Changes in a pseudogene mistakenly mapped as SNPs in the functional gene.
a, AGP1 aligned with AGP2. Boxed amino acids are the substitutions that were nsSNP entries for *AGP1* in dbSNP. These substitutions were derived from GenBank AC# AA700876, T27840, H67119, T73068, T70907, H82213, and AA550779, all of whose nucleotide sequences matched *AGP2*. Amino acids in lowercase bold italic are the substitutions predicted to be deleterious. AGP2 was not included in the alignment used for prediction on AGP1 because it was > 90% identical to AGP1 and thus discarded (see Methods). b, FKSG25 aligned with the translated pseudogene and sequences from which nsSNPs were derived. The pseudogene has a deletion that introduces an early stop codon; the translated pseudogene can be aligned to FKSG25 after the deletion with a frameshift of two bases. AC067884, AL365277, AC024511 are the contigs from which nsSNPs (boxed) were interpreted and deposited into dbSNP and all contain the same deletion as the pseudogene. Substitutions predicted to be deleterious are in lowercase bold italic.

a

```
AGP1       MALSWVLTVL SLLPLLEAQI PLCANLVPVP ITNATLDQIT GKWFYIASAF RNEEYNKSVQ EIQATFFYFT
AGP2       .......... .......... .......... .......R.. .......... .......... ..........

AGP1       PNKTEDTIFL REYQTRQDQC IYNTTYLNVQ RENGTISRYV GGQEHFAHLL ILRDTKTYML AFDVNDEKNW
AGP2       .......... .......N.. F..SS..... .....V...E ..R..V.... f......L.f Gg SYLD.....

AGP1       GLSVYADKPE TTKEQLGEFY EALDCLRIPK SDVVYTDWKK DKCEPLEKQH EKERKQEEGE S
AGP2       ...F...... .......... ......c..R ...m...... .......... .......... .
```

b

```
FKSG25     MAALRLLASV LGRGVPAGGS GLALSQGCAR CFATSPRLRA KFYADPVEMV KDISDGATVM IGGFGLCGIP
pseudogene .......... .......... .......... .......P.. .......... .......... ..........
AC067884   .......... .......... .......... .......... .......... .......... ..........
AL365277   .......... .......... .......... .......P.. .......... .......... ..........
AC024511   .......... .......... .......... .......P.. .......... .......... ..........

FKSG25     ENLIAALLRT RVKDLQVVSS NVGVEDFGLG LLLAARQVRR IVCSYVGENT LCESQYLAGE LELELTPQGT
pseudogene .......... .......... .......... .......... .......... .......... ..........
AC067884   .......... .......... .......... .......... .......... .......... ..........
AL365277   .......... .......... .......... .......... .......... .......... ..........
AC024511   .......... .......... .......... .......... .......... .......... ..........

FKSG25     LAERIRAGGA GVPAFYTPTG YGTLVQEGGA PIRYTPDGHL ALMSQPREVR EFNGDHFLLE RAIRADFALV
pseudogene ..........e. .......... .......... .......... .......... .......... ..........
AC067884   .........e. .......... .......... .......... .......... .......... ..........
AL365277   ..........e .......... .......... .......... .......... .......... ..........
AC024511   .......... .......... .......... .......... .......... .......... ..........

                                          ***** ***
FKSG25     KGWKADRAGN VVFRRSARNF NVPMCKAADV TAVEVEEIVE VGAFPPEDIH VPNIYVDRVI KGQKYEKRIE
pseudogene .......... .......... .......... .....          ....... ......G... ..........
AC067884   .......... .......... .......... .....          ....... ......G... ..........
AL365277   .......... .......... .......... .....          ....... ......G... ..........
AC024511   .......... .......... .......... .....          ....... ......G... ..........

FKSG25     RLTILKEEDG DAGKEEDART RIIRRAALEF EDGMYANLGI GIPLLASNFI SPSMTVHLHS ENGILGLGPF
pseudogene ....R..... .......... .......... .......... .......... .......... ..........
AC067884   ....R..... .......... .......... .......... .........V .......... ..........
AL365277   ....R..... .......... .......... .......... .......... .......... ..........
AC024511   ....R..... .......... .......... .......... .......... .......... ..........

FKSG25     PTEDEVDADL INAGKQTVTV LPGGCFFASD DSFAMIRGGH IQLTMLGAMQ VSKYGDLANW MIPGKKVKGM
pseudogene .......... .......... .......... .......... .......... .......... ..........
AC067884   .......... .......... .......... .......... .......... .......... ...s......
AC024511   .......... .......... .......... .......... .......... .......... ..........

FKSG25     GGAMDLVSSQ KTRVVVTMQH CTKDNTPKIM EKCTMPLTGK RCVDRIITEK AVFDVHRKKE LTLRELWEGL
pseudogene .......... .......... .......... .......... .......... .........G ..........
AC067884   .......... .......... .......... .......... .......... .......... ..........
AL365277   .......... .......... .......... .......... .......... .........R ..........
AC024511   .......... .......... .......... .......... .......... .........G ..........

FKSG25     TVDDIKKSTG CAFAVSPNLR PMQQVAP
pseudogene ...n...... .......... .......
AC067884   .......... .......... .......
AL365277   .......... .......... .....V.
AC024511   ...n...... .......... .......
```