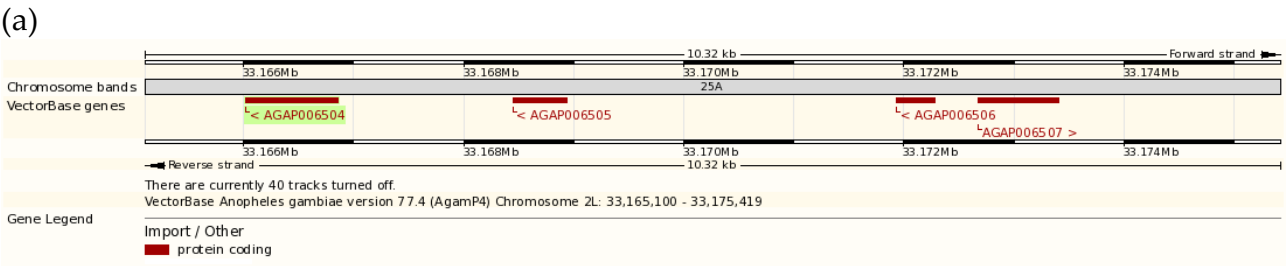


Figure S3: Sg2/sg2a/AGAP006505 gene cluster.



(b)  
Sequence alignment

```
sg2      MKSILVTFATLSVALVVVAIPTSFNYGGGGYFINGTGQSFNFSGESNGTSIPGVPDFG
sg2a     MKLFLTLTLLTSLV--VMVFALPAHHHSRGGDGSSANSTGQSYNNSGSPNNS--AGVPDFG
AGAP006505 MKLFLPLLSTLSV--AMVFALPAHHHSRGGEGSSANNTGHSNDYVD---SS--NSVPAIG
          ** : * : **** :.:.*: :. . ** * * ** : * : . ** : *

sg2      SFL----PNLGNLTQ----QFGSSGAFPPQS-----IPSWTNFTD-----AFSSI
sg2a     FNSQSNVPGFGNGQQPGQQQQGQGFPPFGQGSFSGFNRQPFQGNQQGQDG---
AGAP006505 FNPPSNVPGFGIGKQSRQQQ---QGQFSLFTQGQGFNIGNDQPVFDYNQGPYPVPSA
          * : * * * * . . * * : * *

sg2      FPPFG--NVQGGGFPPFG-----
sg2a     -----NGQQGRGIPFFGQGGGQGGIPSGSQQNGGIPSLGNGQAQSGFSPFGNGQQGGN
AGAP006505 VPKYGVGQENIGPPFYFGARGDTGLPHIGYIDL-----G-S
          : * ** :

sg2      -----
sg2a     FPPFG
AGAP006505 YPIVL
```

(c)

	Most frequent aminoacids			Cysteins
sg2	17.5% Gly	15.8% Phe	12.3% Ser	0 Cys
sg2a	24.3% Gly	16.2% Gln	11.6 Phe	0 Cys
AGAP006505	14.7% Gly	10% Pro	10% Ser	0 Cys

Figure S3: sg2 (AGAP006504)/sg2a (AGAP006506) gene cluster. (a): scheme of the genomic region (chromosome 2L) where sg2, sg2a and AGAP006505 are placed and separated by a short region of around 8 kbp. (b): sequence alignment of sg2, sg2a and AGAP006505. (c): table summarizing the most frequent aminoacids in the sequence of the three candidates: all predicted peptides present a high content of glycine and no cysteine.