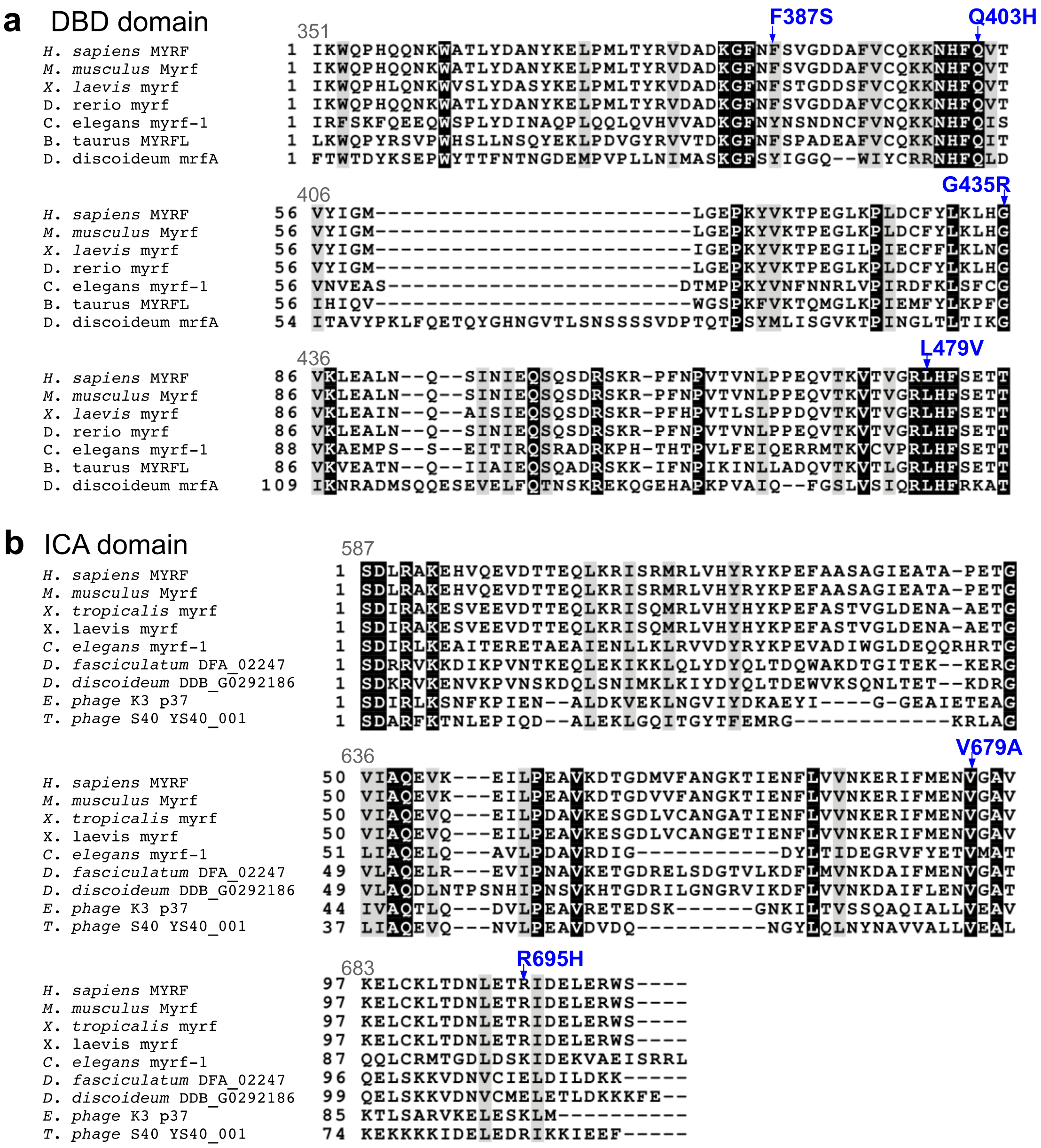
S3 Fig. Multiple sequence alignment (MSA) of MYRF protein domains.



a DBD domain

```
                       351
H. sapiens MYRF          1 IKWQPHQQNKWATLYDANYKELPMLTYRVDADKGFNFSVGDDAFVCQKKNHFQVT
M. musculus Myrf         1 IKWQPHQQNKWATLYDANYKELPMLTYRVDADKGFNFSVGDDAFVCQKKNHFQVT
X. laevis myrf           1 IKWQPHLQNKWVSLYDASYKELPMLTYKVDADKGFNFSTGDDSFVCQKKNHFQVT
D. rerio myrf            1 IKWQPHQQNKWATLYDANYKELPMLTYRVDADKGFNFSVGDDAFVCQKKNHFQVT
C. elegans myrf-1        1 IRFSKFQEEQWSPLYDINAQPLQQLQVHVVADKGFNYNSNDNCFVNQKKNHFQIS
B. taurus MYRFL          1 LKWQPYRSVPWHSLLNSQYEKLPDVGYRVVTDKGFNFSPADEAFVCQKKNHFQIT
D. discoideum mrfA       1 FTWTDYKSEPWYTTFNTNGDEMPVPLLNIMASKGFSYIGGQ--WIYCRRNHFQLD

                       406
H. sapiens MYRF         56 VYIGM-------------------------LGEPKYVKTPEGLKPLDCFYLKLHG
M. musculus Myrf        56 VYIGM-------------------------LGEPKYVKTPEGLKPLDCFYLKLHG
X. laevis myrf          56 VYIGM-------------------------IGEPKYVKTPEGILPIECFFLKLNG
D. rerio myrf           56 VYIGM-------------------------LGEPKYVKTPEGLKPLDCFYLKLHG
C. elegans myrf-1       56 VNVEAS-----------------------DTMPPKYVNFNNRLVPIRDFKLSFCG
B. taurus MYRFL         56 IHIQV-------------------------WGSPKFVKTQMGLKPIEMFYLKPFG
D. discoideum mrfA      54 ITAVYPKLFQETQYGHNGVTLSNSSSSVDPTQTPSYMLISGVKTPINGLTLTIKG

                       436
H. sapiens MYRF         86 VKLEALN--Q--SINIEQSQSDRSKR-PFNPVTVNLPPEQVTKVTVGRLHFSETT
M. musculus Myrf        86 VKLEALN--Q--SINIEQSQSDRSKR-PFNPVTVNLPPEQVTKVTVGRLHFSETT
X. laevis myrf          86 VKLEAIN--Q--AISIEQSQSDRSKR-PFHPVTLSLPPDQVTKVTVGRLHFSETT
D. rerio myrf           86 VKLEALN--Q--SINIEQSQSDRSKR-PFNPVTVNLPPEQVTKVTVGRLHFSETT
C. elegans myrf-1       88 VKAEMPS--S--EITIRQSRADRKPH-THTPVLFEIQERRMTKVCVPRLHFSETT
B. taurus MYRFL         86 VKVEATN--Q--IIAIEQSQADRSKK-IFNPIKINLLADQVTKVTLGRLHFSETT
D. discoideum mrfA     109 IKNRADMSQQESEVELFQTNSKREKQGEHAPKPVAIQ--FGSLVSIQRLHFRKAT
```

b ICA domain

```
                                 587
H. sapiens MYRF                    1 SDLRAKEHVQEVDTTEQLKRISRMRLVHYRYKPEFAASAGIEATA-PETG
M. musculus Myrf                   1 SDLRAKEHVQEVDTTEQLKRISRMRLVHYRYKPEFAASAGIEATA-PETG
X. tropicalis myrf                 1 SDIRAKESVEEVDTTEQLKRISQMRLVHYHYKPEFASTVGLDENA-AETG
X. laevis myrf                     1 SDIRAKESVEEVDTTEQLKRISQMRLVHYHYKPEFASTVGLDENA-AETG
C. elegans myrf-1                  1 SDIRLKEAITERETAEAIENLLKLRVVDYRYKPEVADIWGLDEQQRHRTG
D. fasciculatum DFA_02247          1 SDRRVKKDIKPVNTKEQLEKIKKLQLYDYQLTDQWAKDTGITEK--KERG
D. discoideum DDB_G0292186         1 SDKRVKENVKPVNSKDQLSNIMKLKIYDYQLTDEWVKSQNLTET--KDRG
E. phage K3 p37                    1 SDIRLKSNFKPIEN--ALDKVEKLNGVIYDKAEYI----G-GEAIETEAG
T. phage S40 YS40_001              1 SDARFKTNLEPIQD--ALEKLGQITGYTFEMRG------------KRLAG

                                 636
H. sapiens MYRF                   50 VIAQEVK---EILPEAVKDTGDMVFANGKTIENFLVVNKERIFMENVGAV
M. musculus Myrf                  50 VIAQEVK---EILPEAVKDTGDVVFANGKTIENFLVVNKERIFMENVGAV
X. tropicalis myrf                50 VIAQEVQ---EILPDAVKESGDLVCANGATIENFLVVNKERIFMENVGAV
X. laevis myrf                    50 VIAQEVQ---EILPEAVKESGDLVCANGETIENFLVVNKERIFMENVGAV
C. elegans myrf-1                 51 LIAQELQ---AVLPDAVRDIG-----------DYLTIDEGRVFYETVMAT
D. fasciculatum DFA_02247         49 VLAQELR---EVIPNAVKETGDRELSDGTVLKDFLMVNKDAIFMENVGAT
D. discoideum DDB_G0292186        49 VLAQDLNTPSNHIPNSVKHTGDRILGNGRVIKDFLVVNKDAIFLENVGAT
E. phage K3 p37                   44 IVAQTLQ---DVLPEAVRETEDSK------GNKILTVSSQAQIALLVEAV
T. phage S40 YS40_001             37 LIAQEVQ---NVLPEAVDVDQ----------NGYLQLNYNAVVALLVEAL

                                 683
H. sapiens MYRF                   97 KELCKLTDNLETRIDELERWS----
M. musculus Myrf                  97 KELCKLTDNLETRIDELERWS----
X. tropicalis myrf                97 KELCKLTDNLETRIDELERWS----
X. laevis myrf                    97 KELCKLTDNLETRIDELERWS----
C. elegans myrf-1                 87 QQLCRMTGDLDSKIDEKVAEISRRL
D. fasciculatum DFA_02247         96 QELSKKVDNVCIELDILDKK-----
D. discoideum DDB_G0292186        99 QELSKKVDNVCMELETLDKKKFE--
E. phage K3 p37                   85 KTLSARVKELESKLM----------
T. phage S40 YS40_001             74 KEKKKKIDELEDRIKKIEEF-----
```

Homologous protein sequences from UniProtKB/Swiss-Prot were searched by BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Selected ortholog sequences on DBD and ICA domains were aligned using Clustal Omega(Sievers, Wilm et al. 2011), and the resulting MSA were then visualized by BoxShade (https://embnet.vital-it.ch/software/BOX_form.html). Residues that are completely conserved are highlighted in black; changes to residues with similar biochemical properties are shaded in gray. Amino acid substitutions that result from *de novo* missense variants are shown above the MSA. Most substitutions affect highly conserved residues.