**Supplementary Information**

**Table S1**. Summary Statistics from GWAS result using 154 SNPs

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SNP | Chromosome | Position | P value | MAF | Effect |
| S1\_68992917 | 8 | 68992917 | 6.60E-14 | 0.1529 | 0.6401 |
| S1\_53208128 | 4 | 53208128 | 3.38E-07 | 0.0537 | -0.4741 |
| S1\_24909781 | 6 | 24909781 | 2.80E-05 | 0.0868 | -0.4375 |
| S1\_23794794 | 1 | 23794794 | 0.00033228 | 0.0744 | -0.2892 |
| S1\_94173819 | 3 | 94173819 | 0.00037769 | 0.0785 | -0.2982 |

SNP – Single Nucleotide Polymorphism

MAF – Minor Allele Frequency

A picture containing text, screenshot, graphic design, logo

Description automatically generated

**Figure S1: Outline of the proteome-wide association study**. Genetic variants represent the variants called through exome capture sequencing. Protein features used as local distance difference test generated using the alphafold through 3D structure of proteins.

A picture containing diagram, line, plan, parallel

Description automatically generated

**Figure S2. A Peptide bond (orange lines) links the two amino acids together in a protein molecule**. Here R represents the alkyl group which can have a different number of residues based on the amino acid type. These residues determine the property of the amino acids to which it is linked.

A picture containing text, screenshot, map

Description automatically generated

**Figure S4.3. 3D Structure of the gene MsG0280007716.01.T01.** The confidence scores are color coded from high confidence (dark blue) to low confidence (orange). The prediction accuracy of this protein is high as most of the regions are predicted with high confidence.

A picture containing text, diagram, screenshot, plot

Description automatically generated

**Figure S4.4**. **Distribution of mutation effect prediction from MutPred2**. The mutation's effect value due to the substitution of amino acids lies between 0 and 1. Any prediction effect above 0.5 is considered to have a deleterious effect on the phenotype. While the x-axis represents the effect of amino acid substitution on phenotype and the y-axis their frequency,  the red line represents the number of predicted values that are at or above the threshold used for the selection of deleterious mutation for association study.