**Table S1.** Selected kafirin sequences from Uniprot.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Entry | Protein name | Length | Molecular weight (Da) | Full sequence |
| A9XEC0 | Alpha kafirin B1 | 267 | 28,883 | MAAKIFSLIMLLALFASAATA**TYIPQPHCSQISIAAPHPAYLPPVCANPMLQPYGLQQAIAASILQSSPLFIQQPSALLQQLSLVNLLAQSIRAQQLQQLVLPSINQVTLANLAAYSQQQQFIPFSQLAAVNPAAYLQLQQLQQFSQLAAASPAAFWPQQQLLPFYPQGVANAATLLQLQQLLPFNQLAAVNPAAYLQLQLLLQFSQLGAVSPATFWPQQQLLPFYPQGVANAATLFQQQQLLPFIQLALANPATSCQQPFIGGALF** |
| A9XEC1 | Alpha kafirin B3 | 267 | 28,970 | MAAKIFSLVMLLALFASAATA**TYIPQPHCSQISIAAPHPAYLPPVCANPMLQPYGLQQAIAASILQSSPLFIQQPSALLQQLSLVNLLAQSIRAQQLQQLVLPSIHQVTVANLAAYSQQQQFIPFSQLAAVNPAAYLQLQQLQQFSQLAAASPAAFWPQQQLLPFYPQGVANAATLLQLQQLLSFNQLAAVNPAAYLQLQQLLQFSRLGAVSPATFWPQQQLLRFYPQGVANAATLFQQQQLLPFIQLALANPATSCQQPFIGGALF** |
| Q9XE78 | Alpha-kafirin | 184 | 19,969 | MATMIFAFLALLALSVSTTTA**VIIPQCSLAPNAIISQFLPPFTPMGFEHPALQAYRLQQALANSILQQPFAQLQQQSSAHLTVQTIAAQQQQQFLLALKQLALANPVAYLQQQLLASNPLALVNNAAYLQQQLQQVLPVISQVALANPAAYLQQQQLAYNPLVTANAAAYLQQQQLQQILPALS** |
| D5L0X8 | Beta-kafirin | 193 | 20,840 | MKMVIVLAVCLALSAASASA**LQMPGMGLQDLYGAGALMTMMGAGGGLYPCAEYLRQPQCSPVAAPFYALREQTMWQPNFICQPLRQQCCQQMRMMDMQSRCQAMCGVVQSVVQQLQMTMQLQGVAAAASSLLYQPALVQQWQQLLPAAQALTPLAMAVAQVAQNMPAMCGLYQLPSYCTTPCATSAAIPPYYY** |
| Q6Q299 | Gamma kafirin | 186 | 19,590 | MKVLLVALALLAPAASA**ASTLTTGGCGCQTPHLPPPPVHLPPPVHLPPPVHLPPPVHVPPPPPQCHPHPTLPPHPHPCPTYPPHPSPCHPGHPGSCGVGGGPVTPPILGQCIEFLRHQCSPAATPYCSPQCQALRQQCCQQLRQVEPLHRYQAIFGVVLQSIQQQQPQGQSSPLPALMAAQIAQQL** |
| G3FMW5 | Gamma-kafirin | 212 | 22,112 | MKVLLVALALLALAASAAS**TLTTGGCGCQTPHLPPPPVHLPPPVHLPPPVHLPPPVHVPPPPPQCHPHPTLPPHPHPCPTYPPYPSPCHPGHPGSCGVGGGPVTPPILGQCIEFLRHQCSPAATPYCSPQCQALRQQCCQQLRQVEPLHRYQAIFGVVLQSIQQQQPQGQSSPLPALMAAQIAQQLTAMCGLGVGQPSPCASCSPFAGGVHY** |
| D5L0Y6 | Delta-kafirin | 147 | 16,490 | MAAKMFALFALLALCASATSA**THIPGHFPLVMPLGTMNPCTQYCMMQQRFARLLAWPIPMLQQLLALPLQPAYQTPMTMPNMMPPMTMRPTMMPPMTMMPSMMSLPQCHCDAISQIMQQQQLPFMFNPTAMAIPPMFLQQPFVSSAF** |
| C5Z299 | Delta-kafirin | 147 | 16,516 | MAAKMFALFALLALCASATSA**THIPGHLPLVMPLGTMNPCTQYCMMQQRFARLLAWPIPMLQQLLALPLQPAYQTPMMMPNMMPPMMMRPTMMPPMTMMPSMMSLPQCHCDAISQIMQQQQLPFMFNPTAMAIPPMFLQQPFVSSAF** |

Note: The bold letters are the sequence we used in the hydrolysis simulation and the unbold letters are signal peptides.