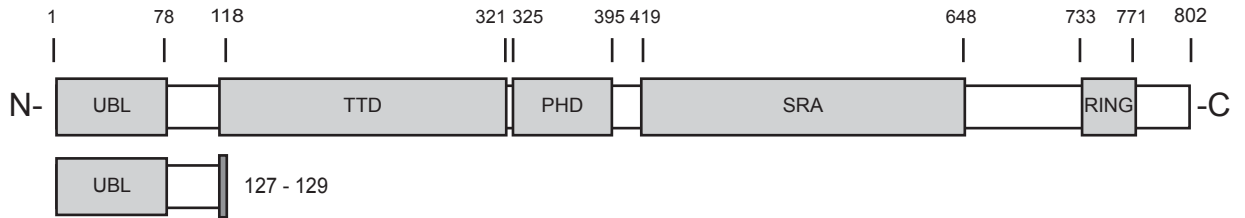


Figure S1

A



B

CGTGCCCGTCTTATTGATCCTGGCTTTGGAATATATAAGGTATG Genomic Sequence
 CGTGCCCGTCTTATTG-TCCTGGCTTTGGAATATATAAGGTATG Allele 1 |Times detected: 2
 CGTGCCCGTCTTATTGA--CTGGCTTTGGAATATATAAGGTATG Allele 2 |Times detected: 2
 CGTGCCCGTCTTATTGA----GGCTTTGGAATATATAAGGTATG Allele 3 |Times detected: 3
 CGTGCCCGTCTTATT--TCCTGGCTTTGGAATATATAAGGTATG Allele 4 |Times detected: 3

118 - RLIDPGFGIYKVNEL---> Endogenous Protein Sequence
 118 - RLIVLALEYIR* - 129 Allele 1
 118 - RLIDWLWNI* - 127 Allele 2
 118 - RLIEALEYIR* - 128 Allele 3
 118 - RLISWLWNI* - 127 Allele 4

C

