Supplementary information:

I.) Methods:

Data mining strategy.

Regular BLAST sequence similarity search ignores the position specific information about protein domains. Since the overall sequence similarity is expected to be very low for RHG proteins and the IBM motif is very short, regular BLAST analysis will not be sensitive and selective enough to identify the potential orthologs. Rather, a model has to be built for the IBM motif to specifically search for its presence in genomic sequences.

In order to build the model, predated sequences of orthologs of *Drosophila* IAP-antagonists (Reaper, Hid, Grim, and Sickle) from divergent *Drosophila* species such as *D. virilis* were collected (Supplementary information II). Models and scoring matrices for the two motifs were built with the MEME program (Bailey and Elkan, 1994), and tested empirically by searching the Swiss-Pro annotated protein data set (Supplementary information III). Genomic sequences for *Anopheles* and *Aedes* were obtained from Ensembl and TIGR, respectively. A motif search program implemented in C was customized to search for the two motifs in the genomic as well as EST sequences from mosquito genomes (Zhou et al., 1999). To search for potential Reaper/Grim-like IBM containing protein, the program was customized to report only qualifying matches that immediately follow a methionine and do not have any stop codon within 20 amino acids after the IBM. DNA sequences for the genomic region encompassing the potential motif matches were then analyzed with gene structure predicting programs (Burge and Karlin, 1997) and used to identify corresponding ESTs.
Cloning mx.

Genomic DNA flanking the predicted ORF of mx was first amplified from Aedes albopictus and Anopheles gambiae genomic DNA and tested for cell killing ability in S2 cells. Primers used for amplifying mx from Aedes are 5’GACTGTAGAAACATGGCAAT and 3’ATACACCTTTCGCAGCAAGT; for Anopheles, they are 5’AACATGGCAATCGCATTCTACAT and 3’ATGTCTTGGGGTTTCGAT. An intron-less cDNA was then obtained by RT-PCR using RNA extracted from UV-irradiated C6/36 mosquito cells. Additional mx cDNA clones were isolated from the Aedes aegypti cDNA libraries. For function/structure studies, Mx cDNA from Aedes albopictus was used as the template for generating Mx(-IBM). Wild type or mutated ORFs were subcloned into the pIE and pRK5 vector. All constructs were verified by sequencing.

Cell culture and in vitro cell death assay.

Culture of the S2 cell line and cell death assay was performed essentially as previously described (Jones et al., 2000). For each test, a total of 1.0 ug of DNA mixed with 8ug of liposome (Cellfectin(Invitrogen)) in serum free media was distributed into 2 wells in a 24-well plate. This included 0.1 ug pIE-lacZ and 0.9 ug of the test DNA sample or a combination of samples in the pIE vector. Intact pIE vector was used as the control. The transfection lasted for 4 hours and was stopped by replacing the transfection mix with culture media supplemented with 10% FBS. At 20 hours post-transfection, cells were fixed and stained with X-Gal/IPTG. Blue cells were counted to calculate the percentage of cell survival. All data reported here were repeated at least three times.
C6/36 cells were cultured with MEM plus 10% FBS and maintained in 29°C incubator with 5% CO2. Transfection of DNA to the C6/36 cells and cell death assays were conducted essentially the same as in the S2 cells except GeneJuice (Novagen) is used instead of Cellfectin.

**UV treatment of mosquito cells and gene expression assay.**

For UV (254nm) treatment, the media was removed briefly during irradiation using the UV Stratalinker (Stratagene). Fresh media was added back immediately afterwards to both treated and control samples. At defined time points, cells were lysed on plate and total RNA was extracted with RNeasy(Qiagen), kept at –80°C before being processed for RT-PCR or real-time RT-PCR assay. For quantitative RT-PCR, gene-specific primer sets were designed for the mosquito gapdh (5’:GGGCTTCCTGTACCACCAACTGC; 3’:CCATCACGCCACAGCTTACCAG) as well as mx (5’:CGAGAAGCGGTATCAGTTGCAA; 3’:GGTTGTGGTGGATGCTGACTTG) by GeneDirect (Princeton Junction, NJ).

**In vivo killing analysis.**

The ORF of mx (from *Aedes albopictus*) was cloned into the PUAST vector. Transgenic fly strains were generated via embryo injection as described before (Zhou et al., 1997). To collect embryos, virgin females carrying P{UAS-mx} were crossed with males carrying P52Gal4::P{UAS-lacZ}(Wing et al., 1998) (obtained from J. Nambu). Embryos were fixed and processed for anti-βGal staining as described previously (Zhou et al., 1995).

**RNA interference assay in C6/36 cells.**
Double stranded RNA for mx or GFP were synthesized according to the method described by Clemens et al. (Clemens et al., 2000). The primer pair used for mx is T7mx5:

GAATTAATACGACTCACTATAGGGAGAGGCACAGCAGCAACAACAAA and T7mx3:

GAATTAATACGACTCACTATAGGGAGAATACACCTTTCGCAGCAAGTA, for GFP it is T7GFP5:

GAATTAATACGACTCACTATAGGGAGAAGCTGGACGGCGACGTAAA and T7GFP3:

GAATTAATACGACTCACTATAGGGAGATTGCCGTCGTCCTTGAAGAA.

Transcription, annealing, and purification of double strand RNA were carried out using the MEGAscript RNAi kit from Ambion.

To test the effect of mx RNAi on UV-induced cell death, several methods of introducing dsRNA were tested. The direct application method that worked for S2 cells (Clemens et al., 2000) did not work for C6/36 cells. Also, we found that transfection immediately following UV irradiation was extremely toxic to cells and had very low efficiency. So we adopted the strategy of transfecting dsRNA prior to UV irradiation.

Transfection of dsRNA to C6/36 cells was carried out using Cellfectin (Invitrogen). 1.8ug of dsRNA for mx or GFP and 0.2ug of pIE-lacZ were mixed with 16ug liposome in serum free MEM 15 minutes before being applied to a column of four wells in a 24 well plate. The transfections were stopped 4 hours later by replacing the transfection mixture with FBS-supplemented MEM. The cells were then allowed to recover in FBS-supplemented media for 2 hours before UV treatment. For UV irradiation, media was
removed from all wells. While the control wells were shielded with two layers of aluminum foil, other wells were subjected to various dosage of UV in the UV Stratalinker (Stratagene). Immediately following UV treatment, FBS-supplemented media were reintroduced to the wells. The cells were fixed and stained with Xgal/IPTG 16-18 hrs after UV treatment. Blue cells were counted for each well, and the survival rate was calculated using the corresponding zero-exposure (shielded control) well as the baseline.

Co-immunoprecipitation and Western blot analysis.

To measure protein stability, Mx or Mx(-IBM) was fused at the C-terminus with either a HA or FLAG tag in the pRK5 vector. NCI-H1299 lung cancer cells were transfected using lipofectamine, and protein expression was determined 48 hours post transfection by Western blot analysis as described previously (Xiao and Lang, 2000). For co-immunoprecipitation, 1 µg of the HA-Diap expression construct was co-transfected with 3 µg of plasmids expressing either Mx-FLAG, or Mx(IBM)-FLAG, or empty vector. At 48 h post transfection, cells were lysed in a modified RIPA buffer (Xiao and Lang, 2000). Equal amounts of whole-cell lysates were immunoprecipitated with a mouse monoclonal anti-HA antibody (12CA5). The immunoprecipitate was resolved using SDS-PAGE, transferred to membrane, followed by immunoblotting.

II.) RHG sequences used to build the model (in FASTA format). Orthologs of Drosophila IAP-antagonists (Reaper, Hid, Grim, and Sickle) from divergent Drosophila species such as D. virilis were identified using batch BLAST analysis (McGinnis and Madden, 2004). The protein sequences were then predicted using GeneWise (Birney et al., 2004). For details, please refer to Zhou (2005) or corresponding with the author.

>Rpr_D.virilis
MAVAFYIPDQAAMLRKAREQEKQLLRLQRWQMMLNTIRQLMDTAGQ
CLQRAGRKGKWRKPQS
Zhou et al. IAP-antagonist in mosquitoes

> Hid_D.viri
MAVPFYLPQEGTDDIAASSSSGTGVSSSSSSSSSSSSASSSSGASSSGVS
SSGASGASSVASSQSPMTTSTATQTPMQLQPLPMDQVLYIALCEWYQQH
QRHYIAPEQIFQYYPFPSPSSCSAQSPSYTGGDVFAPAGHVPRTPTRTSI
SFAAGEELNFRQPYPAPPTPQPMPQQPAPMQAPSMHSYPQAASHLL
QHPASYGHPVQGAYATTPPTPNTANASTTSTVAAAASFGRHGYG
PAMASTPLAPTGPKMLQRQSDAARKRTLSTGEDEREYQSDHETSWE
FDERYDNFATGERLRQEFSGRPRKKTSNPGRSNNTHLCGGSDDANE
NSYSPQRERTAKEKTSFTWPTVFTVFLAMCGFFAAR

> Grim_D.viri
MAIAYFIPDQAPQLLARSQQAQQAGSVDGSRPNPRQNESRQQQQHQQQ
QQQQRQFRHRGNTQFRANISVPLGREQQGAMMSFECWELQUALCYAL
RLYSPNPRQQAPRTVQISFEISSNTSDADADGVGDVDATAK

> Skl_D.viri
MAIAYFEDDEADAVASAAAAATTLESSQITDGDQVDGPAYASDAGFDQVDYE
EPSGATNSQTTTSTETEAPPSTEEOQKLAWKMLALAMCKVLQFYLQHKKQ
QQAQGKSTKLATATVQIAQ

> Rpr_D.yaku
MAVAFIPDQATITHQILRLRESQWRFLATVQLTQLRQYSCT
PKTGRCGKRYKPSQ

> Hid_D.yaku
MAVPFYLPSEGGADDVASSSSGASGNSSPHNHPHPLASSASSSSSVSSGVSSAS
ASSASSASSSSASSGGASSAASQSPNNTTSSATQTPMQLQPLPMDQVLYAL
YEVRMSMQQQASPQFPSYQFPSPSCNFQTTGDVFVFHPGHPNPNSNPH
PRTPTSVSFSQGEEYNFFFRQQQPPQPHYSYPAPSTPQPMPQQPSAPPMHCS
HSYQQQSASHMPHSAFPGMGGTYAYAATTPPTPNTASAGTSSSSAAFGW
HGHLHLSPLTSPLATPAVPMKRLQRQSDAARKRTLSTGEDEREYQ
DHEATWDEFGRDNYDFTGERLQEFNGRIPPRKKNSHNSSTSSNPPVC
HTDSQPDNASQAAGATNQHISQQRRQQPSFTWPTVFTVFLAMCGFFAAR

> Grim_D.yaku
MAIAYFIPDPQALLARSQQAQQAGSVDGSRPNPRQNESRQQQQHQQQ
QQQQRQFRHRGNTQFRANISVPLGREQQGAMMSFECWELQUALCYAL
RLYSPNPRQQAPRTVQISFEISSNTSDADADGVGDVDATAK

> Skl_D.yaku
MAIAYFEDDEADAVASAAAAATTLESSQITDGDQVDGPAYASDAGFDQVDYE
EPSGATNSQTTTSTETEAPPSTEEOQKLAWKMLALAMCKVLQFYLQHKKQ
QQAQGKSTKLATATVQIAQ

> Grim_D.pseu
MAIAYFIPDPQALLARSQQAQQAGSVDGSRPNPRQNESRQQQQHQQQ
QQQQRQFRHRGNTQFRANISVPLGREQQGAMMSFECWELQUALCYAL
RLYSPNPRQQAPRTVQISFEISSNTSDADADGVGDVDATAK
Zhou et al. IAP-antagonist in mosquitoes

PQPQTSQQRQQRFRANISVPLGRQQGAMSSEFGCWDLLAQIFCYALRIYSYNSSQRRPTVIQISFEIRSGGQEEQQQQEEDQEEHQADATDATGKGN

>Hid_D.pseu
MAVPFYFLPeggaddiassgasgssgasgssgasgsagstgasgssgdassqasqspnnttsatqtpmqspqlpteqvylalcqvrhysqqsaaqifqypppspcscnytvgdvvfphpppvpprptrptsrvsfaageeynnfrrqqtqpppapaqstpqqmpfqqfqqqsappmhysfhpqspqlqhqhhhsavymqghqpgfgsyatypppptpntanagtstssasafghshhppqlastplatpqmgskmr

>Lqsqsdsaarrrkrltstgedereyqsdhetswddeydrydntagerrlqefngriipprrkkknshssshshshshssnhnpisqqnflhssnaecsssa
gsasahgsnsqrererereqraekekkipqvfwtptvvtvfbvmamcgfgfaar

>Rpr_D.pseu
Mavaifyipdqatllrrkaaqeqqellrlrsqswflatvvelslrqfmdsa

>SkI_D.pseu
Maiifyenenasatatapsdnatasatagdqvdgpplldpgfdhvyegpppa1aaaaaapssatsepppsaeeqllawklmlalmtckvlkqfcqhhkqatgkpptkmr
tatvqiqpqs

>Rpr_D.mela
Mavaifyipdqatllrrreaekekqqlrlrsqswflatvvelsrqytsch

>GrIm_D.mela
Maiayfipdqallarsyqqngqtaasprrttataaapsqqqqqqqqsqqqqqqqqrrhhqrrqfrqfranisvplgsqqggsmtpmsefgcwwllaqilcyalriysyssqqptviqisfeissgggnnededvtdatsken

>Hid_D.mela
MAVPFYLPeggaddvassgasgssgasgssgasgsasssasassssasssasgassgsagstgasgssgdassqasqspnnttsatqtpmqspqlpteqvylalyewe

>SkI_D.mela
Maiipfieehapksapdpmpmpmghpmmnhppnpnsnhpsrhptpr
tsvsfssgeeynnfrrqqqpqhpshysapstppmpmqppmpmqppmmpmhchshypq

7
IAP-antagonist in mosquitoes

>Sk1_D.moja
MAIPFYEDEPETATAANLDASQITDADQVDGPLSYSSNDANGYDQVDYE
APSASASASASAATSNTQATSTDNEMPPSTEEQLLVWKMLALAMCKVLK
QFYLQQKKQQQPGKTTTLATVQIQPQAQ

>Rpr_D.moja
MAVAFYIPDQKEMLRKAREQEKQLLRLRQWQMATVVLNTIRQLMDSTASE
CVQRAGRKCGKWRKPSQ

>Grim_D.moja
MAIAYFIPDQALLARSRQQASAQQAGGGVGSSPNPRQNETQQQQQP
QQQQQRQORQQQQRTGSSQFRANISVPLGREQGAMSSEFCWELLAQIFC
YALRLYSYNPAQRRTVIQISFEISSSGAGEANADGGAADAAGDVTA

III.) Matrix and search results.

a.) Scoring matrix for the Iap-Binding Motif built with MEME (Bailey and Elkan, 1994).

log-odds matrix: alength= 20 w= 9 n= 5 bayes= -664.386 E= 4.1e-005

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>333</td>
<td>-326</td>
<td>-611</td>
<td>-330</td>
<td>-343</td>
<td>-239</td>
<td>-231</td>
<td>-263</td>
<td>-335</td>
<td>-452</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-423</td>
<td>-159</td>
<td>-479</td>
<td>-181</td>
<td>-151</td>
<td>132</td>
<td>2</td>
<td>-100</td>
<td>-353</td>
<td>-478</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-265</td>
<td>-212</td>
<td>-551</td>
<td>-244</td>
<td>-226</td>
<td>-133</td>
<td>-47</td>
<td>244</td>
<td>-382</td>
<td>-429</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-467</td>
<td>-113</td>
<td>196</td>
<td>-112</td>
<td>-115</td>
<td>104</td>
<td>19</td>
<td>-179</td>
<td>-395</td>
<td>-490</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-438</td>
<td>-205</td>
<td>-408</td>
<td>-456</td>
<td>291</td>
<td>-230</td>
<td>-189</td>
<td>-325</td>
<td>-229</td>
<td>-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-385</td>
<td>-228</td>
<td>-507</td>
<td>-260</td>
<td>-240</td>
<td>-104</td>
<td>-192</td>
<td>-211</td>
<td>-142</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-383</td>
<td>-115</td>
<td>-460</td>
<td>-133</td>
<td>-107</td>
<td>-40</td>
<td>-100</td>
<td>-197</td>
<td>-38</td>
<td>277</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-348</td>
<td>-216</td>
<td>-335</td>
<td>-103</td>
<td>-340</td>
<td>-177</td>
<td>-264</td>
<td>253</td>
<td>-112</td>
<td>181</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-232</td>
<td>-156</td>
<td>-505</td>
<td>-166</td>
<td>-156</td>
<td>-75</td>
<td>-26</td>
<td>140</td>
<td>-318</td>
<td>-376</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-467</td>
<td>-101</td>
<td>280</td>
<td>-54</td>
<td>-68</td>
<td>45</td>
<td>-13</td>
<td>-217</td>
<td>-401</td>
<td>-483</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-471</td>
<td>25</td>
<td>-384</td>
<td>88</td>
<td>-54</td>
<td>19</td>
<td>-26</td>
<td>-245</td>
<td>-450</td>
<td>-473</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b.) The above matrix were used for testing search to the Swiss-Prot release 43.3 (total 150,889 annotated protein sequences)(Bairoch and Apweiler, 2000). Total 11 hits were identified as having a score larger than 1500 based on the matrix. The top three are the “true” hits, i.e. Reaper, Grim, and Hid. Sickle was not included in this release of Swiss-Prot.
Zhou et al.  

IAP-antagonist in mosquitoes

dbname:>RPR_DROME (Q24475) Cell death protein rpr (Reaper protein) len 65
Matches at:0
score 2360
MAVAFYIPD QATLLREAEOQEQQLRLRESQNRFQATLVV

$dbname:>HID_DROME (Q24106) Cell death protein W (Head involution defective protein) (Wrinkled protein) len 410
Matches at:0
score 2266
MAVFFYLPQ GAADDVASSSGASGNNSSHPHHLPFSSLAS

$dbname:>GRIM_DROME (Q24570) Cell death protein Grim len 138
Matches at:0
score 1949
MAIAFIPD QAQLLARYQMGYQQTASSRFTTATAAPS

$dbname:>SERC_XANAC (Q8PLY7) Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT) len 361
Matches at:296
score 1818
NAEKAALVYGAIDGSGGFYRNLIKPAVSR SNIPFFLPD ERLDALFVSESKAAGLLALKGHKAVGGIRA

$dbname:>O10A_DROME (Q9VYZ1) Putative odorant receptor 10a len 406
Matches at:375
score 1683
AMFSPCWQLFKQPKQRLVQQLILRSPQRFPS MAVFFFSPS LATFAAILQTSGSIIALVKSFQ NINQVVK

$dbname:>Y022_NPVUP (O10281) Hypothetical 43.4 kDa protein (ORF20) len 382
Matches at:80
score 1585
RRQVYPMKQPELPVFQDNLQTLGKTYQV DAVFQQVQDVPN SAVLYVSE IDLPMFDCGECOPNAVVFGEQDTYV

$dbname:>SERC_XYLFA (Q9PB19) Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT) len 362
RD:>SERC_XYLFA
Matches at:297
score 1548
NQAAQALYVQTDQSGFHYRNGVAAVSR SNIPFFLPQ VEQQADFQAAEAAAGLLSLKGGHALGGIRA

$dbname:>YIAN_HAEIN (P44993) Hypothetical protein HI1029 len 425
Matches at:56
score 1545
CGIALMLHLDFFNQILAQQILVGGADSFSL MAIPIIFLA GEINMENGLSKRIDDLMKLVGHKRGGLGF

$dbname:>UNG_CHRVO (Q7P111) Uracil-DNA glycosylase (EC 3.2.2.-) (UDG) len 238
Matches at:80
score 1541
QVFNALAPAAPADDVKKVLQDPFNYGDEA MGLSFSVPD GARVFPSSLKNIKELAALDGLGVPASGDLT

$dbname:>Y022_NPVAC (P41427) Hypothetical 43.8 kDa protein in EGT-IAP1 intergenic region (ORF 8/6) len 382
Matches at:80
score 1534
RRHYPFIALPQFRDNLQTLGKTYQV MAVFQFVSIDLPMFDCQIIDNPSAYFFVNETDFVYV

$dbname:>ATU1_YEAST (P38360) Probable copper-transporting ATPase (EC 3.6.3.4) (Cu(2+)-ATPase) len 1216
Matches at:943
score 1521
TVVHETGRDSRLNQSSLLGLLGGKHPVs MAIAYLKE KGVSQMVNTKAVGKRVEGSTSYSLKLQ

150889 entries searched

References:


