EVALUATION OF THE USEFULNESS OF TESTING FOR P53 MUTATION IN SOME HUMAN CANCERS

Nadia Helal[1,*], Rizk Abdel Moneim[2] & Marwa Yassein[2]
[1]Nuclear and Radiological Regulatory Authority, Nasr City 11762, P. O. Box 7551, Cairo, Egypt
[2]Physics Department, Faculty of Science, Helwan University, Ain Helwan 11795, Cairo, Egypt
*__Email__: Nadia_Helal2003@yahoo.com

ABSTRACT

The effectiveness of cancer radiotherapy is compromised by the small proportion of patients who sustain severe normal tissue damage after standard radiotherapy treatments. Predictive tests are required to identify these highly radiosensitive cases and screening for mutations in genes could be used as a predictive test.

Mutations in the tumor suppressor gene P53 are associated with a wide range of different cancers and may have prognostic and therapeutic implications. Methods for rapid and sensitive detection of mutations in this gene are therefore required. Some gene tests are used to clarify a diagnosis and direct a physician toward appropriate treatments, while others allow families to avoid having children with devastating diseases or identify people at high risk for conditions that may be preventable. One type of DNA testing involves comparing the sequence of DNA bases in a patient's gene to a normal version of the gene.

In this work, we propose a gene testing method, for this sake, we have 1) constructed a program using Matlab for screening P53-gene and to estimate the percentage of the four nuclides (A, G, C and T) along the gene sequence. 2) The constructed program has been applied for wild-type P53 (reference) and nine different cancers (test) for about 500 patients (the data has been taken from the IARC database, http://www.iarc.fr). 3) We evaluate the efficiency of the constructed program and compare it to DNA-counter software for mutation detection in nine different cancers and a high correlation has been obtained. We conclude that the constructed program is a robust, rapid, and comprehensive screening tool for sequence alterations in tumors. It allows one to make comparison between different sequences of any gene. It is believed that this study of P53 mutation may provide useful information for the diagnosis, prognosis and treatment of cancer. Estimating the percentage of mutations within cancer patients may also be of value in predicting adverse reactions to radiotherapy.

Keywords*: P53 Gene, DNA Counter, MATLAB program, and gene testing.*

1. INTRODUCTION

There is a range in the severity of normal tissue reactions when cancer patients receive standard radiotherapy treatment. Dose schedules have evolved to limit the proportion of highly radiosensitive (HR) adverse responses to about 5% of cases, Appleby *et al.,* [1]. If it were possible to identify these HR cases in advance of therapy, their treatment could be adjusted and it might then be possible to escalate the dose in the remaining patients to improve local control and cure rates, Athma *et al.,* [2]. P53 is a tumor suppressor gene, also known as "**Guardian of the genome**", "**Death star**", "**Good and bad cop**", and "**An acrobat in tumorigenesis**"**,** are just a few of the names that have been attributed to the P53-gene over recent years, Wiman *et al.,* and Soussi *et al.,* [3& 4]. The tumor suppressor P53 is a phosphoprotein barely detectable in the nucleus of normal cells. Upon cellular stress, particularly that induced by DNA damage, P53 can arrest cell cycle progression, thus allowing the DNA to be repaired; or it can lead to apoptosis. The most common changes of P53 in human cancers are point missense mutations within the coding sequences of the gene. Such mutations are found in all major histogenetic groups, including cancers of the colon (60%), stomach (60%), breast (20%), lung (70%), brain (40%), and esophagus (60%). It is estimated that P53 mutations are the most frequent genetic event in human cancers, accounting for more than 50% of cases, Sunthornwat *et al.,* [5]. Genetic code of P53-gene consists of sequences of four nucleotides (bases): adenine (A), cytosine (C), guanine (G), and thymine (T) which is arranged into codons, where a codon is a triplet of bases which codes for the production of an amino acid, Helal *et al.,* [6]. Recent studies have shown that cells carrying P53 mutations are more resistant to radiation and chemotherapy than cells with functional P53, Helal [7].

1.1 WHAT IS GENETIC TESTING?

Genetic testing is the scientific testing of a person's genes. Genetic testing is usually offered when someone is at a high risk of having inherited a faulty gene, based on a strong family history of cancer. It is possible to test for some faulty genes that increase the risk of certain cancers. For example, two genes called BRCA1 and BRCA2 are now known to be important in the development of hereditary breast and ovarian cancer, Sharon *et al.,* [8]. One of the

benefits of genetic testing is screening embryos for diseases. Pre-implantation genetic diagnosis (PGD) is a test that screens for genetic flaws among embryos used in *in-vitro* fertilization. With PGD, DNA samples from embryos created *in-vitro* by the combination of a mother's egg and a father's sperm are analyzed for gene abnormalities that can cause disorders. Fertility specialists can use the results of this analysis to select only mutation-free embryos for implantation into the mother's uterus. With PGD, couples are much more likely to have healthy babies, Pedro *et al.*, [9]. Gene-testing for mutations among breast cancer patients provided a predictive assay for highly radiosensitive responses, Appleby *et al.*, [1].

2. METHODOLOGY

P53-gene sequence missense mutation for 500 patients with P53-gene mutation was downloaded from IARC database, http://www.iarc.fr. Percentage of P53 nucleotides variations was calculated using Matlab constructed program. The Matlab output has been compared with the widely used DNA counter software.

2.1. MATLAB Program

MATLAB is a high-level language and interactive environment that enables one to perform computationally intensive tasks faster than with traditional programming languages such as C, C++, and FORTRAN. One can use MATLAB in a wide range of applications, including signal and image processing, communications, control design, test and measurement [10], financial modeling and analysis, and computational biology [10].

2.1.1 Constructed Program

The present work provides computer program with Matlab language for wild type and mutant type of P53-gene, using a series of cods to represent the percentage of each nucleotide in P53-gene. The inputs of constructed program are two P53-gene sequence text files, one has wild-type P53 sequence (reference) and the other has mutant type P53-gene sequence (test). Program compares these two files, and calculates the percentage of each nucleotide. The structure (flow chart) of present computer program is shown in figure 1.

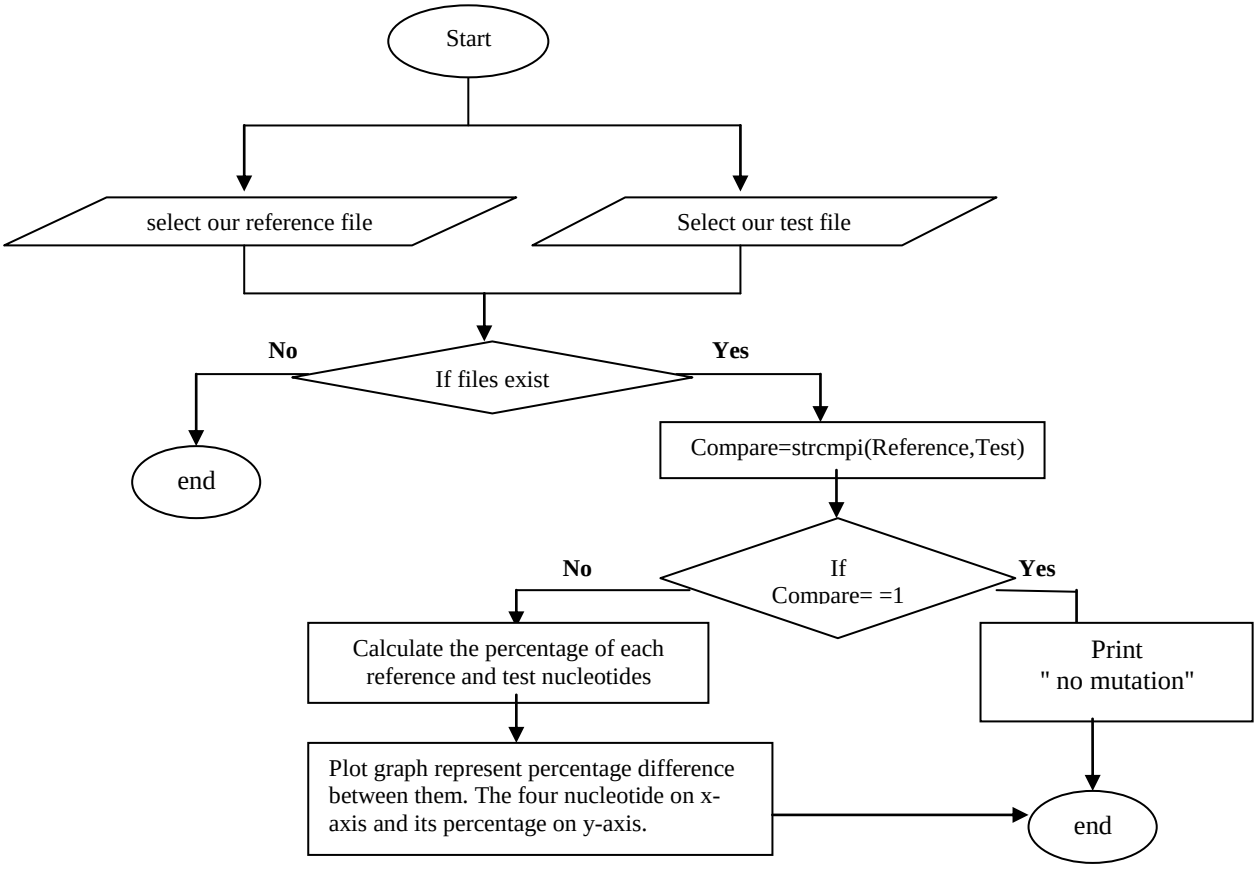


Figure 1: The flowchart represents MATLAB constructed program.

2.2. DNA Counter

The personal computer-based software: DNA Counter (DNA Nucleotide Counter) used in this work provides a useful tool for calculating the proportions of A,C,G, T nucleotides and CG, AT nucleotides in a DNA sequence. You do not need to install it or unpack it. Just download the program and double click to run it. Start DNA Nucleotide Counter. Paste your DNA sequence into the program. Press the 'Start' button. DNA Nucleotide Counter will show you the percent of A, C, G and T bases and the proportion between AT and CG bases as shown in figure 2.

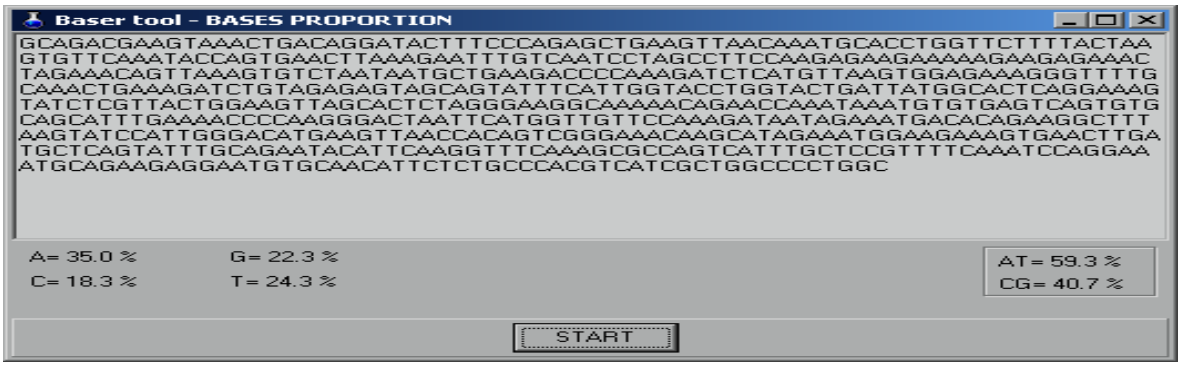


Figure 2: flowchart of the DNA-counter, with P53 sequence as an input.

3. RESULTS AND DISCUSSIONS

The present work provides a computer program with MATLAB language for nine cancer types: breast, bladder, lung, rectal, head and neck, ovarian, gastric, esophagus and hepatocellular carcinoma. Using a series of codes and functions to represent wild type and mutant p53-gene was presented. The data has been taken from the IARC database, http://www.iarc.fr .

The Matlab program provides a powerful programming language, as well as an interactive computational environment. You can enter commands from the language one when the Matlab command line in command window, or you can write a series of commands to a file that you then execute as any Matlab function.

Then, you can use the Matlab editor or any other text editor to create your own program file and save it in the matlab current folder. To view the contents of a program file, for example, "myfile.m", double click on it and to run it type the file name "myfile" in command window as shown in figure 3 below.

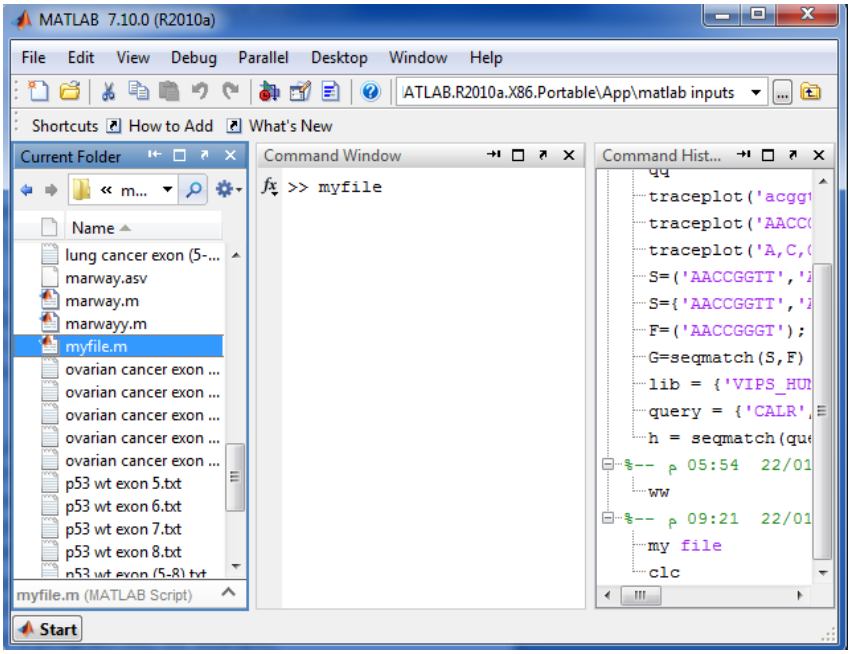


Figure 3: Matlab main window program file "myfile" created in current folders, to run it type the program file name in the command window.

Exons 5-8 have been screened for both wild-type and mutant type P53 of the 500 patients. Applying the constructed program with Matlab, the percentage of each nucleotide (A, G, C, and T) along the gene sequence is given. The following graphs represent the wild type and mutant type P53 gene in different cancers types using MATLAB constructed program.

The constructed program window for wild-type P53 and mutant P53 sequence is shown in figure 4.

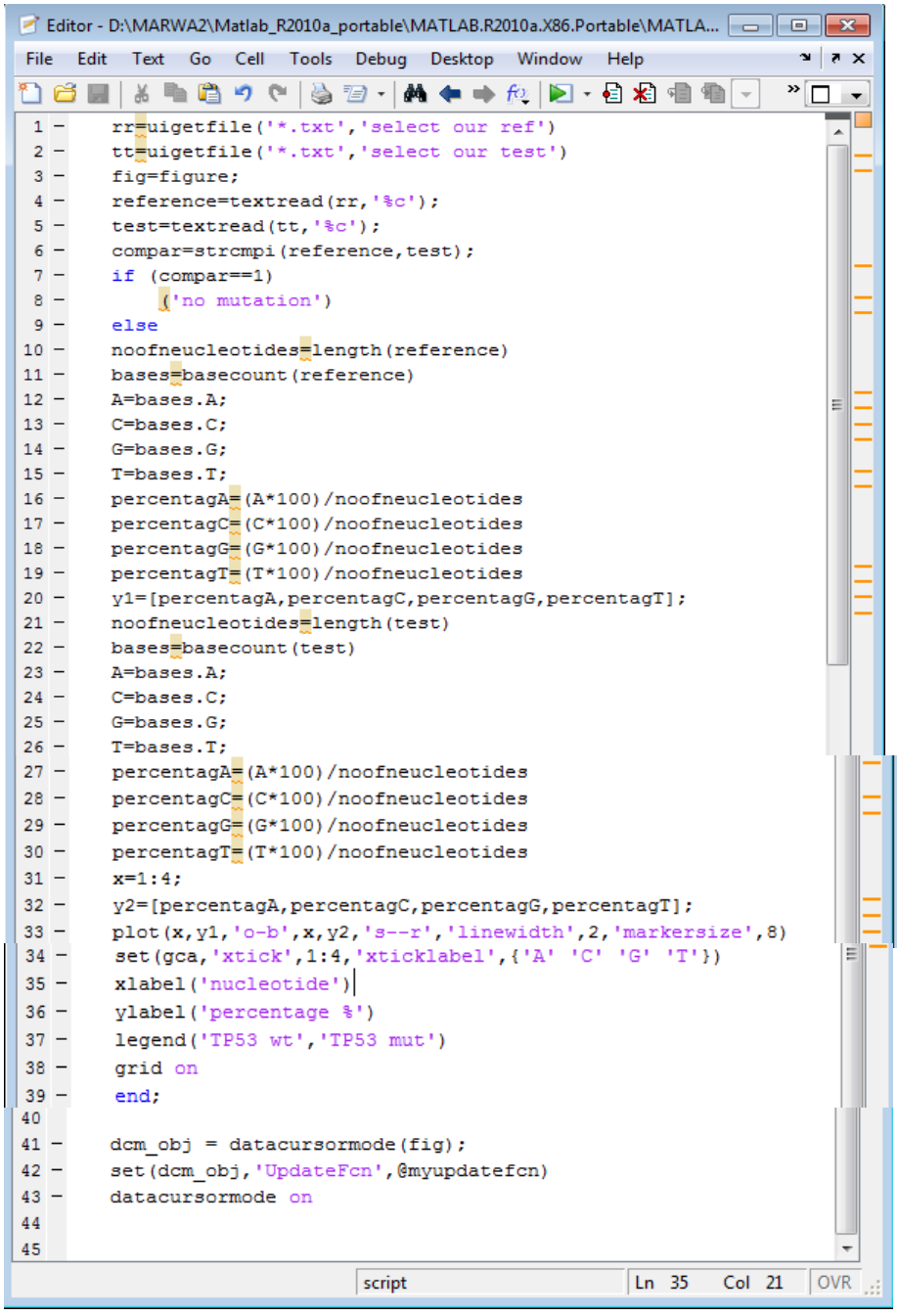
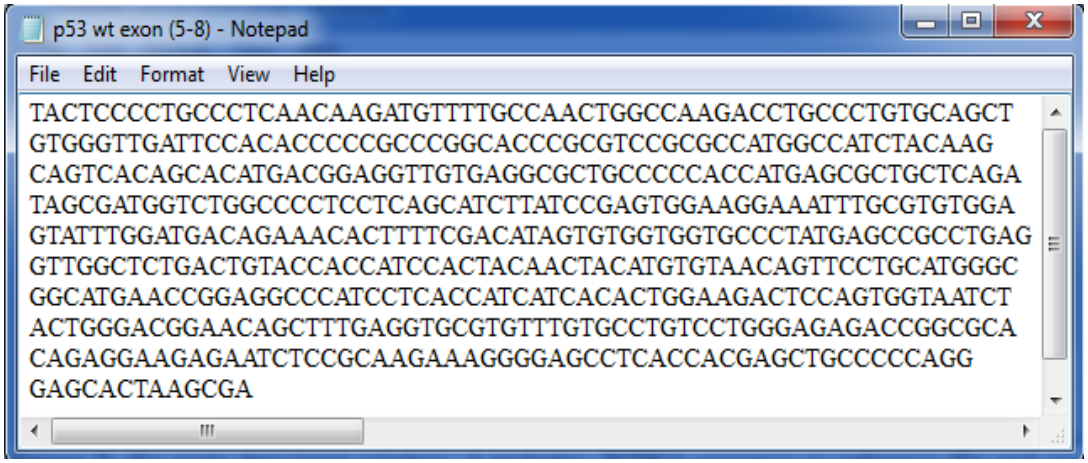


Figure 4: the constructed matlab program in editor matlab window.

Figure 5, shows an example of the input data as a text file (wild-type P53 gene as a reference file and mutant sequences of breast cancer as a test files), which saved in matlab current folder.



P53-gene wild type sequence.

Breast cancer P53-gene exon (5-8) sequence.

Figure 5: the main data-input window, it includes: WT P53 sequence as reference file and mutant sequences of breast as test files.

Figure 6, represents the two dimensional graphical outputs of analyzed sequence data of breast, bladder, lung and rectal cancers. The solid line represents the wild-type P53 gene, while the dotted line represents the mutant sequences under investigation.

The constructed model and the obtained graphs have the potential to detect the gene mutations since the varying degree of nucleotides suggesting the presence of mutation. To know the percentage of each nucleotide along P53-gene, clicking the mouse on appoint of interest, displays data value of the point clicked.

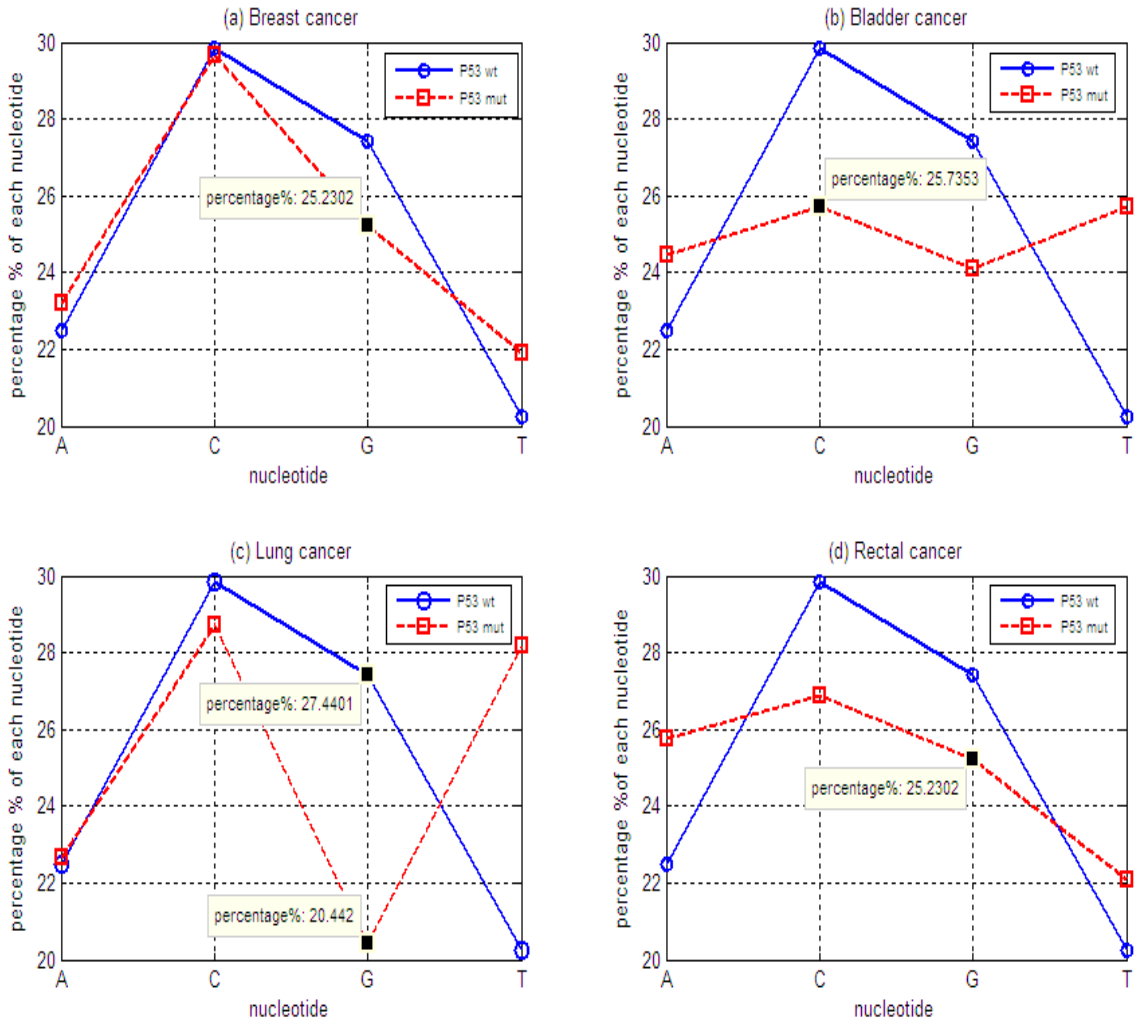


Figure 6: main project window presented by Matlab program once data analysis of exons 5-8 is complete and showed in 2D graphical form. a) Breast cancer, b) Bladder cancer, c) Lung cancer and d) Rectal cancer.

The samples had a varying percentage of the nucleotides, and based on the visual inspections, these percentages are varied from one cancer to another. From figure 5 (a) of breast cancer, the percentage of A (ratio of 23.20: 22.46) and C (ratio of 29.65:29.83) are so close to the wild-gene. Both guanine (ratio of 25.23:27.44) and thymine (ratio of 21.91:20.25) also show slight deviation from the normal gene.

In case of bladder cancer at figure 5 (b), a remarkable deviation from the normal sequence for the four nucleotides: A, C. G and T is seen. A high percentage of guanine (ratio of 20.44:27.44) and thymine (ratio of 28.17:20.25) bases is obvious in lung-cancer sequence (fig. 5c) in comparison to adenine (ratio of 22.65:22.46) and cytosine (ratio of 28.72:29.83).

The fourth graph (5d) represents rectal cancer, where moderate deviation of the four bases from the normal graph is remarkable A (ratio of 25.78:22.46), C (ratio of 26.88:29.83), G (ratio of 25.23:27.44) and T (ratio of 22.09:20.25).

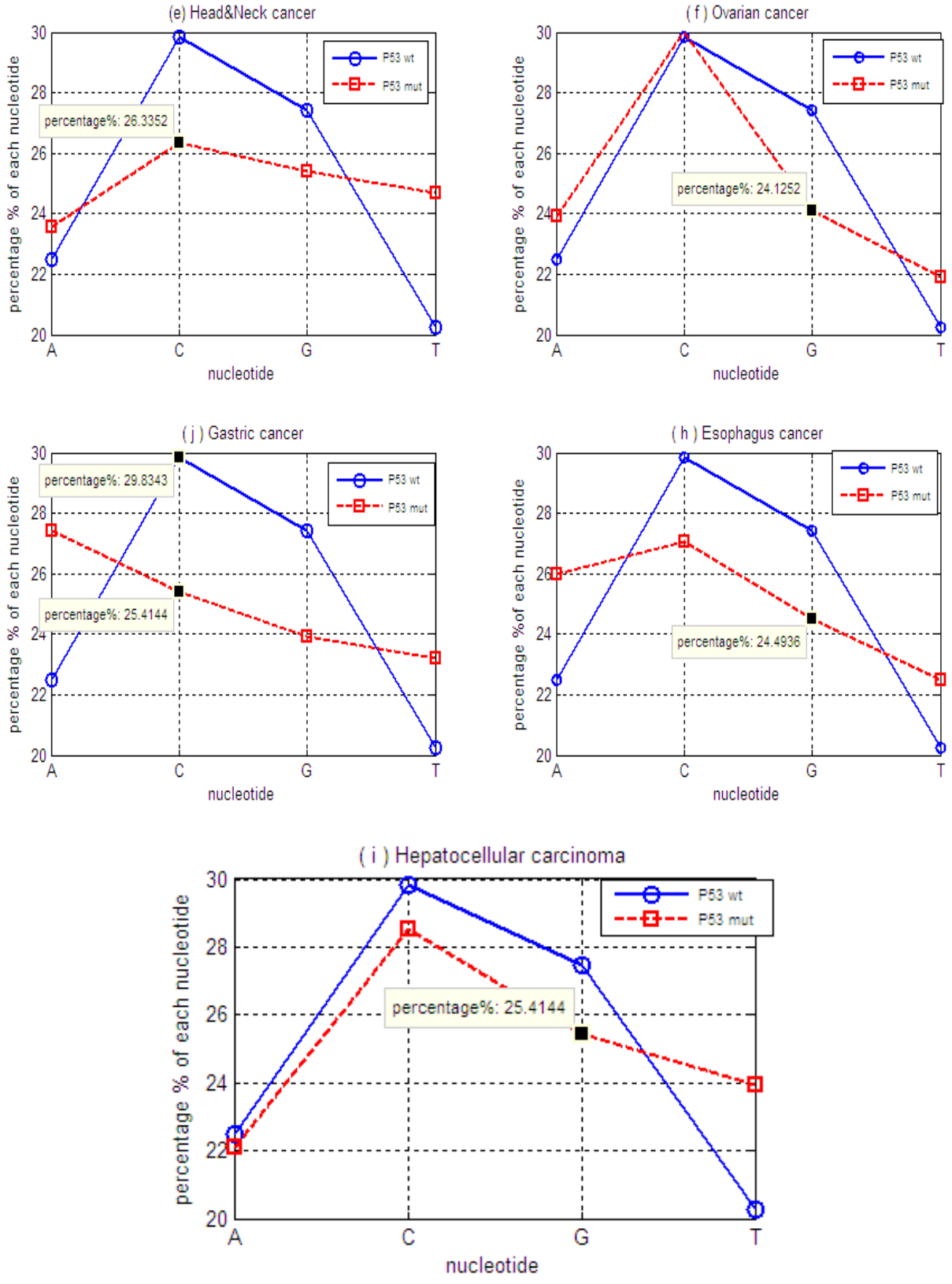


Figure 6: main project window presented by Matlab program once data analysis of exons 5-8 is complete and showed in 2D graphical form. e) head& neck, f) ovarian, g) gastric, h) esophagus and i) hepatocellular carcinoma.

For head and neck graph 6 (e), the mutation percentage in base A (ratio of 23.57:22.46) is slighter than other bases C (ratio of 26.33:29.83), G (ratio of 25.41:27.44), and T (ratio of 24.67:20.25).

In case of ovarian cancer 6 (f), the calculated variations in mutant sequence in comparison to normal one is given by A (ratio of 23.94:22.46), C (ratio of 30.01:29.83), G (ratio of 24.12:27.44) and T (ratio of 21.91:20.25).

For gastric cancer 6 (g), A (ratio of 27.44:22.46), C (ratio of 25.41:29.83), G (ratio of 23.94: 27.44), and T (ratio of 23.20:20.25). The calculated percentages for esophagus cancer 6 (h) are A (ratio of 25.96:22.46), C (ratio of 27.07:29.83), G (24.49:27.44) and T (ratio of 24.49:20.25). Finally, the hepatocellular carcinoma 6 (i) A (ratio of 22.09:22.46), C (ratio of 28.54:29.83), G (ratio of 25.41:27.44) and T (ratio of 23.94:20.25).

In this part, a comparison between output results of the constructed Matlab program and DNA Counter software has been done for further validation of the constructed program. Table 1 represents this comparison.

Table 1: the percentage of each nucleotide in wild-type P53 gene and the corresponding mutant type in different cancers, using both Matlab and DNA-counter programs.

Nucleotide	**A%**		**C%**		**G%**		**T%**	
	Counter [a]	**Matlab**	**Counter[a]**	**Matlab**	**Counter[a]**	**Matlab**	**Counter[a]**	**Matlab**
W.T P53	22.5	22.4678	29.8	29.8343	27.4	27.4401	20.3	20.2578
Breast cancer	23.2	23.2044	29.7	29.6501	25.2	25.2302	21.9	21.9153
Bladder cancer	24.4	24.4485	25.7	25.7353	24.1	24.0809	25.7	25.7353
Lung cancer	22.7	22.6519	28.7	28.7293	20.4	20.442	28.2	28.1768
Rectal cancer	25.8	25.7827	26.9	26.8870	25.2	25.2302	22.1	22.0994
H&N cancer	23.6	23.5727	26.3	26.3352	25.4	25.4144	24.7	24.6777
Ovarin cancer	23.9	23.9411	30	30.0184	24.1	24.1252	21.9	21.9153
Gastric cancer	27.4	27.4401	25.4	25.4144	23.9	23.9411	23.2	23.2044
Esoph. cancer	26	25.9669	27.1	27.0718	24.5	24.4936	24.5	24.4936
Hepat. cancer	22.1	22.0994	28.5	28.5451	25.4	25.4144	23.9	23.9411

[a] *DNA Counter, W.T P53 (wild-type P53).*

From the above table, the calculated percentage of the mutant P53 nucleotides by DNA-Counter software are: 1) adenine nucleotide (A) percentage is varies from 22.7 to 27.4 (WT=22.5), 2) cytosine nucleotide (C) percentage is varies from 25.4 to 29.7 (WT=29.8), 3), guanine nucleotide (G) percentage varies from 20.4 to 25.4 (WT=27.4) and 4) and thymine nucleotide (T)percentage is varies from 21.9 to 28.2 (WT=20.3).
The correlation coefficient was estimated for the data obtained by both DNA-Counter software and the constructed model using Matlab. The goodness of fit was listed in table 2.

Table 2: Correlation coefficient for different cancers.

Cancer type	**Correlation coefficient**
Wild type P53-gene	1
Breast cancer	0.99998
Bladder cancer	0.99956
Rectal cancer	0.99997
Lung cancer	0.99995
Head and Neck	0.99996
Ovarian cancer	1
Gastric cancer	1

Esophagus cancer	0.99998
Hepato. carcinoma	0.99998

From the above table, the correlation coefficient between DNA-Counter and Matlab constructed program is remarkably high (wild-type is the "reference" and the mutant genes are the "test"). For example: ovarian and gastric cancers provided the best fit (correlation coefficient=1), followed by breast, esophagus and hepatocellular carcinoma (correlation coefficient = 0.99998), rectal cancer (correlation coefficient = 0.99997), head and neck (correlation coefficient = 0.99996), lung cancer (correlation coefficient = 0.99995), and finally bladder cancer (correlation coefficient = 0.99956).

In this work, the constructed program has been shown to be accurate and to have a high sensitivity in detecting low amounts of mutated P53 in the presence of wild-type one. The potential of this method for mutation screening led us to investigate whether Matlab program could match the high sensitivity, specificity, and high-throughput of DNA-counter as a screening method.

4. CONCLUSION

We have presented in this paper a mathematical methodology for clinical reporting of genetic variants in tumor suppressor gene (P53) conferring a high risk of cancer. Both under- and over-interpretation of such a result are common in clinical practice. The purpose of this study is to improve the clinical utilization of genetic testing results, to maximize the opportunity to learn more about variants for the benefit of other families and to minimize the risk of incorrect interpretation of variants in the clinical setting. If applied successfully, the principles of this approach where analysis of the gene detects sequence variants for which a clear functional consequence cannot be immediately assigned.

Applying the above work for any normal gene and mutant one, a positive test result means that there is a change found in gene of interest. Depending on the purpose of the test, this result may confirm a diagnosis, indicate that a person is a carrier of a particular genetic mutation, identify an increased risk of developing a disease (such as cancer) in the future, or suggest a need for further testing. Because family members have some genetic material in common, a positive test result may also have implications for certain blood relatives of the person undergoing testing.

Finally, we have evaluated the efficiency of the constructed MATLAB program and compare it to DNA-counter software for mutation detection in nine different cancers. We conclude that the constructed program is a robust, rapid, and comprehensive screening tool for sequence alterations in different cancer patients. The program is usable to detect base-percentages or base-rich which is an important tool to represent any imbalances of the four nucleotides along any gene sequence.

5. REFERENCES

[1]. J.M.Appleby, J.B.Barber, E.Levine, J.M.Varley, A.M.Taylor, T.Stankovic, J.Heighway, C.Warren, D.Scott. "Absence of Mutations in the ATM gene in Breast Cancer Patients with Severe Responses to Radiotherapy". British Journal of Cancer, **76**(12), 1546-1549, 1997.

[2]. P.Athma, R.Rappaport, M.Swift. "Molecular Genotyping shows that Ataxiatelangiectasia Heterozygotes are Predisposed to Breast Cancer". Cancer Genet. Cytogenet, **92**: 130-134, 1996**.**

[3]. K Wiman, P.Hainaut. "25 Years of p53 Research". *Springer*, 1st edition, PP: 225-292, 2005.

[4]. T.Soussi. "The p53 Tumor Suppressor gene: from Molecular Biology to Clinical Investigation". *Annals of the NewYork Acadimy of Sciences*, vol. **910**, 2000.

[5]. R.Sunthornwat, E.Moore, Y.Temtanapat." Detecting and Classifying Mutations in Genetic Code with an Application to β-thalassaemia". *Science Asia*, vol. **37**, PP: 51-61, 2011.

[6]. N.L.Helal, Dorrah M, C. Li "Ladder Like Graphical Representation of p53 Gene Alterations in Some Human Cancers". Isotope& Rad. Res., 37, **6**, 1477-1487, 2005.

[7]. N. L. Helal, "Simple Mathematical Method to Quantify P53 Mutations in Lung Cancer Associated with Occupational Radon" VII Radiation Physics& Protection Conference, 27-30, November 2004, Ismalia-Egypt.

[8]. E.Sharon Poln, M.Diana Eccles, E.Douglas , D.William Foulkes, Maurizio Genuardi, S.Marc Greenblatt, B. L. Frans Hogervorst, H.Nicoline, B.Amanda . T.Spurdle, and Sean . "Sequence Variant Classification and Reporting: Recommendations for Improving the Interpretation of Cancer Susceptibility Genetic test Results". Hum. Mutat. **29**(11): 1282-1291, 2008**.**

[9]. K.Pedro, B.Anna, D.Eldri Undlien, W.Yun, T.Elda ,M. Jahn Nesland, N.Aune, T.Neeme, B.Anne-Lise. "Evaluation of Arrayed Primer Extension for TP53 Mutation Detection in Breast and Ovarian Carcinomas". *BioTechniques* **39***:* 755-761, 2005**.**

[10]. http://www.mathworks.com/products/matlab/