 doi:10.20944/preprints202008.0562.v1

# Mutational Hotspots and Conserved Domains in P53 Tumour Suppressor Protein

Olabode E. Omotoso *

Cancer Research and Molecular Biology Laboratories, Department of Biochemistry, University of Ibadan

***Correspondences to:**

Olabode Omotoso

Email: olabodeomotoso@gmail.com

Tel.: +234703 0927854

 doi:10.20944/preprints202008.0562.v1

**Abstract**

**Introduction:** The tumour suppressor protein p53 commonly referred to as guardian of the genome plays important role in preserving the genome through the regulation of programmed cell death, DNA repair, energy metabolism, cell cycle entry or exit and senescence. Mutations in p53 can either result to a loss of tumour suppressor function or gain of oncogenic properties. Hence, mutations in p53 are the most frequent genetic mutational alteration in human cancers, associated with worse prognosis and more aggressive disease outcome.

**Methods:** To assess the mutational hotspots and conserved regions of p53, I analyzed 76 complete p53 protein sequences covering whole exons were mined from the NCBI GenBank database. Multiple sequence alignment (MSA) was done using ClustalX version 2.1.

**Results:**

Thirty-five (35) mutations were identified with more frequent mutations in amino acid (aa) position 72 and 79 (Exon 4), amino acid deletion in codon 112-122 (Exon 4), codon 213 (Exon 6), codon 248 (Exon 7), codon 273 (Exon 8) and codon 278 (Exon 8). Mutations at amino acid position 79, 248, 278 located in the DNA-binding domain exhibited more than one alteration in same position.

**Conclusions:** Findings from this study revealed the prevalence of mutations in the DNA binding domain of p53 and the structure-function effect of the mutations. Assessing the pattern and frequency of p53 alterations, and analyzing it thoroughly for each carrier, could help in identifying correlations between p53 status, disease outcome and possible candidate for gene therapy.

**Keywords: mutations, exon, p53, regulation, DNA, tumour suppressor**

 doi:10.20944/preprints202008.0562.v1

## Introduction

The *p53* gene is a very crucial tumour suppressor gene, commonly referred to as guardian of the genome because of its role in conserving genomic stability. The tumour suppressor, *p53* gene, a 53-kDa nuclear phosphoprotein, composed of 19180 bp, spans 11 exons and 10 introns which codes for 393 amino acids is located on the short arm of human chromosome 17 (17p13.1 locus), structurally and functionally divided to distinct domains; the transactivation domain (TAD); proline rich domain; DNA-binding domain (DBD); oligomerization domain; and c-terminus regulatory domain (CRD) [1,2]. Exon 1 is non-coding but notable for the very large intron that comes after it and the ability to form a tightly bound stem-loop structure to wild type p53, while exons 2 and 3 encodes the TAD; a serine-threonine-rich region which permits the interaction of p53 with MDM2 (murine double minute 2) and other proteins and a phosphorylation site by ATM. Exons 4 to 8 encodes the DNA binding domain and majority of reported mutations in p53 are in this region [2]. Exons 9 and 10 encode the oligomerization domain, which is responsible for the interaction with Rad51 and BRCA1 and formation of an active tetramer [3]. Functionally, p53 plays important role in programmed cell death (apoptosis), cell cycle regulation, angiogenesis, cellular stress response, cell proliferation, DNA repair by activating GADD 45 (Growth arrest and DNA damage) transcription and binding to ERCC3, modulation of senescence in response to cellular insult and ageing [4–7].

In addition to its role in cell cycle regulation, facilitating DNA repair and programmed cell death, p53 also plays important role in the antioxidant and energy metabolism regulation [8,9]. Tumour cells depend largely on energy metabolism as precursors for macromolecule biosynthesis, in order to cater for their rapid growth and proliferation requirement. Reactive oxygen species (ROS) generated through oxidative stress can initiate cellular insult resulting to DNA damage. P53 in a crosstalk controls intracellular ROS by up-regulating nuclear factor

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2020 doi:10.20944/preprints202008.0562.v1

erythroid-related factor 2 (NRF2) in response to mild stress and inhibits NRF2 in response to severe stress, in order to induce apoptosis. In mild stress scenario, NRF2 links with p21 in the nucleus and transactivate antioxidant enzymes to stabilize intracellular ROS level [4,8]. Since first published findings on p53 mutations involvement in carcinogenesis in 1989 [10], mutations in p53 have been a subject of interest in cancer biology, being the most frequent genetic mutational alteration accounting for more than 50% of cases in human cancers [6,11,12]. Mutations in p53 have been associated with worse prognosis and more aggressive disease outcome in several cancer types [6,12–14] which are also reported in the International Agency for Research on Cancer (IARC) database [15]. High frequency mutations in p53 resulting to increased risk, chemoresistance and poor prognosis have been reported in endometriosis [16], skin cancer [17], breast cancer [18], cervical cancer [9,19,20], testicular cancer [21], colorectal cancer [2,5] and many other cancer types [6,9,11,22]. Earlier studies have identified missense mutation as the most prevalent in p53, targeting exons 5-8 which codes for the DNA binding domain [8,14].

Most mutations observed in p53 impair its DNA-binding ability, thereby allowing cellular proliferation in state where cells with intact p53 function are regulated or suppressed [6,8]. Hence, during cellular insult or damage, cells with mutated p53 are unable to initiate cell cycle arrest, DNA repair or apoptosis, thereby resulting to unregulated proliferation, metastasis, and invasion of damaged cells [2,12,23]. Mutation pattern in p53 can help understand the contribution of endogenous events and exogenous agents in carcinogenesis or progression. This may help with respect to diagnosis, detection and in definitive approach to cancer therapy. Thus, this study was carried out using multiple sequence alignment (MSA) to investigate the frequency of mutations, mutational hotspots and conserved domains in p53 spanning exons 1 through 11.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2020 doi:10.20944/preprints202008.0562.v1

## Materials and methods

### *Data acquisition*

From the available p53 sequence in NCBI database, total 76 p53 protein sequences fit into the inclusive criteria for this study; filtered as "high coverage only, *Homo sapiens*, complete (393 aa), > 97% percent identity, all exons, and low coverage excl" were mined from the NCBI database (https://ncbi.nim.nih.gov/protein/?term=P53). The protein sequences in the NCBI database used in this study are listed in Table S1 (with accession numbers with corresponding amino acid sequences in FASTA format).

### *Multiple sequence alignment (MSA)*

The retrieved p53 protein sequences were aligned with the reference protein [24] by ClustalX version 2.1 [25] MSA software using the default parameters [26].

### *Sequence and mutational analysis*

In this study, the filtered data were used to analyze the polymorphisms and mutations in the p53 sequences in order to identify similar or conserved domains and mutational patterns. Mutations that occur multiple times independently were focused on as they are likely candidates for understanding the mutational landscape in p53 while 28 mutation sites with only a single mutation candidate were not fully explored. After exclusion, we observed mutations in multiple sequences at 7 different positions spread across exons 4, 6, 7, and 8 in the p53 protein sequence.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2020 doi:10.20944/preprints202008.0562.v1

**Results**

The result of the MSA of the protein sequences with NP_001119584.1 as reference protein sequence using ClustalX version 2.1. Mutations observed are reported in Table 1 with more frequent mutations in codon 72 and 79 (Exon 4), amino acid deletion in codon 112-122 (Exon 4), codon 213 (Exon 6), codon 248 (Exon 7), codon 273 (Exon 8) and codon 278 (Exon 8).

**Table 1.** Mutations observed in the p53 protein sequence with the corresponding mutation type

| S/N | Amino acid change | Exon | Frequency (%) | Mutation type |
|---|---|---|---|---|
| 1 | L14fs* | 1 | 1.3% | Frameshift (del) |
| 2 | P72R | 4 | 54% | Polymorphism, Missense |
| 3 | A76G | 4 | 1.3% | Missense |
| 4 | A79T, A79fs* | 4 | 6.6% | Missense, Frameshift (del) |
| 5 | 112-122fs* | 4 | 2.6% | Frameshift (del) |
| 6 | K139N | 5 | 1.3% | Missense |
| 7 | C141Y | 5 | 1.3% | Missense |
| 8 | V143M | 5 | 1.3% | Missense |
| 9 | T155P | 5 | 1.3% | Missense |
| 10 | P158H | 5 | 1.3% | Missense |
| 11 | Y163H | 5 | 1.3% | Missense |
| 12 | K164R | 5 | 1.3% | Missense |
| 13 | C176F | 5 | 1.3% | Missense |
| 14 | H179R | 5 | 1.3% | Missense |
| 15 | R181C | 5 | 1.3% | Missense |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2020 doi:10.20944/preprints202008.0562.v1

| | | | | |
|---|---|---|---|---|
| 16 | L188M | 6 | 1.3% | Missense |
| 17 | H193R | 6 | 1.3% | Missense |
| 18 | G199V | 6 | 1.3% | Missense |
| 19 | R213Q | 6 | 2.6% | Missense |
| 20 | M237I | 7 | 1.3% | Missense |
| 21 | C238Y | 7 | 1.3% | Missense |
| 22 | S241F | 7 | 1.3% | Missense |
| 23 | M246T | 7 | 1.3% | Missense |
| 24 | R248Q, R248N | 7 | 7.9% | Missense |
| 25 | I254D | 7 | 1.3% | Missense |
| 26 | 261fs* | 8 | 1.3% | Frameshift (ins) |
| 27 | R267L | 8 | 1.3% | Missense |
| 28 | R273H | 8 | 5.3% | Missense |
| 29 | C275W | 8 | 1.3% | Missense |
| 30 | P278A, P278L | 8 | 4% | Missense |
| 31 | E286K | 8 | 1.3% | Missense |
| 32 | L289F | 8 | 1.3% | Missense |
| 33 | P309S | 9 | 1.3% | Missense |
| 34 | R337H | 10 | 1.3% | Missense |
| 35 | S362G | 10 | 1.3% | Missense |

*Key: One letter code for corresponding amino acid: A – Alanine, R – Arginine, N – Asparagine, D – Aspartic acid, C – Cysteine, E – Glutamate, Q – Glutamine, G – Glycine, H – Histidine, I – Isoleucine, L – Leucine, K – Lysine, M – Methionine, F – Phenylalanine, P – Proline, S – Serine, T – Threonine, W – Tryptophan, Y – Tyrosine, V – Valine

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2020 doi:10.20944/preprints202008.0562.v1

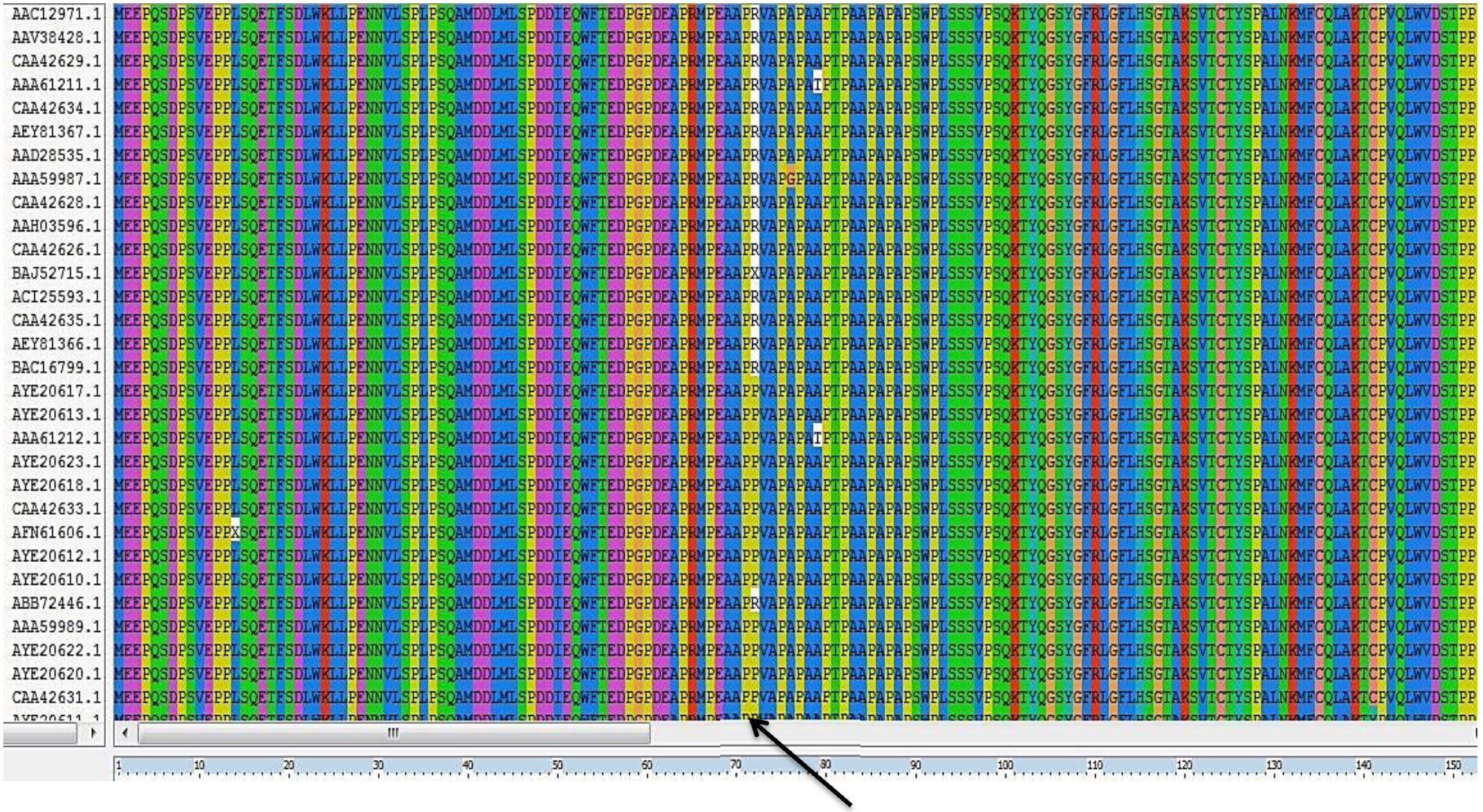


**Figure 1: Multiple sequence alignment showing the P72R polymorphism with respect to the Reference protein sequence (**NP_001119584.1**).** The highly recurrent polymorphism showing the proline (P) and arginine (R) polymorphs at position 72 (as indicated) present in exon 4 coding for the DNA binding domain; the most prevalent polymorphic site in p53.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2020 doi:10.20944/preprints202008.0562.v1

## Discussion

Mutations in the *p53* gene promotes genomic instability which plays a crucial role in the pathogenesis of several cancer types [1,27]. Wild-type p53 is only available in minute quantities in normal cells due to the short half-life. Mutations lead to more stable p53 resulting in elevated mutated p53 capable of genomic instability. Mutations in p53 either result to gain of function to facilitate tumour progression or loss of tumour suppression ability. In the present study we used multiple sequence alignment to detect mutations and conservations in all exons of 76 human p53 protein sequences. As observed in this study, the most frequent mutations in p53 were missense mutations, this is corroborated by other studies [12,18,27]. The present study in support with earlier findings [2,27,28] observed primary localization of p53 mutations in exons 4, 6, 7 and 8 and as mutation hotspots.

Mutations in the DNA binding domain of p53, alters the capacity to bind DNA and are deleterious to survival [3]. Earlier reports [3,8] have linked missense mutations in the DNA binding domain of p53 to poor prognosis, modification of transcriptional function and also loss of the oligomerization domain, which can induce a truncated protein with no capacity to form tetramers.

Three (3) recurrent mutations (P72R, A79T and 112fs*) localized in exon 4 of the p53 was observed. The exon 4 of the *p53* gene was characterized by two types of polymorphism at amino acid 72 located in the proline-rich domain; the arginine and proline polymorph. This polymorphism have been reported to increase cervical cancer risk by increasing p53 sensitivity to HPV- E6-mediated ubiquitin-dependent proteolysis [15,29]. The allelic frequencies, 0.54p53Arg and 0.46p53Pro, found in this study, corresponded with those described in earlier reports [1,9,16,30]. Racial variability exists in the frequency of 72 amino acid mutation. The P72 polymorph is more frequent in African-American and Chinese, while the

 doi:10.20944/preprints202008.0562.v1

R72 is common in Caucasians [18]. The P72 variant facilitates DNA repair and cell-cycle arrest, while the R72 variant interacts well with MDM2 than the P72 variant resulting to a higher efficiency at inducing apoptosis. A study [9] of 80 Greek women positive for Human papilloma virus (HPV) infection revealed a 4.17-fold higher risk of high grade intraepithelial lesions in women homozygous for the R72 variant. On the other hand, other investigations suggested the contribution of the pro/pro genotype to cervical cancer development in Chinese, Indian and Korean women [9].

C176, H179, and C238 localized in exons 5 and 7 coordinates zinc binding which plays critical role for correct folding [21]. Hence, missense mutations in these positions (C176F, H179R and C238Y) observed in this study may be deleterious as it affects DNA binding. The p53 mutant R248Q have been shown to be involved in chromatin remodeling, promoting cell proliferation, invasion and migration while R248W facilitates induction of angiogenesis and increasing colony formation ability [2]. The hotspot p53 mutants R248W and R273H interact with p63 and p73, which prevents gene expression and inhibits their binding to targeted promoters. A mutant p53 R273H/miR-27a/EGFR pathway has been reported to play crucial roles in facilitating tumour cell proliferation, colony formation and carcinogenesis [2]. In hepatocarcinoma, R267L and R273H mutants localized in exon 8 observed in this study are shown to upregulate CXCL5 and Axl expression, resulting to uncontrolled cell growth, cell migration, invasiveness and carcinogenesis [2]. The R273H mutant attenuates NRF2 activation, leading to NRF2 functional loss in ROS detoxification. Hence, the resultant elevated level of ROS generated promotes unregulated cell growth, colony formation and inability to induce cell growth arrest. The R337H mutant protein sequence in the present study was derived from a breast cancer patient [31]. This corroborates previous finding [18] which demonstrated that the R337H mutation might significantly increase the breast cancer risk in carriers (P=0.0442).

## Conclusion

The present study revealed the hotspot mutations in p53 protein sequences and their impact to the structure-function as the guardian of the genome. The results showed the prevalence of missense mutations as the most recurrent mutation type in p53. Synonymous (silent) mutations do not change the amino acid, hence, this study could not account for them due to the focus on the protein sequences. Most mutations observed are present in exons 4, 6, 7 and 8 encoding the proline rich and DNA binding domains of p53, with notable conservation of the transactivation and oligomerization domains. These mutational hotspots can serve as important biomarker in screening, diagnosis, monitoring metastasis and follow-up in case of possible remission.

## Declaration of Competing Interest

The author declare no conflict of interest

## Acknowledgements

The author would like to thank those who sequenced complete p53 gene sequence and deposited in the NCBI repository.

## References

1. Verheijen, F. M. *et al.* TP53 mutations in human meningiomas. *Int. J. Biol. Markers* **17**, 42–48 (2002).

2. Li, H. *et al.* Targeting the Oncogenic p53 Mutants in Colorectal Cancer and Other Solid Tumors. *Int. J. Mol. Sci.* **20**, 1–19 (2019).

3. Vegran, F. *et al.* Only Missense Mutations Affecting the DNA Binding Domain of P53 Influence Outcomes in Patients with Breast Carcinoma. *PLoS One* **8**, 1–8 (2013).

4. Rufini, A., Tucci, P., Celardo, I. & Melino, G. Senescence and aging: the critical roles of p53. *Oncogene* 1–15 (2013). doi:10.1038/onc.2012.640

5. Vogelstein, B., Surojit, S. & Prives, C. p53 : The Most Frequently Altered Gene in Human Cancers. *Nat. Educ.* **3**, 4–7 (2010).

6. Hainaut, P. *et al.* Database of p53 gene somatic mutations in human tumors and cell lines : updated compilation and future prospects. *Nucleic Acids Res.* **25**, 151–157 (1997).

7. Fridman, J. S. & Lowe, S. W. Control of apoptosis by p53. *Oncogene* **22**, 9030–9040 (2003).

8. Perri, F., Pisconti, S. & Scarpati, G. D. V. P53 mutations and cancer: a tight linkage. *Ann. Transl. Med.* **4**, 522 (2016).

9. Moschonas, G. D. *et al.* Association of codon 72 polymorphism of p53 with the severity of cervical dysplasia , E6-T350G and HPV16 variant lineages in HPV16-infected women Cervical specimens. *J. Med. Microbiol.* **66**, 1358–1365 (2017).

10. Baker, S. J. *et al.* Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science (80-. ).* **244**, 217–221 (1989).

11. Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science* **253**, 49–53 (1991).

12. Souza, D. M. B. *et al.* Detection of mutations within exons 4 to 8 of the p53 tumor suppressor gene in canine mammary glands. *Arq Bras Med Vet Zootec* **64**, 341–348 (2012).

13. Chizumi, F. *et al.* A Comprehensive System to Explore p53 Mutations. *Anat. Pathol.*

 doi:10.20944/preprints202008.0562.v1

**110**, 368–373 (1998).

14. Makwane, N. & Saxena, A. STUDY OF MUTATIONS IN p53 TUMOUR SUPPRESSOR GENE IN HUMAN SPORADIC BREAST CANCERS. *Indian J. Clin. Biochem.* **24**, 223–228 (2009).

15. Mutation Nomenclatures. *IARC TP53 Database* (2019). Available at: http://p53.iarc.fr/HGVSstandards. (Accessed: 14th January 2019)

16. Hsieh, Y. & Lin, C. P53 codon 11, 72, and 248 gene polymorphisms in endometriosis. *Int. J. Biol. Sci.* **2**, 188–193 (2006).

17. Mcgregor, J. M. *et al.* Relationship Between p53 Codon 72 Polymorphism and Susceptibility to Sunburn and Skin Cancer. *J. Invest. Dermatol.* **119**, 84–90 (2002).

18. Huszno, J. & Grzybowska, E. W. A. TP53 mutations and SNPs as prognostic and predictive factors in patients with breast cancer (Review). *Oncol. Lett.* **16**, 34–40 (2018).

19. Garima, Pandey, S., Saxena, L. K., K, P. A. & Patel, N. The Role of p53 Gene in Cervical Carcinogenesis. *J. Obstet. Gynecol. India* **66**, 383–388 (2016).

20. Stiasny, A. *et al.* The involvement of E6, p53, p16, MDM2 and Gal‐3 in the clinical outcome of patients with cervical cancer. *Oncol. Lett.* **14**, 4467–4476 (2017).

21. Olivier, M., Hollstein, M. & Hainaut, P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harb Perspect Biol* **2:a001008**, 1–17 (2010).

22. Tommasino, M. *et al.* The role of TP53 in Cervical carcinogenesis. *Hum. Mutat.* **21**, 307–12 (2003).

23. Freier, C. *et al.* Immunohistochemical Evaluation of the Role of p53 Mutation in Cervical Cancer : Ser-20 p53 -Mutant Correlates with Better Prognosis. *Anticancer Res.* **36**, 3131–3138 (2016).

24. GenBank. P53 Protein Reference Sequence. *NP_001119584.1* (2020). Available at: https://ncbi.nim.nih.gov/protein/187830777#comment_NP_001119584.1. (Accessed: 27th June 2020)

25. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

26. Sievers, F. & Higgins, D. G. Clustal Omega, accurate alignment of very large numbers of sequences, Multiple Sequence Alignment Methods. *Springer* 105–116 (2014).

27. Blandino, G. & Agostino, S. Di. New therapeutic strategies to treat human cancers expressing mutant p53 proteins. *J. Exp. Clin. Cancer Res.* **37**, 1–13 (2018).

28. Donehower, L. A. *et al.* Integrated Analysis of TP53 Gene and Pathway Alterations in The Cancer Genome Atlas. *Cell Rep.* **28**, 1370–1384.e5 (2019).

29. Calhoun, E. S. *et al.* Host Genetic Polymorphism Analysis in Cervical Cancer. *Clin. Chem.* **48**, 1218–1224 (2002).

30. Lin, H., Huang, C., Wu, W., Chang, L. & Lung, F. TP53 codon 72 Gene Polymorphism Paradox in Associated with Various Carcinoma Incidences , Invasiveness and Chemotherapy Responses. **4**, 248–254 (2008).

31. Kwong, A. *et al.* Full spectrum of TP53 Mutations in Hong Kong. *GenBank* (2018). Available at: ncbi.nlm.nih.gov/nuccore/MG595979.1.