

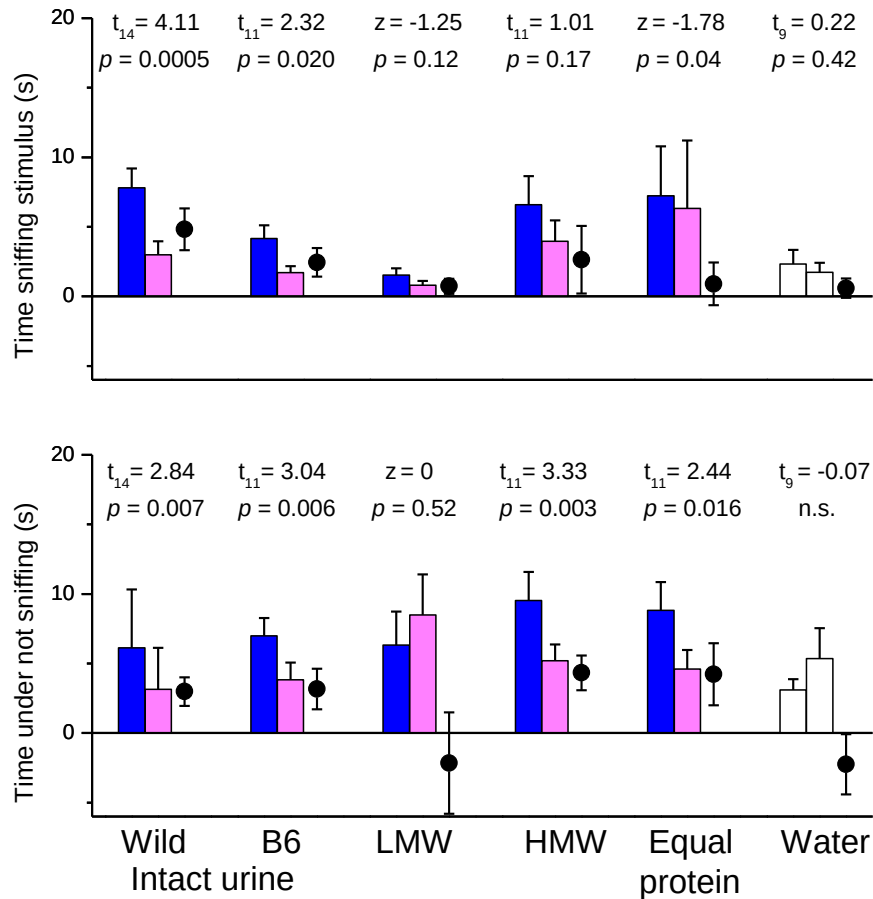


Figure S1. Female sexual attraction to male urine: time sniffing and not sniffing.

Time spent sniffing up or not sniffing when under a male (blue bars) and a matched female (pink bars) urine stimulus, together with the difference in response to male minus female stimulus (circles), plotted as means $\pm$ s.e.m. Significant P values indicate greater response to the male stimulus (matched pair t -tests of log transformed data, or Wilcoxon matched pair tests when transformed data did not approximate normality), provided as a rough indication of which fractions retained the attraction response. Wild: urine from a random selection of wild males and females ($n=15$); B6: normal intact B6 urine ($n=12$); LMW: low molecular weight urine fraction ($< 3\text{kDa}$, $n=12$); HMW: high molecular weight urine fraction ($\geq 3\text{kDa}$, $n=12$); Equal protein: male urine diluted to the same protein concentration as female urine ($n=12$); Water (open bars): two water stimuli ($n=10$). Total time under (sniffing + not sniffing) is shown in Figure 1.

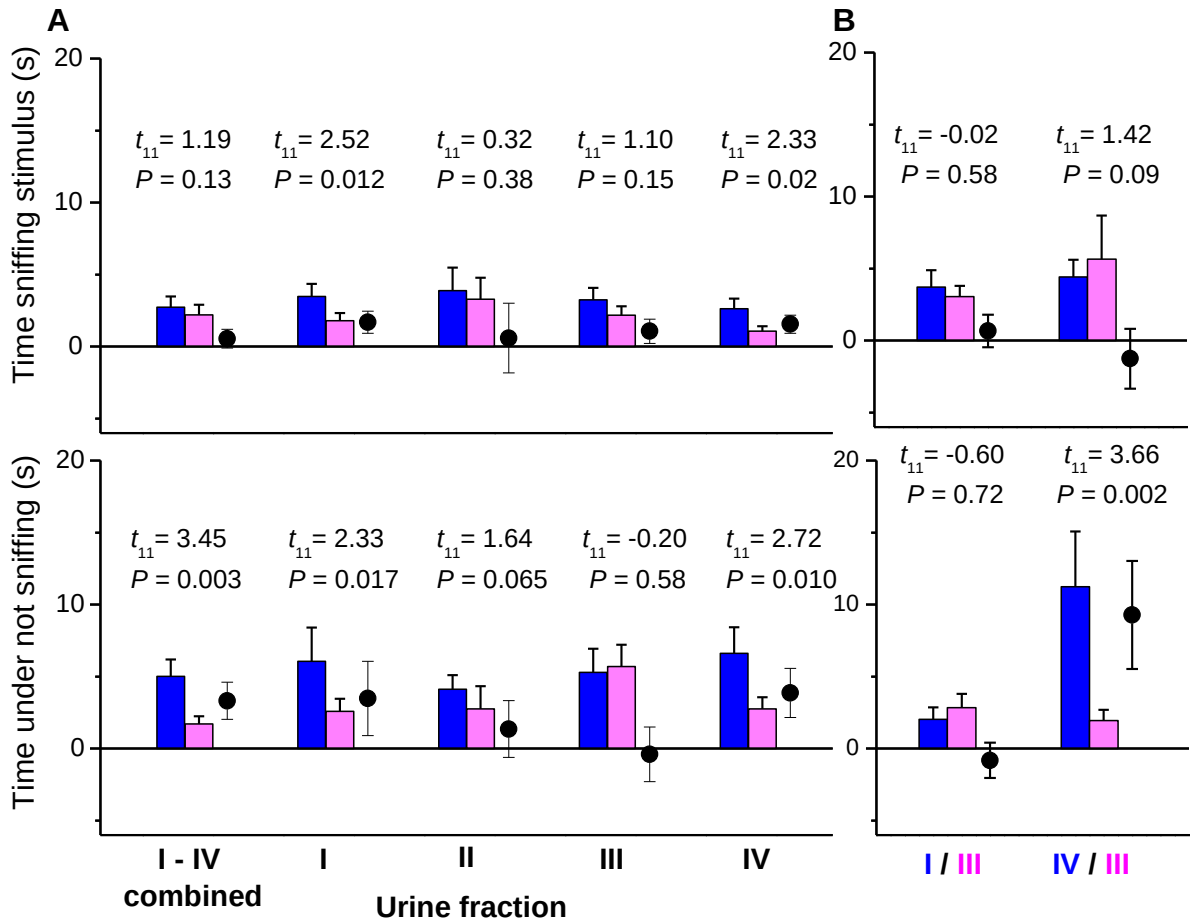


Figure S2. Female sexual attraction to MUP urine fractions: time sniffing and not sniffing.

Time spent sniffing up or not sniffing when under a male (blue bars) and a matched female (pink bars) urine stimulus, together with the difference in response to male minus female stimulus (circles), plotted as means $\pm$ s.e.m. Significant P values indicate greater attraction to the male stimulus (matched pair t -tests of log transformed data), provided as a rough indication of which fractions retained the attraction response. (A) Response to urine fractions containing different MUPs (separation shown in Figure 2). I: 18645Da MUP (in male fraction only); II: 18709Da MUP (both sexes); III: 18694Da and 18713Da MUPs (both sexes); IV: 18893Da MUP (r-darcin, male only). The protein concentration in each male and female fraction is given in Table 2. (B) Male fractions I and IV were subsequently tested against female fraction III. Total time under (sniffing + not sniffing) is shown in Figure 3.

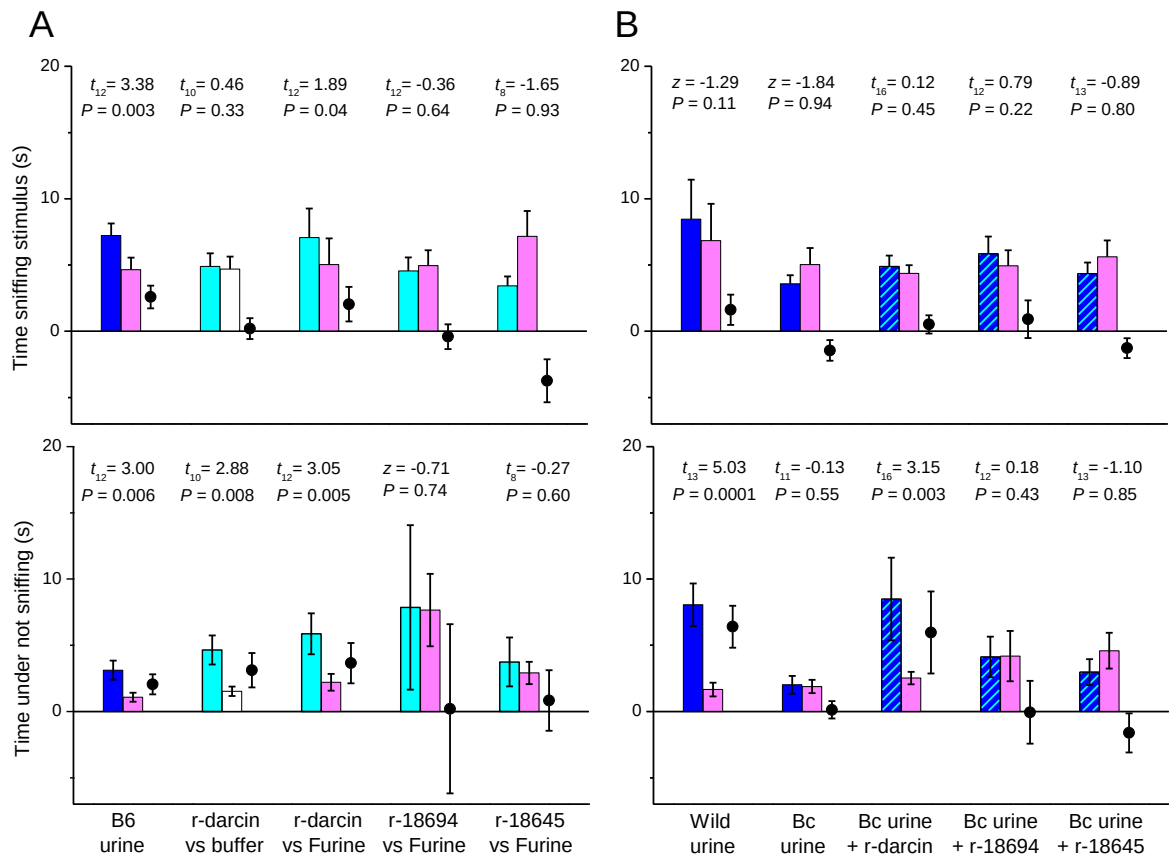


Figure S3. Female sexual attraction to male urine is elicited by darcin: time sniffing and not sniffing.

Time spent sniffing up or not sniffing when under a male stimulus (blue bars: male urine; cyan bars: recombinant MUP alone; hatched bars: male urine plus recombinant MUP) and a matched control stimulus (pink bars: BALB/c female urine with or without buffer; open bars: buffer), together with the difference in response to male minus female stimulus (circles), plotted as means $\pm$ s.e.m. Significant P values indicate greater attraction to the male stimulus (matched pair t -tests of log transformed data, or Wilcoxon matched pair tests when log transformed data did not approximate normality). Response to 11 μ g recombinant MUP presented alone (A) or added to 10 μ l BALB/c male urine (B). B6: C57BL/6 urine control; r-darcin: recombinant darcin; r-18694: recombinant 18694Da MUP; r-18645: recombinant 18645Da MUP; Furine: female urine plus buffer; Wild: wild male versus BALB/c female urine control; Bc: BALB/c urine. The total time under each stimulus is given in Figure 7.

rDarcin	- - - - - 10 - - - - - 20 - - - - - 30 - - - - - 40 - - - - - 50 - - - - - 60 - - - - - 70 - - - - - 80
B6 darcin	A T G A A G C T G C T G G T G C T G C T G C T G T G T T T G G G A C T G A C C C T A G T C T G T G T G C A T G C A G A A G A A G C T A G T T C T A T G G A A C G
rDarcin	C A A C T T C A A C G T C G A G A A A A T C A A C G G T G A A T G G T A C A C T A T C A T G C T G G C T A C G G A C A A A C G T G A A A A A T C G A A G A G C
B6 darcin	G A A C T T T A A T G T A G A A A A G A T T A A T G G G A A T G G T A T A C T A T T A T G C T G G C A C T G A C A A A A G A G A A A A G A T A G A A G A A C
rDarcin	A C G G C T C T A T G C G T G T T T T C G T T G A A T A C A T C C A C G T T C T G G A A A A C T C T C T G G C G C T G A A A T T C C A C A T C A T C A T C A A C
B6 darcin	A T G G C A G C A T G A G A G T T T T T G T C G A G T A C A T C C A T G T C T T G G A G A A T T C C T T A G C T C T T A A A T T C C A T A T T A T T A T A A A T
rDarcin	G A A G A G T G C T C T G A A A T C T T C C T C G T T G C G G A C A A A A C C G A A A A A G C T G G T G A A T A C T C T G T A A C C T A C G A C G G C T C T A A
B6 darcin	G A G G A G T G C T C T G A A A T A T T T T T G G T T G C T G A C A A A A C A G A A A A G G C T G G T G A A T A T T C T G T A A C G T A T G A T G G A T C G A A
rDarcin	C A C C T T T A C C A T C C T G A A A A C G G A C T A C G A C A A C T A C A T C A T G A T C C A C C T G A T C A A C A A A A A G A C G G T G A A A C G T T C C
B6 darcin	T A C A T T T A C T A T A C T T A A G A C A G A C T A T G A T A A T T A T A T T A T G A T T C A T C T C A T T A A C A A A A A G G A T G G A G A A A C C T T T C C
rDarcin	A G C T G A T G G A A C T G T A C G G T C G T G A A C C G G A T C T G T C T A G C G A C A T C A A A G A A A A A T T C G C A C A G C T G T C T G A A G A A C A T
B6 darcin	A G C T G A T G G A C T C T A T G G C C G A G A A C C A G A T T T G A G T T C A G A C A T C A A G G A A A A G T T T G C A C A A C T A A G C G A G G A G C A T
rDarcin	G G T A T C G T G C G T G A G A A C A T C A T T G A T C T G A C T A A C G C G A A C C G T T G C C T G G A A G C C C G T G A A
B6 darcin	G G A A T C G T T A G A G A A A A T A T C A T T G A C C T A A C C A A T G C C A A T C G C T G C C T T G A G G C C C G A G A A

His tag

rDarcin	M G S S H H H H H H I E G R E E A S S M E R N F N V E K I N G E W Y T I M L A T D K R E K I E E
B6 darcin	- - - - - - - - - - - - E E A S S M E R N F N V E K I N G E W Y T I M L A T D K R E K I E E
rDarcin	H G S M R V F V F E Y I H V L E N S L A L K F H I I I N E E C S E I F L V A D K T E K A G E Y S V
B6 darcin	H G S M R V F V F E Y I H V L E N S L A L K F H I I I N E E C S E I F L V A D K T E K A G E Y S V
rDarcin	T Y D G S N T F T I L K T D Y D N Y I M I H L I N K K D G E T F Q L M E L Y G R E P D L S S D I
B6 darcin	T Y D G S N T F T I L K T D Y D N Y I M I H L I N K K D G E T F Q L M E L Y G R E P D L S S D I
rDarcin	K E K F A Q L S E E H G I V R E N I I D L T N A N R C L E A R E
B6 darcin	K E K F A O L S E E H G I V R E N I I D L T N A N R C L E A R E

Figure S4. Expression of recombinant darcin.

The published sequence of darcin, verified by mass spectrometry was used to direct the design and synthesis of a gene, codon optimized for high level expression in *E. coli*. Top panel: The C57BL/6J cDNA sequence (B6 darcin) is aligned to the optimized DNA sequence (rDarcin). Bottom panel: translation of both sequences, emphasising the identity of the coding regions.

Translation of MUP 18645 synthetic gene, codon optimised

```

      M G S S H H H H H H I E G R E E A S S T
1  ATGGGCTCTAGCCATCACCACCATCACCATATCGAAGGTCGTGAAGAGGCATCCTCTACT
      10      20      30      40      50
      G R N F N V E K I N G E W H T I I L A S
61 GGTCGCAACTTCAACGTCGAGAAAATCAACGGCGAATGGCACACTATCATCCTGGCGAGC
      70      80      90      100     110
      D K R E K I E D N G N F R L F L E Q I H
121 GATAAACGCGAAAAAATTGAAGACAACGGCAACTTCCGTCTGTTTCTGGAGCAGATTCAC
      130     140     150     160     170
      V L E N S L V L K V H T V R D E E C S E
181 GTTCTGGAAAACCTCCCTCGTACTGAAAGTTCACACTGTTTCGTGATGAAGAGTGCTCCGAA
      190     200     210     220     230
      L S M V A D K T E K A G E Y S V T Y D G
241 CTGTCTATGGTGCGTGACAAAACCGAGAAAGCTGGCGAATACTCTGTGACCTATGATGGT
      250     260     270     280     290
      F N T F T I P K T D Y D N F L M A H L I
301 TTCAACACCTTCACGATCCCGAAAACCGACTACGATAACTTCCTGATGGCACACCTGATC
      310     320     330     340     350
      N E K D G E T F Q L M G L Y G R E P D L
361 AACGAGAAAGACGGCGAAAACCTTCCAGCTGATGGGTCTGTATGGTCGTGAACCGGATCTG
      370     380     390     400     410
      S S D I K E R F A Q L C E K H G I L R E
421 AGCTCTGACATCAAAGAACGCTTTGCCAGCTGTGTGAGAAACACGGTATCCTGCGTGAG
      430     440     450     460     470
      N I I D L S N A N R C L Q A R E * *
481 AACATTATTGACCTGTCTAACGCGAACCGTTGCCTGCAGGCGCGTGAATAATGA
      490     500     510     520     530

```

Translation of MUP 18694 synthetic gene, codon optimised

```

      M G S S H H H H H H I E G R E E A S S T
1  ATGGGCTCTAGCCATCACCACCATCACCATATCGAAGGTCGTGAAGAGGCATCCTCTACT
      10      20      30      40      50
      G R N F N V E K I N G E W H T I I L A S
61 GGTCGCAACTTCAACGTCGAGAAAATCAACGGCGAATGGCACACTATCATCCTGGCGAGC
      70      80      90      100     110
      D K R E K I E D N G N F R L F L E Q I H
121 GATAAACGCGAAAAAATTGAAGACAACGGCAACTTCCGTCTGTTTCTGGAGCAGATTCAC
      130     140     150     160     170
      V L E N S L V L K F H T V R D E E C S E
181 GTTCTGGAAAACCTCCCTCGTACTGAAATTCCACACTGTTTCGTGATGAAGAGTGCTCCGAA
      190     200     210     220     230
      L S M V A D K T E K A G E Y S V T Y D G
241 CTGTCTATGGTGCGTGACAAAACCGAGAAAGCTGGCGAATACTCTGTGACCTATGATGGT
      250     260     270     280     290
      F N T F T I P K T D Y D N F L M A H L I
301 TTCAACACCTTCACGATCCCGAAAACCGACTACGATAACTTCCTGATGGCACACCTGATC
      310     320     330     340     350
      N E K D G E T F Q L M G L Y G R E P D L
361 AACGAGAAAGACGGCGAAAACCTTCCAGCTGATGGGTCTGTATGGTCGTGAACCGGATCTG
      370     380     390     400     410
      S S D I K E R F A Q L C E E H G I L R E
421 AGCTCTGACATCAAAGAACGCTTTGCCAGCTGTGTGAGGAACACGGTATCCTGCGTGAG
      430     440     450     460     470
      N I I D L S N A N R C L Q A R E * *
481 AACATTATTGACCTGTCTAACGCGAACCGTTGCCTGCAGGCGCGTGAATAATGA
      490     500     510     520     530

```

Figure S5. Expression of 18645Da and 18694Da recombinant MUPs.

The published sequences of MUPs expressed by C57BL/6 mice, verified by mass spectrometry [Robertson DH, Cox KA, Gaskell SJ, Evershed RP, Beynon RJ: Molecular heterogeneity in the Major Urinary Proteins of the house mouse *Mus musculus*. *Biochem J* 1996, 316:265-272] was used to direct the design and synthesis of genes, codon optimized for high level expression in *E. coli*. The protein sequence (ORF) is aligned to the DNA sequence that was optimized and synthesized for maximal expression -this will differ from the published cDNA sequence.