.1 mM, then decreases till 3.5 ms and thereafter achieves equilibrium at ~ 1.2 mM.
- In case of PMS2, the concentration is ~ 3.15 mM, and then it slightly increases till 6 ms, after that it gradually decreases till 12.5 ms and thereafter achieves equilibrium.

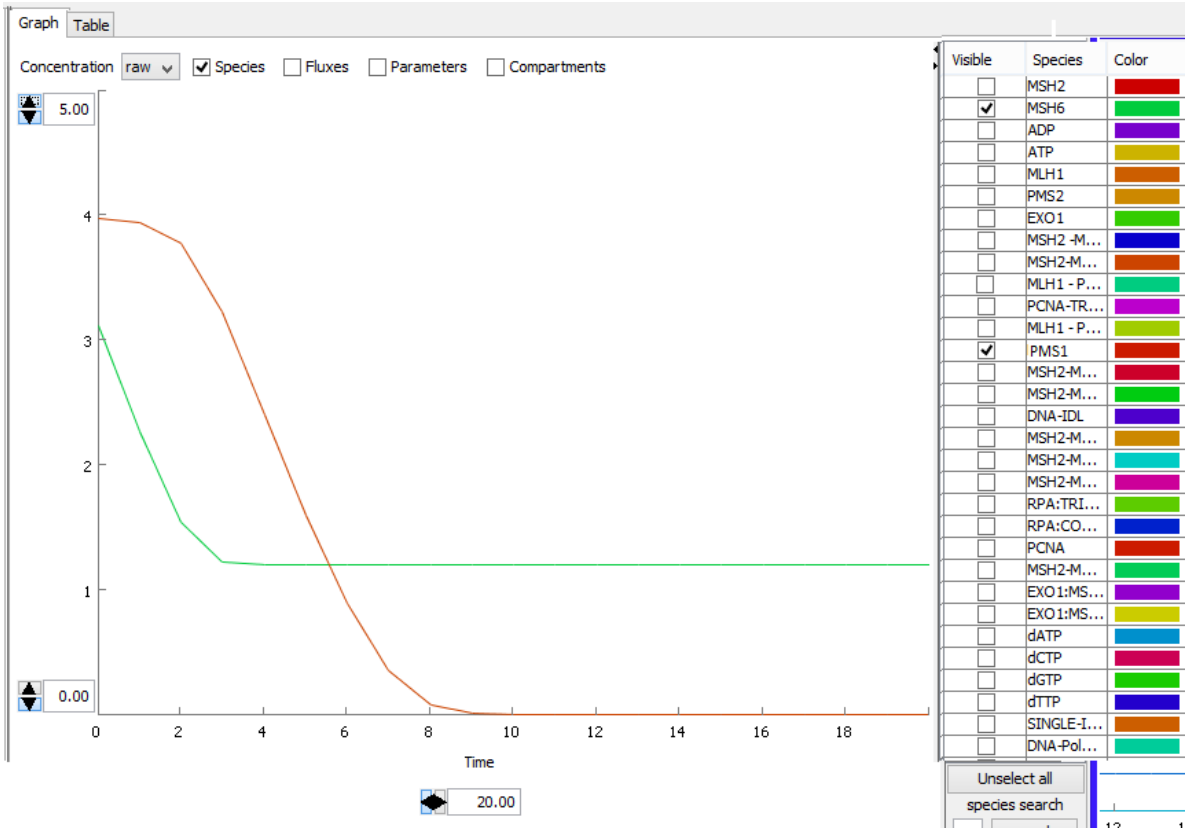


Fig11: Graph representing behavior of MSH6 and PMS1

- In case of MSH6, the concentration is ~ 3.1 mM, then decreases till 3.5 ms and thereafter achieves equilibrium at ~ 1.2 mM.
- In case of PMS1, the concentration is ~ 3.99 mM, then it gradually decreases till 9.0 ms, then thereafter achieves equilibrium at 9 ms.

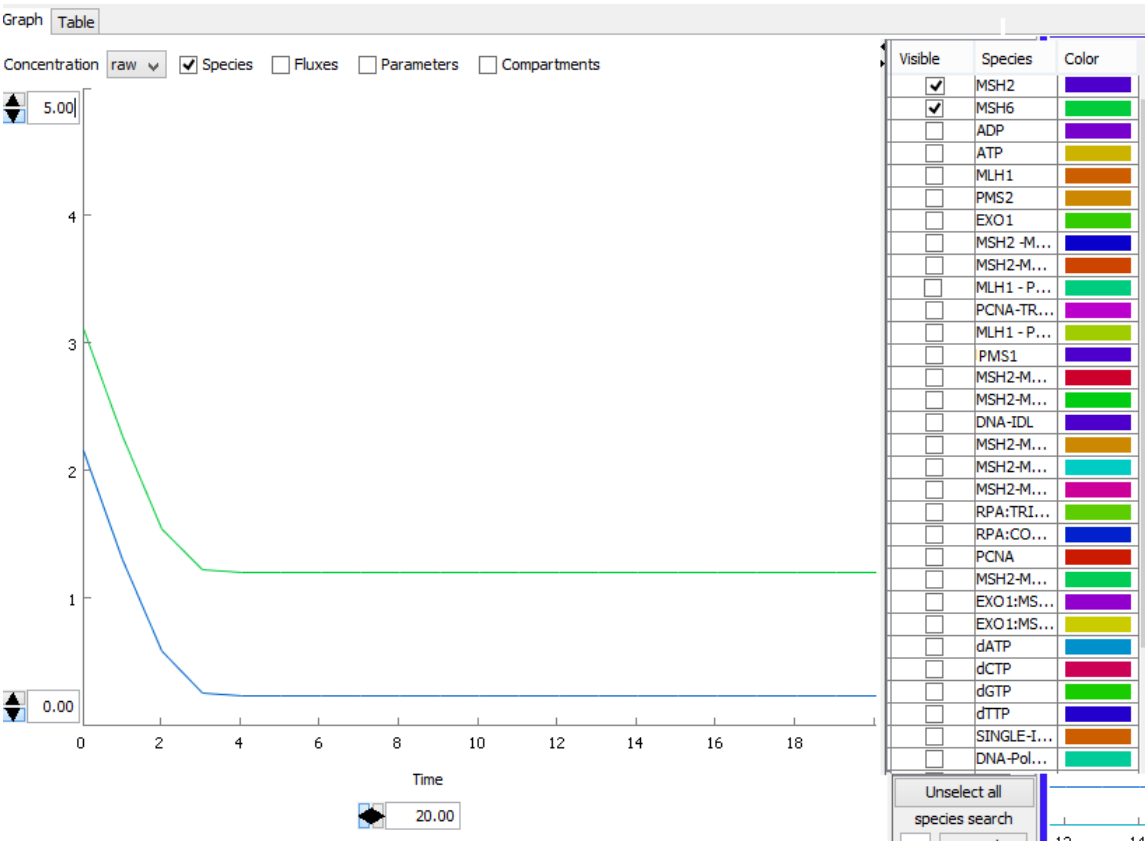


Fig12: Graph representing behavior of MSH6 and MSH2

- In case of MSH6, the concentration is ~ 3.1 mM, then decreases till 3.5 ms and thereafter achieves equilibrium at ~ 1.2 mM.
- In case of MSH2, the concentration is ~ 2.1 mM, and then it decreases till 2.2 ms to 3.5 ms up to ~ 0.25 mM, and then thereafter achieves equilibrium.

- **For Endometrial Cancer**

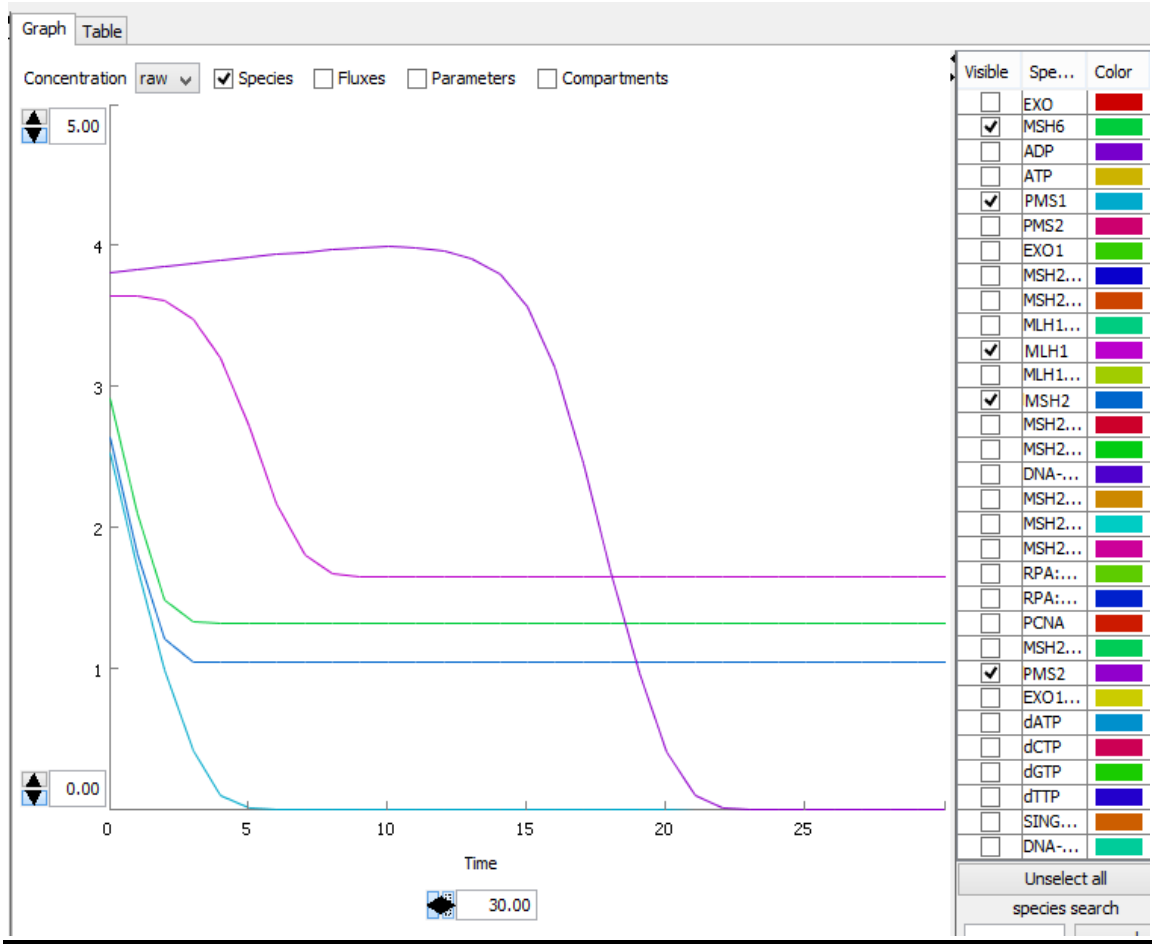


Fig13: Graph representing behavior of five genes involved in Endometrial Cancer

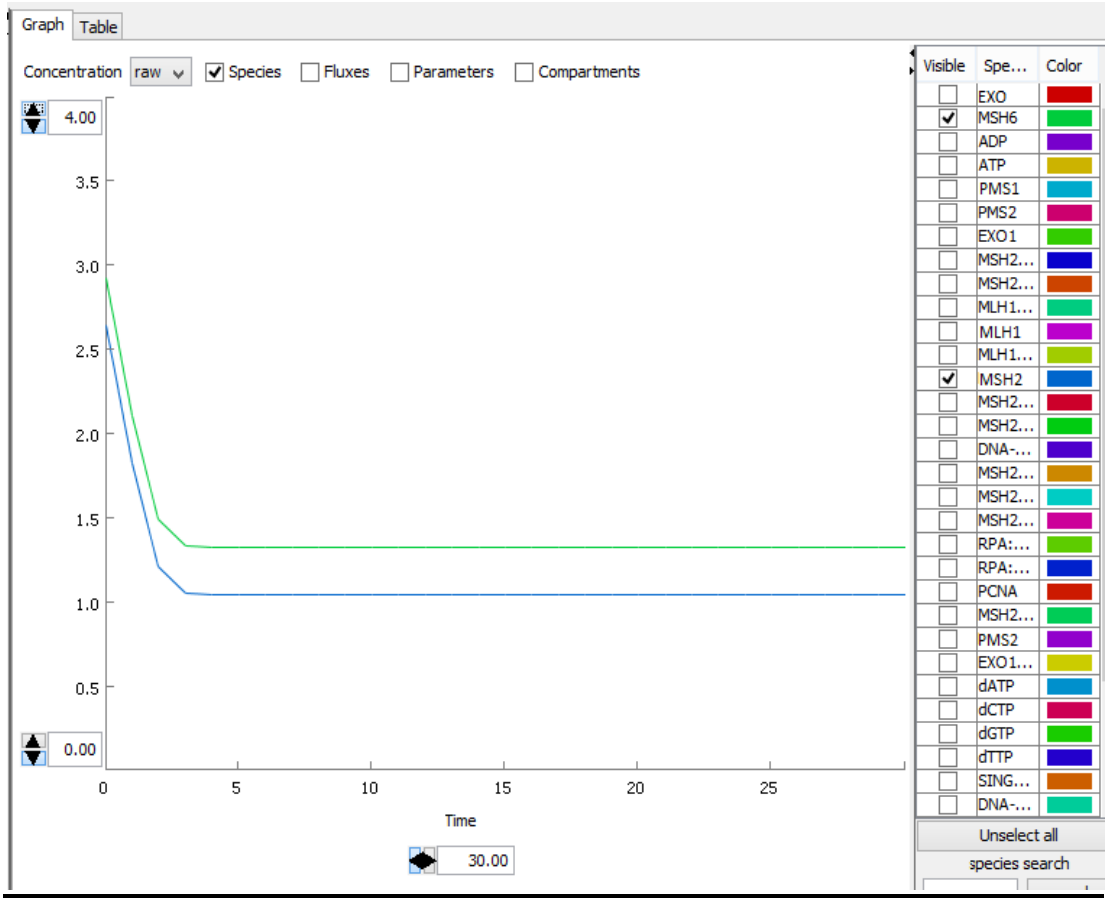


Fig14: Graph representing behavior of MSH6 and MSH2

- In case of MSH6, the concentration is ~ 2.9 mM at time 0 ms, then decreases till 2.8 ms to 3.5 ms up to ~ 1.3 mM and thereafter achieves equilibrium.
- In case of MSH2, the concentration is ~ 2.6 mM at time 0 ms, and then it decreases till 2.5 ms to 3.2 ms up to ~ 1mM, and then thereafter achieves equilibrium.

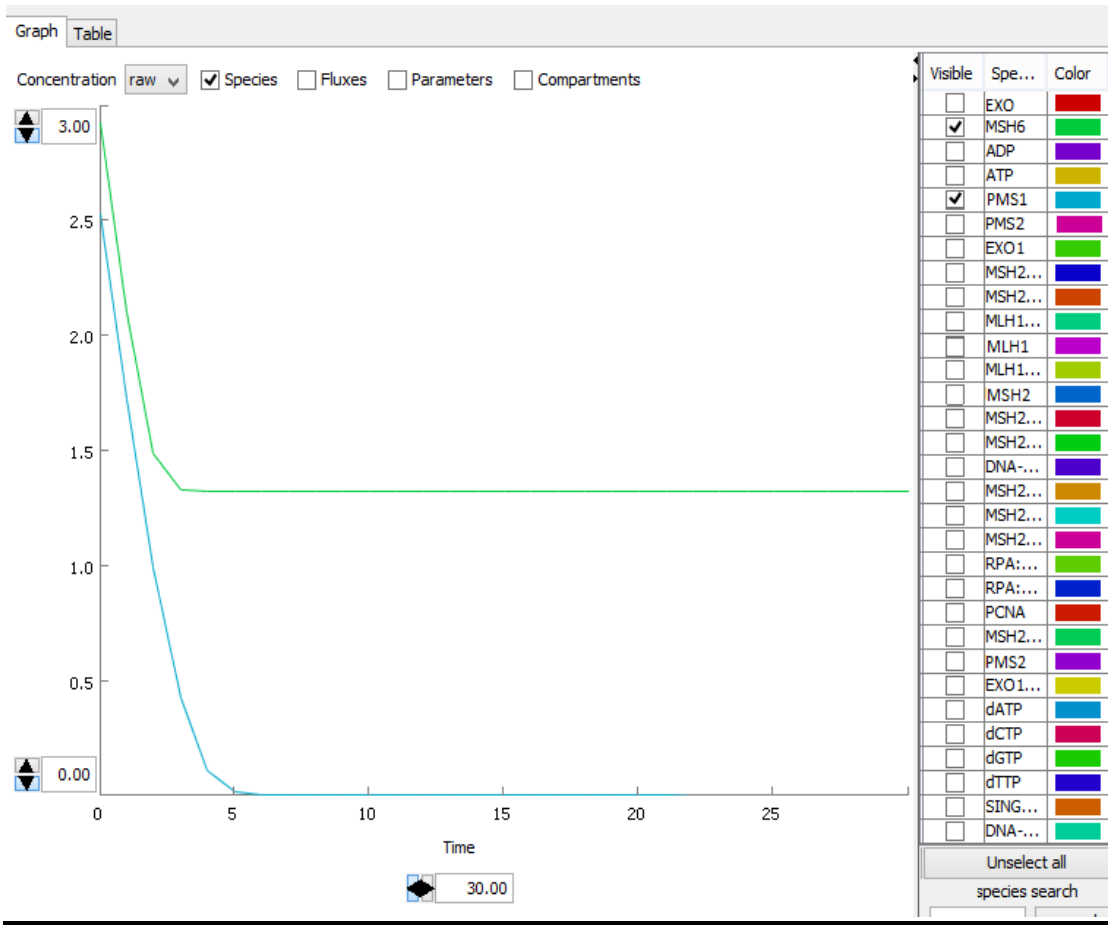


Fig15: Graph representing behavior of MSH6 and PMS1

- In case of MSH6, the concentration is ~ 2.9 mM at time 0 ms, then decreases till 2.8 ms to 3.5 ms upto ~ 1.3 mM and thereafter achieve equilibrium.
- In case of PMS1, the concentration is ~ 2.5 mM at time 0 ms, and then it decreases till 5 ms, and then thereafter achieves equilibrium.

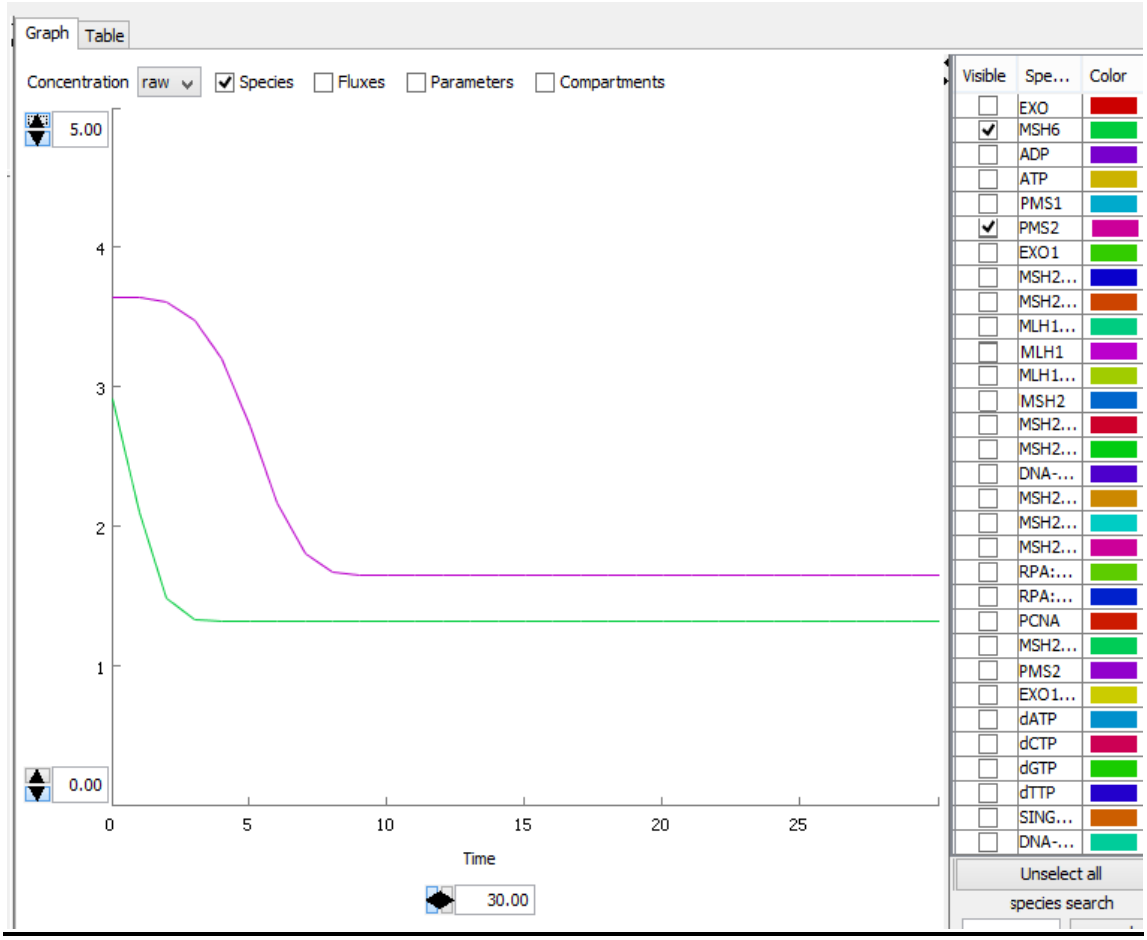


Fig16: Graph representing behavior of MSH6 and PMS2

- In case of MSH6, the concentration is ~ 2.9 mM at time 0 ms, then decreases till 2.8 ms to 3.5 ms up to ~ 1.3 mM and thereafter achieves equilibrium.
- In case of PMS2, the concentration is ~ 3.8 mM at time 0 ms, and then it gradually decreases till 8 ms up to ~ 1.7 mM, and then thereafter achieves equilibrium.

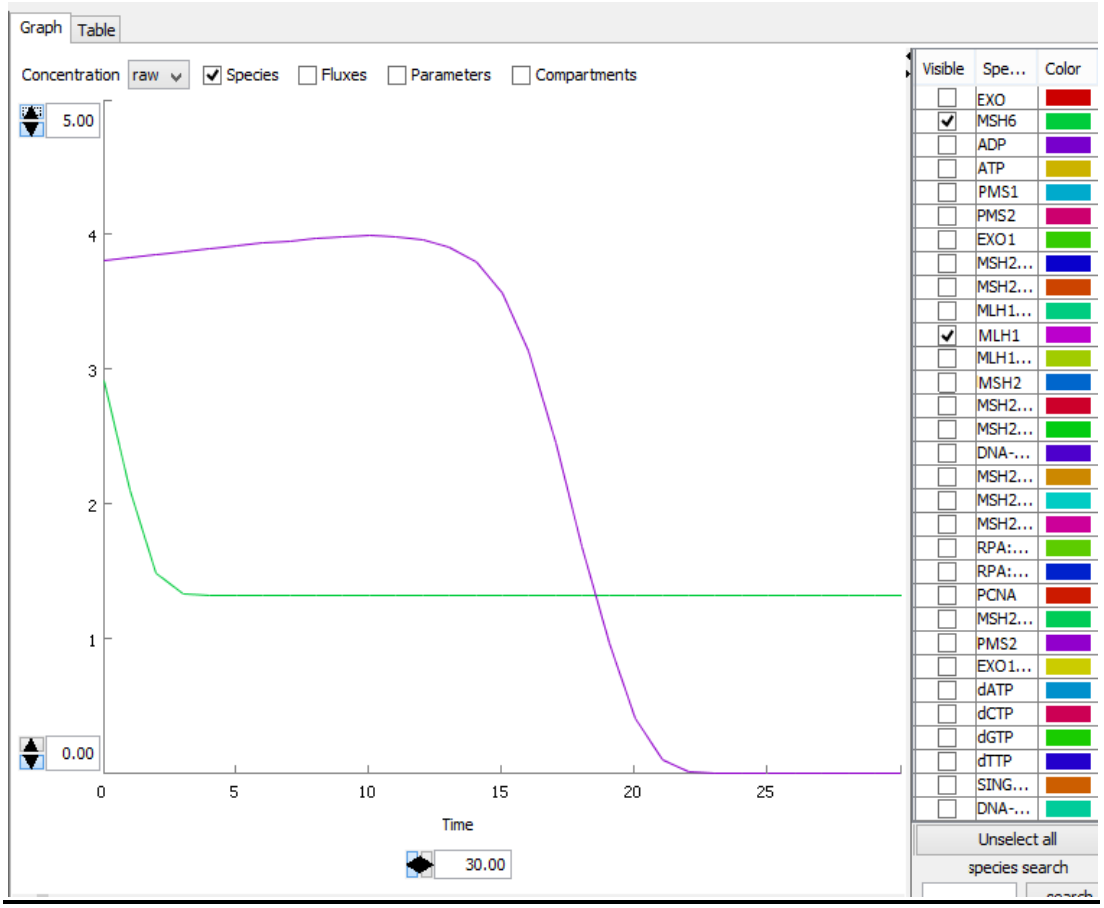


Fig17: Graph representing behavior of MSH6 and MLH1

- In case of MSH6, the concentration is ~ 2.9 mM at time 0 ms, then decreases till 2.8 ms to 3.5 ms up to ~ 1.3 mM and thereafter achieves equilibrium.
- In case of MLH1, the concentration is ~ 3.85 mM at time 0 ms, and then it slightly increases to 4.0 mM up to 16 ms, and then it gradually decreases till 22 ms, and thereafter achieves equilibrium.

CHAPTER 4

Discussion

- According to the literature review, over-expression rates of MSH6, MSH2, PMS1 and MLH1 protein were 16.18% ,12.14%, 7.51% and 5.78% respectively which is also in accordance with the data retrieved from cosmic database.
- The total loss rate of MMR protein was 29.89% as mentioned in the literature.
- Mutations in MLH3, MLH1, MSH2, PMS1, MSH6 and MSH3 accord significantly to endometrial cancer and ovarian cancer.
- The study of association of PMS2 with endometrial cancer and ovarian cancer should be done independently, as the analysis is difficult due to presence of psuedogenes.
- There were 10, 12, 17, 10, 14, 11 interacting partners retrieved from STRING database for He4, MLH1, MSH2, PMS1, MSH6, PMS2 respectively. Out of these 13 unique common interacting partners were found which were further analyzed.
- HE4 does not have any pathway data available as even KO number is not assigned to it, therefore we could only point out its role in endometrial and ovarian cancer according to the expression level data available in the literature and the databases.

CHAPTER 5

Conclusion

Network modeling and simulation provides thorough understanding into the mechanisms of molecular entities of any pathway under the study. The simulations depicted the change in the concentration of the substrate in unit time till it achieves the steady state. The simulation analysis done for the MMR pathway model gave the insight of important genes i.e. MLH1, MSH2, hMSH6, PMS1 and PMS2 providing emphasis on the substrate concentration change in unit time and achievement of steady state. HE4 does not have any pathway data available as even KO number is not assigned to it, therefore we could only point out its role in endometrial and ovarian cancer according to the expression level data available in the literature and the databases.

Studies suggest that He4 gene is found to be over-expressed in ovarian cancer, but the involvement of HE4 gene in endometrial cancer is not supported by facts. This gene is involved in processes like proteolysis, spermatogenesis as depicted by Gene Ontology. As this gene functions in inhibiting peptide activity this results in degrading the protein products of MMR genes which are involved in endometrial and ovarian cancer, namely MLH1, MSH2, hMSH6, PMS1 and PMS2. Moreover, data retrieved from COSMIC database highlight the involvement of HE4 in both the cancers, as the gene is over-expressed in both the cases.

CHAPTER 6

Gaps

- During the germline DNA testing, rules for the detection of mutations can be determined, but the analysis of PMS2 gene is difficult due to presence of psuedogenes.

- Because of the limited studies on unselected endometrial cancer which is not because of age restriction, studies show that the number of mutation carriers is small and therefore it is difficult to find mutation spectrum due to difficulty in estimating the level of confidence and age of the patient under study.

- The performance of IHC and MSI also show some uncertainties in analysis of endometrial cancers and tumors, because mostly the data is available for the colorectal tumors, therefore there is an immediate need for optimal test parameters, which are specific and reliable, so that clinical sensitivity and specificity of these tests can be estimated.

- Next-generation sequencing technologies now-a-days plays an important role as well as they begin to enter in the clinical procedures, because of the high cost of clinical DNA sequencing NGS should also be taken into account.

References

1. Hofestädt, Ralf, et al. "Modeling and simulation of metabolic pathways, gene regulation and cell differentiation." BioEssays 18.4 (1996): 333-335.
2. Meng, Tan Chee, Sandeep Somani, and Pawan Dhar. "Modeling and simulation of biological systems with stochasticity." In silico biology 4.3 (2004): 293-309.
3. Bellomo, Nicola, N. K. Li, and Ph K. Maini. "On the foundations of cancer modelling: selected topics, speculations, and perspectives." Mathematical Models and Methods in Applied Sciences 18.04 (2008): 593-646.
4. Dada, Joseph O., and Pedro Mendes. "Multi-scale modelling and simulation in systems biology." Integrative Biology 3.2 (2011): 86-96.
5. Haoula, Zeina, Maisa Salman, and William Atiomo. "Evaluating the association between endometrial cancer and polycystic ovary syndrome." *Human reproduction* 27.5 (2012): 1327-1331.
6. Simmons, Archana R., Keith Baggerly, and Robert C. Bast Jr. "The emerging role of He4 in the evaluation of advanced epithelial ovarian and endometrial carcinomas." Oncology (Williston Park, NY) 27.6 (2013): 548.
7. Kahsar-Miller, Melissa D., et al. "Prevalence of polycystic ovary syndrome (PCOS) in first-degree relatives of patients with PCOS." *Fertility and sterility* 75.1 (2001): 53-58.
8. Azziz, Ricardo, et al. "The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report." *Fertility and sterility* 91.2 (2009): 456-488.
9. Geoffrey M Cooper *The Cell: A Molecular Approach*, 2nd ed. 2000.
10. Dexheimer, Thomas S. "DNA repair pathways and mechanisms." DNA repair of cancer stem cells. Springer Netherlands, 2013. 19-32.
11. Moore, Richard G., et al. "HE4 (WFDC2) gene overexpression promotes ovarian tumor growth." Scientific reports 4 (2014).
12. Manika Sehgal, and Tiratha Raj Singh (2014) DR-GAS: A database of functional genetic variants and their phosphorylation states in human DNA repair systems, DNA Repair , 16: 97-103 .

13. Shukla, Ankita, Ahmed Moussa, and Tiratha Raj Singh. "DREMECELS: A Curated Database for Base Excision and Mismatch Repair Mechanisms Associated Human Malignancies." *PloS one* 11.6 (2016): e0157031.
14. Funahashi, Akira, et al. "CellDesigner: a process diagram editor for gene-regulatory and biochemical networks." *Biosilico* 1.5 (2003): 159-162.
15. "Reactome Pathway Database". Reactome.org. N.p., 2017. Web. 27 Apr. 2017.
16. "COSMIC: Catalogue Of Somatic Mutations In Cancer - Home Page". Cancer.sanger.ac.uk. N.p., 2017. Web. 27 Apr. 2017.
17. "STRING: Functional Protein Association Networks". String-db.org. N.p., 2017. Web. 27 Apr. 2017.
18. "Gene Ontology Consortium | Gene Ontology Consortium". Geneontology.org. N.p., 2017. Web. 27 Apr. 2017.
19. Shukla, Ankita, Manika Sehgal, and Tiratha Raj Singh. "Hydroxymethylation and its potential implication in DNA repair system: A review and future perspectives." Gene 564.2 (2015): 109-118.
20. Milanowska, Kaja, et al. "REPAIRtoire—a database of DNA repair pathways." *Nucleic acids research* 39.suppl_1 (2011): D788-D792.

Appendix I

List of research papers consulted which are cited below:

S No.	Name	Author
1.	Genetic Testing Strategies in Newly Diagnosed Endometrial Cancer Patients Aimed at Reducing Morbidity or Mortality from Lynch Syndrome in the Index Case or Her Relatives.	Alison Steward Dated: 16 September 2013
2.	Role of endometrial cancer abnormal MMR protein in screening Lynch-syndrome families.	Qiongxian Long etal. Dated: 15 September 2014
3.	Mismatch repair gene expression defects contribute to microsatellite instability in ovarian carcinoma.	John P. Geisler et al. Dated: 25 September 2003
4.	MSH2 Mutation Carriers Are at Higher Risk of Cancer Than MLH1 Mutation Carriers: A Study of Hereditary Nonpolyposis Colorectal Cancer Families.	H. F.A. Vasen et al. Dated: October 2001
5.	Endometrial Carcinogenesis and Molecular Signaling Pathways.	Xianyong Ma1* et al. Dated: 1 June 2014
6.	DNA Mismatch Repair and Infertility	Sarmistha Mukherjee et al. Dated: November 2010
7.	Features of Ovarian Cancer in Lynch Syndrome	Kanako Nakamura et al. Dated: November 2014
8.	Common variants in mismatch repair genes and risk of invasive ovarian cancer.	Song H et al. Dated: November 2006
9.	Hereditary Syndromes manifesting as Endometrial Carcinoma: How can pathological features aid risk assessment?	Adele Wong et al. Dated: 23 November 2014
10.	Mismatch repair genes in Lynch syndrome: a review	Felipe Cavalcanti Carneiro da SilvaI et al. Dated: Jan. 2009
11.	Human Epididymis Protein 4 (HE4) Is a Secreted Glycoprotein that Is Overexpressed by Serous and Endometrioid Ovarian Carcinomas.	Ronny Drapkin et al. Dated: March 2005
12.	Expression of HE4 in Endometrial Cancer and Its Clinical Significance	Xiao Li et al. Dated: 11 January 2015