In-sillico analysis and 2D structure prediction of protein; decoding the genomic aspect of *Withania coagulans Dunal*

Jawad Ali[1], Maliha Ghaffar[2], Samiya rehman[1], Fouzia Tanvir[3], Yasir Nawaz[4,*], Ayesha Aslam[1], Durre Nayab[5], Muhammad Sajid[6], Muhammad Arslan[7], Anam Abbas[2], Taimoor Nawaz[1], Muhammad Luqman[4]

1 Department of Biochemistry, University of Okara, Pakistan

2 Department of Biology, University of Okara, Pakistan

3 Department of Zoology, Faculty of Life Sciences, University of Okara, Pakistan

4 Jiangsu Key Laboratory for Microbes and Functional Genomics, School of Life Sciences, Nanjing Normal University, Nanjing 210023, China

5 Department of Plant Sciences, Quaid-e-Azam University, Islamabad, Pakistan

6 Department of Biotechnology, University of Okara, Pakistan

7 Department of Agricultural Sciences, Allama Iqbal Open University, Islamabad, Pakistan

Corresponding author:

Yasir Nawaz

Abstract

Withania coagulans L. Dunal, member of Solonaceae family is a valuable plant holding a key place in Ayurvedic medicines. Presence of biomedical compounds as secondary metabolites imparts anti-microbial, anti-inflammatory, anti-oxidant and anti-cancerous properties to the *W. Coagualans Dunal.* Various analysis was performed to evaluate the genomic aspects of the *Withania coagulans Dunal.* In this study the *W. coagulans Dunal* genes were selected and their genomic structure, evolutionary analysis, gene structure analysis, motif enrichment analysis and secondary structure prediction was being performed to decode the genomic aspect. All genes were divided in 4 families based on their evolutionary relationship, Family 1,2,3 and 4 contain 5,3,1 and 6 genes respectively. Gene structure analysis confirmed that four genes include only CDS sequences, one contains two introns, and six contain both introns and CDS. Meme Suite predicted the existence of five motifs. Out of the five motifs anticipated, two were not predicted, while three motifs with 50, 30, and 50 amino acids were predicted, with the description Ribosomal protein uS12. With 41% Predominant structure of accD protein was Random coil.

Majority of amino-acids inthe accD form extracellular portion of the protein. These findings will help the scientists to do novel work on the genomic aspects of Withania coagulans Dunal in future.

Keywords: accD, 2D structure, in sillico analysis, proteomic analysis

Introduction

Reactive oxygen species (ROS) damage living organisms. Hydrogen peroxide (H2O2), free radicals like the hydroxyl radical (OH) and superoxide anion (O2) are among the reactive oxygen species produced in cells (Madkour, 2019). Oxidants initiate chemical chain reactions that damage DNA or proteins (A. Gupta, 2023). Imperfect DNA repair mechanisms damage the DNA and ultimately lead to cancer (Valko et al., 2016), whereas protein damage causes enzyme inhibition, denaturation, and degradation (Berlett & Stadtman, 1997).

Presence of antioxidant molecules in food systems decreases the harmful effects produced by the Reactive oxygen species (ROS) in the human body (Cakmakci et al., 2014). Antioxidants Molecules are defined as molecules that inhibit the oxidation of other molecules or they can scavenge the ROS directly or indirectly (Lü et al.,2002.). Superoxide dismutase enzymes (SODs) glutathione peroxidase (GPx), and glutathione Reductase (GRd, catalase (CAT) are the antioxidant enzymes that exhibit the highest antioxidant Defense effectiveness (Balick & Cox, 2020). Humans have tried to cure ailments using nature since ancient times (Jablonska et al., 2015).

Globally medicinal herbs trade is worth over USD 100 billion per year. *W. Coagulans* is a significant medicinal plant found in, Pakistan (Gilani et al., 2009), Iran (Lateef, T., Qureshi, S.A., 2020), Nepal, Afghanistan, and dried parts of India. *Withania coagulans* are generally found as a straight shrub 60-120cm in height. The leaves of the plant are covered by the tomentum on both sides (Chandra Gupta et al., 2012). The ripe fruits are used for wound healing (Melguizo-rodríguez et al., 2021) while dry fruits of *W. coagulans* are traditionally used as anti-diabetic (Ram et al., 2021) anti-microbial (Peerzade et al., 2018), diabetic treatment, hepatoprotective, (Lateef, T., Qureshi, S.A., 2020) anti-oxidant (Kumar Keshari et al., 2018), anti-depressant (Kamdi et al., 2020). Seeds have anti-inflammatory, diuretic, and ophthalmia-curing properties, while flower buds have anthelmintic properties(Maurya, 2010) (Mudassir et al., 2018). In this latest study, we performed in sillico analysis and isolated several genes from

NCBI related to *Withania coagulans Dunal*, which were then examined for evolutionary history, motif analysis, and gene structure. The secondary structure of a given protein was then predicted using several techniques, as well as its hydrophobicity and hydrophilicity, structural composition, extracellular and membrane interactions, and genomic composition were determined.

Material and Methodology

Databases

Genomic, Protein sequence, Complementary DNA Sequence (CDS) for the genes of W. coagulans plant were identified and retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov).

Identification of gene family members

Along with the Genomic, Complementary DNA Sequence and protein sequence for the genes, we also retrieved CDS length, gDNA length, protein length, CD number and start and end position of gene on chromosome using NCBI (https://www.ncbi.nlm.nih.gov) (Pruitt et al., 2007).

Gene structure visualization and detection of conserved motif

Motif analysis was performed using the online tool meme suite 5.4.1 (https://meme-suite.org/meme/index.html) (Bailey & Elkan, n.d.). Gene structure Display was performed by using the online gene structure visualization server (http://gsds.gao-lab.org). Gene structure visualization helps us visualize the distribution of introns and exons inside the gene using coding DNA sequence and DNA sequence of candidate genes (Hu et al., 2015).

Phylogenetic tree and multiple sequence alignment

For understanding the evolutionary relationship among the genes of interest, protein sequences of *W. coagulans* plant were aligned using the Clustal W Program and then evolutionary tree was conducted in Molecular Evolutionary Genetic Analysis X (MEGA X) (https://www.megasoftware.net) (Tamura et al., 2021).

Determination of physiochemical properties

Physiochemical Properties of Proteins such as molecular weight, isoelectric points, and Grand average of hydropathy, aliphatic index and instability index of protein was determined using ProtParam tool Online. ProtParam is online tool (https://web.expasy.org/protparam/) that provides us the information about the proteins stored in the TrEMBLE and Swiss Prot.

Protein primary and secondary structure determination

Protein Primary structure was obtained from the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov).

Protein structure prediction

Secondary structure of proteins and enzymes were predicted by the software PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/). PSI-blast based secondary structure Prediction (PSIPRED) uses the artificial neural network to predict the secondary Structures of proteins using the protein primary sequence. It predicts the 3-units of protein secondary structure Alpha helixes beta sheets and coils (Jones, 1999). Percentage of various secondary structures was obtained by the Self-Optimized Method with Alignment (SOPMA) tool Online (https://npsa-prabi.ibcp.fr/NPSA/npsa_sopma.html) (Geourjon & Deléage, 1995).

Results and Discussion

Phylogenetic Analysis

To understand the Evolutionary history of Genes, we aligned the 12 genes using Clustal W (https://www.genome.jp/tools-bin/clustalw) to construct the evolutionary tree using molecular evolutionary genetic analysis X (MEGA X). The Neighbor-Joining method was used for inferring the evolutionary history (Tamura et al., 2021). Evolutionary analysis is given below in Figure. 1.

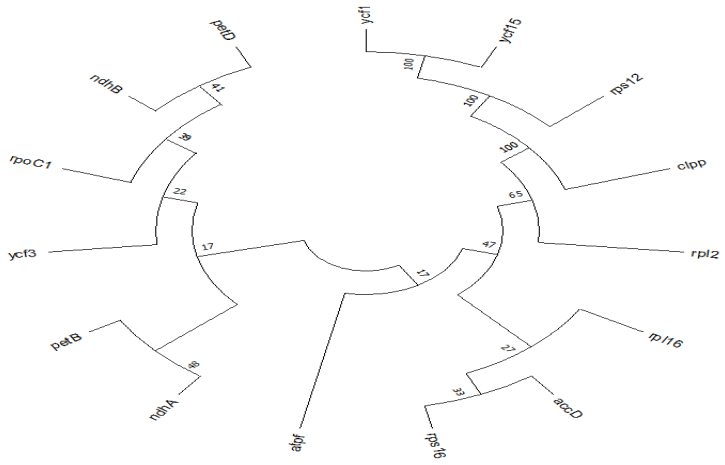


Figure 1: Evolutionary Tree of selected Genes in *W. coagulans*

Gene Structure Analysis

During the evolutionary process introns and exons can associate with the gene function that makes them significant. So, we have compared the coding DNA sequence and genomic DNA

sequences of genes associated with the antioxidant activity *W. Coagulans* using gene structure display server (GSDS) (http://gsds.gao-lab.org) to examine the diversity of intron/Exon in 12 genes associated with the antioxidant potential (Hu et al., 2015) as shown in the Figure 2.
The GSDS results shows that 5 out of 12 genes have no introns insertions while 6 out of 12 genes have just single intron insertion and ycf3 have the 2 intron insertions, the highest number of intron insertions in given set of genes.

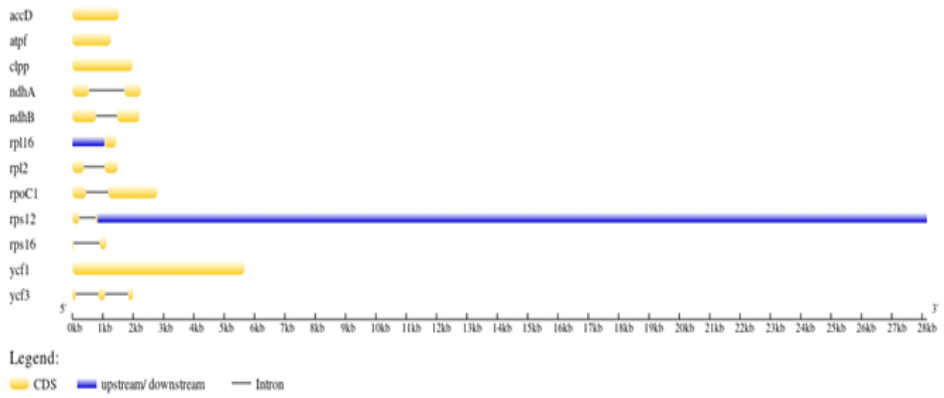


Figure 2: Out of the 12 genes 4 genes contain just CDS sequences and just 1 gene contain 2 introns while 6 genes consist of introns and CDS both a time.

Motif Enrichment Analysis

We have found out the statistically significant conserved motifs inside the encoded protein by using the online tool meme suite 5.4.1 (Figure 3).

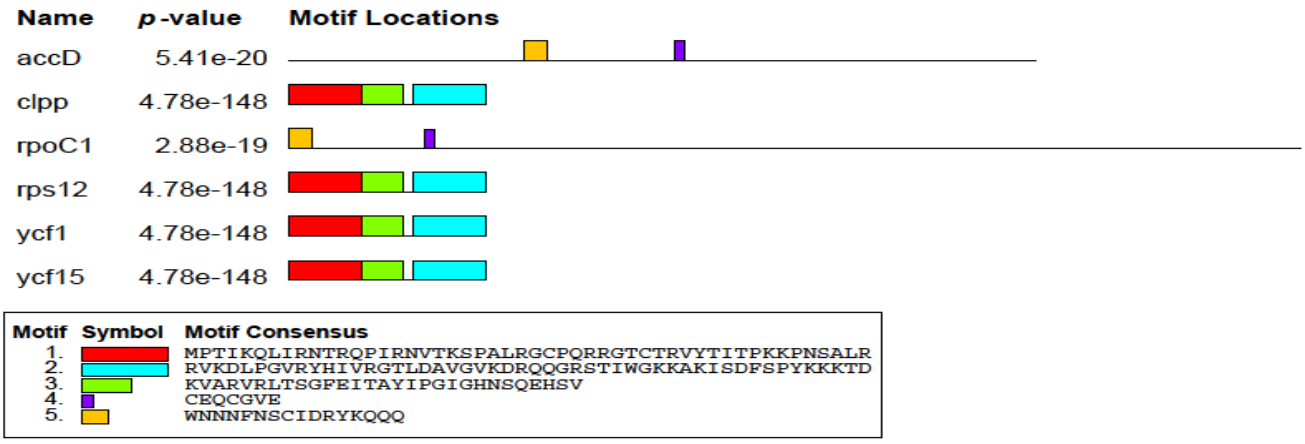


Figure 3: Location of amino-acid sequence of motifs

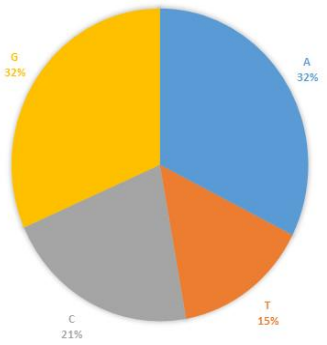
These motifs were assigned based on their order in MEME, corresponding amino acids motif sequence is given in Table 1. Motif Analysis was carried out using Meme Suite. Meme Suite predicted the presence of 5 motifs. Out of 5 motifs predicted 2 motifs were not being predicted while 3 motifs with 50, 30 and 50 Amino-acids were being predicted having description Ribosomal protein uS12.

Table 1: Indicated the motif analysis

#	Sequence	Amino Acids	Description
1	RVKDLPGVRYHIVRGTLDAVGVKDRQQGRSTIWG KKAKISDFSPYKKKTD	50	Ribosomal protein uS12
2	KVARVRLTSGVFEITAYIPGIGHNSQEHSV	30	Ribosomal protein uS12
3	CEQCGVE	8	None predicted
4	WNNNFNSCIDRYKQQQ	16	None predicted
5	MPTIQKLIRNTRQPIRNVTKSPALRGCPQRRGTCTR VYTITPKKPNSALR	50	Ribosomal protein uS12

Secondary structure of accD protein

Secondary structure analysis predicted by the bioinformatics analysis has confirmed that protein has 172 alpha helix (33.59%), 95 extended strands (18.55%), and Random coils 35 Beta turns (6.84%) and Zero percent Bend regions (Geourjon & Deléage, 1995) as given below in Figure 4b and genomic base content is also given in the Figure 4a.



a) Genomic Base content in the accD gene

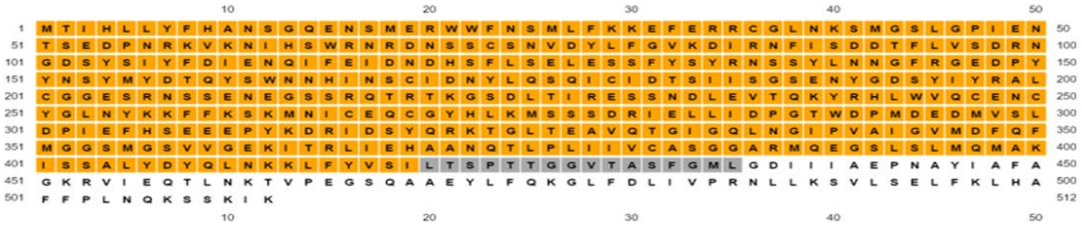
```
SOPMA :
    Alpha helix          (Hh) :     172 is   33.59%
    3₁₀  helix           (Gg) :       0 is    0.00%
    Pi helix             (Ii) :       0 is    0.00%
    Beta bridge          (Bb) :       0 is    0.00%
    Extended strand      (Ee) :      95 is   18.55%
    Beta turn            (Tt) :      35 is    6.84%
    Bend region          (Ss) :       0 is    0.00%
    Random coil          (Cc) :     210 is   41.02%
    Ambiguous states     (?)  :       0 is    0.00%
    Other states              :       0 is    0.00%
```

b) Secondary structures percentage

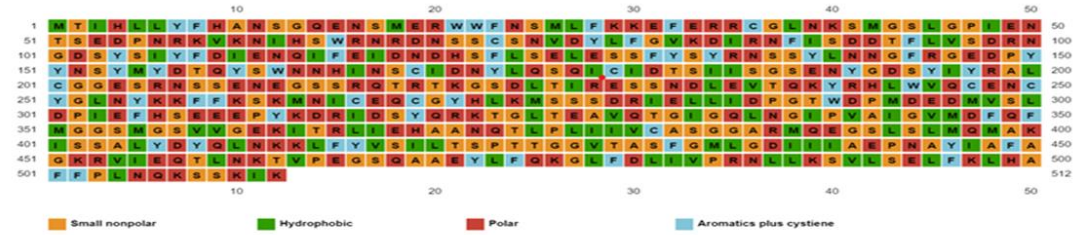
Figure 4: a) the genomic base content in the accD gene (b) the Secondary structures percentage

Predicting the nature of amino acid in accD protein

In this study, 16 out of 435 amino acids of protein form a trans-membrane helix while remaining form extracellular portion of protein that is dominant composition of protein shown in figure 5a. Amino acids can be hydrophilic or hydrophobic in nature, depending on their interaction with water. Whether they are attracted toward water or repelled by it. The various colors in the Figure 5b represent the nature of amino acids based on their interaction with water and polarity. Green in the figure represents hydrophobic amino acids, while red and yellow represents polar amino acids non-polar amino acids having hydrophobic and hydrophilic nature respectively.



a) The amino acids forming protein and their role



B) The hydrophilicity and hydrophobicity of accD amino acids

Figure 5: a) The amino acids forming protein and their role, b) The hydrophilicity and hydrophobicity of accD Amino acids

Conclusion

It is concluded that, *W. coagulans* is a highly beneficial medicinal plant having antioxidants for the treatment of many ailments. Antioxidants biosynthesis is catalyzed by the various enzymes encoded by the genes. Manipulating the biosynthesis of antioxidants involves manipulating the genes associated with them. The comprehensive data available can help us to understand and predict the structure and function of Genes.

Acknowledgment

Authors are thankful to the university department for helping during the completion of work.

Funding

N/A

Authors contribution

All authors contributed equally in the manuscript.

Conflict of interest

None

References:

Bailey, T. L., & Elkan, C. (n.d.). *FITTING A MIXTURE MODEL BY EXPECTATION MAXIMIZATION TO DISCOVER MOTIFS IN BIOPOLYMERS*.

Balick, M. J., & Cox, P. A. (2020). Plants, People, and Culture. *Plants, People, and Culture*. https://doi.org/10.1201/9781003049074/PLANTS-PEOPLE-CULTURE-MICHAEL-BALICK-PAUL-ALAN-COX

Berlett, B. S., & Stadtman, E. R. (1997). Protein oxidation in aging, disease, and oxidative stress. In *Journal of Biological Chemistry* (Vol. 272, Issue 33, pp. 20313–20316). https://doi.org/10.1074/jbc.272.33.20313

Cakmakci, S., Topdas, E. F., Ul C ß Akmakc ßı, S., Topdas, E. F., Kalın, P., Han, H., ßekerci, P. S., K€ Ose, L. P., Ilhami, & _, & Ulc ßin, G. (2014). Antioxidant capacity and functionality of oleaster (Elaeagnus angustifolia L.) flour and crust in a new kind of fruity ice cream. *Wiley Online LibraryS Çakmakçı, EF Topdaş, P Kalın, H Han, P Şekerci, L P. Köse, İ GülçinInternational Journal of Food Science & Technology, 2015•Wiley Online Library*, *50*(2), 472–481. https://doi.org/10.1111/ijfs.12637

Chandra Gupta Keshav Prasad Mishra Rajkiya Mahila Mahavidyalaya Aurai, P., affliated to Mahatma Gandhi Kashi Vidyapith, B., & Chandra Gupta, P. (2012). Withania coagulans Dunal-An overview. In *Article in International Journal of Pharmaceutical Sciences Review and Research* (Vol. 12, Issue 2). https://www.researchgate.net/publication/286747008

Geourjon, C., & Deléage, G. (1995). Sopma: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics*, *11*(6), 681–684. https://doi.org/10.1093/BIOINFORMATICS/11.6.681

Gilani, S. A., Kikuchi, A., & Watanabe, K. N. (2009). Genetic variation within and among fragmented populations of endangered medicinal plant, Withania coagulans (Solanaceae) from Pakistan and its implications for conservation. *African Journal of Biotechnology*, *8*(13), 2948–2958. http://www.academicjournals.org/AJB

Gupta, A. (2023). *Hydrogen peroxide-dependent and-independent DNA damage and repair*. https://www.ideals.illinois.edu/items/127538

Gupta, R., international, V. S.-T., & 2011, undefined. (n.d.). Ameliorative effects of Tinospora cordifolia root extract on histopathological and biochemical changes induced by aflatoxin-B1 in mice kidney. *Ncbi.Nlm.Nih.GovR Gupta, V SharmaToxicology International*,

2011•ncbi.Nlm.Nih.Gov. Retrieved September 22, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3183631/

Hu, B., Jin, J., Guo, A., Zhang, H., Luo, J., Bioinformatics, G. G.-, & 2015, undefined. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Academic.Oup.ComB Hu, J Jin, AY Guo, H Zhang, J Luo, G GaoBioinformatics, 2015•academic.Oup.Com*, *31*(8), 1296–1297. https://doi.org/10.1093/bioinformatics/btu817

Jablonska, E., Gromadzinska, J., Peplonska, B., Fendler, W., Reszka, E., Krol, M. B., Wieczorek, E., Bukowska, A., Gresner, P., Galicki, M., Zambrano Quispe, O., Morawiec, Z., & Wasowicz, W. (2015). Lipid peroxidation and glutathione peroxidase activity relationship in breast cancer depends on functional polymorphism of GPX1. *BMC Cancer*, *15*(1). https://doi.org/10.1186/s12885-015-1680-4

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, *292*(2), 195–202. https://doi.org/10.1006/JMBI.1999.3091

Kamdi, A. S., Gosavi, D. D., Kalambe, S. M., & Bohra, P. N. (2020). The motor coordination activity of alcoholic extract of Withania coagulans fruits in Swiss albino mice by rota rod test. *Indian Journal of Pharmacy and Pharmacology*, *7*(2), 73–76. https://doi.org/10.18231/j.ijpp.2020.015

Kumar Keshari, A., Srivastava, A., Upadhayaya Associate Professor, M., Srivastava Associate Professor, R., Ragini Srivastava Associate Professor, C., Upadhayaya, M., & Srivastava, R. (2018). Antioxidants and free radicals scavenging activity of Medicinal Plants. *~ 1499 ~ Journal of Pharmacognosy and Phytochemistry*, *7*(3).

Lateef, T., Qureshi, S.A., 2020. Centratherum anthelmint... - Google Scholar. (n.d.). Retrieved September 22, 2024, from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Lateef%2C+T.%2C+Qureshi%2C+S.A.%2C+2020.+Centratherum++anthelminticum++and++Withania+coagulans++improves+lipid+profile+and+oxidative+stress+in++triton+X-100+induced+hyperlipidemic+rabbits.++Group%2C+1%3A+1%E2%80%9322&btnG=

Lü, J., Lin, P., Yao, Q., and, C. C.-J. of cellular, & 2010, undefined. (n.d.). Chemical and molecular mechanisms of antioxidants: experimental approaches and model systems. *Wiley Online LibraryJM Lü, PH Lin, Q Yao, C ChenJournal of Cellular and Molecular Medicine, 2010•Wiley*

Online Library. Retrieved September 22, 2024, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1582-4934.2009.00897.x

Madkour, L. H. (2019). Citation: Loutfy H Madkour. Function of Reactive Oxygen Species (ROS) Inside the Living Organisms and Sources of Oxidants Function of Reactive Oxygen Species (ROS) Inside the. In *Article in Journal of Global Pharma Technology*. https://chembiopublishers.com/submit-manuscript.php

Maurya, R. (2010). Chemistry and pharmacology of Withania coagulans: an Ayurvedic remedy. *JPP*, *62*, 153–160. https://doi.org/10.1211/jpp/62.02.0001

Melguizo-rodríguez, L., de Luna-Bertos, E., Ramos-torrecillas, J., Illescas-montesa, R., Costela-ruiz, V. J., & García-martínez, O. (2021). Potential effects of phenolic compounds that can be found in olive oil on wound healing. In *Foods* (Vol. 10, Issue 7). MDPI AG. https://doi.org/10.3390/foods10071642

Mudassir, H. A., Nazim, K., Khan, U. A., & Qureshi, S. A. (2018). COMPARATIVE EVALUATION OF HYPOGLYCEMIC ACTIVITY, TRACE MINERALS AND PHYTOCHEMICAL CONTENTS OF SOME POTENTIAL MEDICINAL PLANT EXTRACTS. In *INT. J. BIOL. BIOTECH* (Vol. 15, Issue 1).

Peerzade, N., Sayed, N., & Das, N. (2018). Antimicrobial and phytochemical screening of methanolic fruit extract of Withania coagulans L. Dunal for evaluating the antidiabetic activity. *The Pharma Innovation Journal*, *7*(1), 197–204. www.thepharmajournal.com

Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2007). NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, *35*(SUPPL. 1). https://doi.org/10.1093/nar/gkl842

Ram, H., Kumar, P., Purohit, A., Kashyap, P., Kumar, S., Kumar, S., Singh, G., Alqarawi, A. A., Hashem, A., Abd-Allah, E. F., Al-Arjani, A. B. F., & Singh, B. P. (2021). Improvements in HOMA indices and pancreatic endocrinal tissues in type 2-diabetic rats by DPP-4 inhibition and antioxidant potential of an ethanol fruit extract of Withania coagulans. *Nutrition and Metabolism*, *18*(1). https://doi.org/10.1186/s12986-021-00547-2

Sofowora, A., Ogunbodede, E., & Onayade, A. (2013). The role and place of medicinal plants in the strategies for disease prevention. *African Journal of Traditional, Complementary, and Alternative Medicines : AJTCAM / African Networks on Ethnomedicines*, *10*(5), 210–229. https://doi.org/10.4314/AJTCAM.V10I5.2

Tamura, K., Stecher, G., evolution, S. K.-M. biology and, & 2021, undefined. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Academic.Oup.ComK Tamura, G Stecher, S KumarMolecular Biology and Evolution, 2021•academic.Oup.Com*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Valko, M., Jomova, K., Rhodes, C. J., Kuča, K., & Musílek, K. (2016). Redox- and non-redox-metal-induced formation of free radicals and their role in human disease. In *Archives of Toxicology* (Vol. 90, Issue 1, pp. 1–37). Springer Verlag. https://doi.org/10.1007/s00204-015-1579-5