

## Supplementary Data

Gene <sup>1</sup>	Forward primer (5'-3') <sup>2</sup>	Reverse primer (5'-3') <sup>2</sup>	Length (bp)	Template	Vector	Integration site	Transgenic line code	Source
<i>Ir7a</i>	AGATCTGCTGAAGAAATAGAGTGTGGC	GAATTCCTTGAACGAAACTGTTGCG	2318	OR genomic DNA	pGAL4 attB	attP2	BT58.1	(a)
<i>Ir7b</i>	AGATCTGGGATGAGAAGACGACATCGAT	GAATTCGGCTAAAGAGTTGCCAAAGG	578	OR genomic DNA	pGAL4 attB	attP2	BT47.1	(b)
<i>Ir7c</i>	CTCCGAGCCCGTTAGTGGTCCAAATA	AGATCTATCGTGTTCATCCGGTGGCT	1595	OR genomic DNA	pGAL4 attB	attP2	BT110.1	
<i>Ir7d</i>	AGATCTAACTTGTGCAATCCGATCC	AGATCTGGCGAATGTGAACATTTGG	1010	OR genomic DNA	pGAL4 attB	attP2	BT79.1	(b)
<i>Ir7e</i>	AGATCTTACTCCGAGAGGAACATAG	GAATTCCTGCCCGGACAAATCGT	600	OR genomic DNA	pGAL4 attB	attP2	BT59.1	(b)
<i>Ir7f</i>	AGATCTGTCCGCTATCCGAAATCCGG	AGATCTATCGATCCCTCGAATTTCTCCA	766	OR genomic DNA	pGAL4 attB	attP2	BT98.1	
<i>Ir7g</i>	AGATCTATCGATCCCGAATTTCTCCA	AGATCTGCTCCCTATCCGAAATCCGG	766	OR genomic DNA	pGAL4 attB	attP2	BT99.1	(b)
<i>Ir10a</i>	GC GGCCGC GACACATATAGTCCCACTACC	CGGGCCGCTCGTATGGGATTTGTTAGCAC	2429	OR genomic DNA	pGAL4 attB	attP2	BT96.1	
<i>Ir11a</i>	AGATCTATGATATGTCATGCCACCAGC	GAATTCGACTGAATGCCGTTGTGAA	2099	OR genomic DNA	pGAL4 attB	attP2	BT51.1	(a)
<i>Ir20a</i>	AGATCTACATTGTTCCGGCCTCCGACAG	AGATCTGTCGCCGACATCGAAGGAAT	2488	OR genomic DNA	pGAL4 attB	attP2	BT61.1	
<i>Ir47a</i>	AGATCTGCTGAGTGGGTGACGAAATC	GAATCTTTTTATGGCCCTTTTGAAC	2488	OR genomic DNA	pGAL4 attB	attP2	BT53.1	
<i>Ir47b<sup>4</sup></i>								
<i>Ir48a<sup>4</sup></i>								
<i>Ir48b</i>	AGATCTCCAGCTCCAGTTCCAGATTGC	AGATCTCTGAAAGATATATAGAGCGT	2575	OR genomic DNA	pGAL4 attB	attP2	BT97.1	
<i>Ir51a<sup>4</sup></i>								
<i>Ir51b</i>	AGATCTCAACCAATCAAGCTGGATAC	AGATCTGGTGGTTGATCAATTGTGACA	2499	OR genomic DNA	pGAL4 attB	attP2	BT82.1	
<i>Ir52a</i>	AGATCTCCGACATTTCTCCGCTAAC	AGATCTCAGCAAACTGTTGACAAATCC	2931	OR genomic DNA	pGAL4 attB	attP2	BT100.1	
<i>Ir52b</i>	AGATCTAGCTGGAGATATTCGGTTGGC	GAATTCGTGTTTCAAACGAACTGTTT	473	OR genomic DNA	pGAL4 attB	attP2	BT55.1	(a)
<i>Ir52c</i>	AGATCTAAACGCTGGATGAAATCCG	AGATCTGGTGGCTAAAGTGAATAATGG	644	OR genomic DNA	pGAL4 attB	attP2	BT80.1	
<i>Ir52d</i>	AGATCTTGAGACTGCGAAGGAACTGC	GAATCTGGTCCAAGAGTTACTATTGC	664	OR genomic DNA	pGAL4 attB	attP2	BT56.1	
<i>Ir52e<sup>4</sup></i>								
<i>Ir54a</i>	AGATCTGACCCAAGTCGACTCAITTTG	AGATCTGTCTTCAATATGTTGCC	2493	OR genomic DNA	pGAL4 attB	attP2	BT101.1	
<i>Ir56a<sup>4</sup></i>	AGATCTATCATCAGTGGCTGTCAATCC	GAATTCGGCTGCTTACACACTTTGAC	2400	OR genomic DNA	pGAL4 attB	attP2	BT52.1	(a)
<i>Ir56b</i>	CGGGCCGCATATCCTTCGGTCCGAAAGTGC	CGGGCCGCTGAAATAATTCTGCACTTGA	2461	OR genomic DNA	pGAL4 attB	attP2	BT82.1	
<i>Ir56c</i>	AGATCTGCAAGAGGACTCCACAGATAG	GAATTCGACTTCCCTAGAAAGCAAC	319	OR genomic DNA	pGAL4 attB	attP2	BT57.1	(b)
<i>Ir56d<sup>4</sup></i>	GAATCTATGAGGACTGCAACATGCTC	AGATCTATATTTGATACGGCAGCTGCC	858	OR genomic DNA	pGAL4 attB	attP2	BT76.1	
<i>Ir56e<sup>4</sup></i>								
<i>Ir60a</i>	GAATCTAGTCCCGGAGCTGATTATC	GAATCTATTGCTCTGTGACGCTCGG	2523	OR genomic DNA	pGAL4 attB	attP2	BT83.1	
<i>Ir60b</i>	AGATCTCATACGATTTCCCGAACAGC	AGATCTTTTCGAGTTGTCTGCTCTGG	2368	OR genomic DNA	pGAL4 attB	attP2	BT104.1	
<i>Ir60c<sup>4</sup></i>	GAATTCGATTGGATACCAAGGTTGGC	GAATTCGGCAGCTATCCGAAACGAGC	560	OR genomic DNA	pGAL4 attB	attP2	BT81.1	(b)
<i>Ir60d</i>	AGATCTAGATGTTGGTACCACAGATGG	GAATCTTTTAAGGCGACTGCTCACCA	339	OR genomic DNA	pGAL4 attB	attP2	BT63.1	
<i>Ir60e</i>	AGATCTAAATTAATGAGCAGTCCGAT	GAATTCAGGCGAGGGAAAATGCTT	2466	OR genomic DNA	pGAL4 attB	attP2	BT72.1	
<i>Ir60f<sup>4</sup></i>								
<i>Ir62a</i>	GAATTCAAATACCAGGTTCAATGGC	GAATTCATTTTCGCTCGTGAACCATG	2439	OR genomic DNA	pGAL4 attB	attP2	BT118.1	
<i>Ir67a</i>	AGATCTACAGACGTTTATCAGCAAAAG	GAATTCATATCCTGGCTGAATGGCTG	2496	OR genomic DNA	pGAL4 attB	attP2	BT73.1	
<i>Ir67b</i>	AGATCTTGGTGTGTCAGCAGTATAGC	GAATTCGAAATGTCTCTGAAATCTCT	515	OR genomic DNA	pGAL4 attB	attP2	BT74.1	
<i>Ir67c</i>	CGGGCCGC GGTTGCTCCATCGTATCCTTTC	CGGGCCGCGATGACGTCCTGCCCGAAAA	2736	OR genomic DNA	pGAL4 attB	attP2	BT84.1	
<i>Ir68b</i>	AGATCTCCGCTTACTCGAAGATATG	GAATTCGTTCTAGCAGCAGCTAACCC	637	OR genomic DNA	pGAL4 attB	attP2	BT48.1	
<i>Ir68a</i>	AGATCTAAGTCCCTCTCAAGTTGTC	GAATCTGCAATGCCAAGTGTGAG	1369	OR genomic DNA	pGAL4 attB	attP2	BT49.1	
<i>Ir87a</i>	CTCGAGACATACCCATATGGACACCG	AGATCTGCCGCAAAAGAACTGATGAT	2097	OR genomic DNA	pGAL4 attB	attP2	BT122.1	
<i>Ir94a</i>	AGATCTGCACACAGATAGATTCCGAC	AGATCTTTTCTACTTTAGCCAAACAAT	2294	OR genomic DNA	pGAL4 attB	attP2	BT64.1	
<i>Ir94b</i>	AGATCTAAGATCAAGCGAAGATGACG	GAATTCATTCTGTAATTCACGATAGT	482	OR genomic DNA	pGAL4 attB	attP2	BT77.1	
<i>Ir94c</i>	AGATCTTTCTGGCAGCGTCTCTATC	AGATCTTTTATGTAAGCTTGGGTTA	2303	OR genomic DNA	pGAL4 attB	attP2	BT65.1	
<i>Ir94d</i>	CTCGAGACATTTGTTCCGGTACGTTG	AGATCTTTTGAATGTGGGAAATGTTGGT	2420	OR genomic DNA	pGAL4 attB	attP2	BT111.1	
<i>Ir94e</i>	AGATCTTTGGCGACATAAGATGTGGC	GAATCTTCCCGAGGGATTACACAAA	322	OR genomic DNA	pGAL4 attB	attP2	BT78.1	(b)
<i>Ir94f</i>	AGATCTGATTTGGAGCGGATCGATTG	GAATTCCTGTGACAGCAGTATGATG	2493	OR genomic DNA	pGAL4 attB	attP2	BT102.1	
<i>Ir94g</i>	GAATTCGAGCTCACTGTTCAATATGC	GAATTCCTTATAACTGACTTCAATT	388	OR genomic DNA	pGAL4 attB	attP2	BT75.1	
<i>Ir94h</i>	GAATCTTGTTCACGCGGCAATTAGC	GAATTCGACTTATACCGAAACCGACG	2000	OR genomic DNA	pGAL4 attB	attP2	BT60.1	(b)
<i>Ir100a</i>	AGATCTTTTACCGAGCTGATAGTAG	GAATCTGTCAGAGTTACTGAAACCGT	512	OR genomic DNA	pGAL4 attB	attP2	BT50.1	(a)

## Footnotes

- <sup>1</sup>: <sup>4</sup> = predicted pseudogene in the reference *D. melanogaster* genome (Adams et al., Science 2000); for most of these we therefore did not construct a driver line. For *Ir48a-Gal4* we observed very variable expression in the central and peripheral nervous system (data not shown). For *Ir60c*, this locus is predicted to be intact in a *w<sup>1118</sup>* strain
- <sup>2</sup>: Restriction enzyme sites in cloning primers are highlighted in blue
- <sup>3</sup>: Lines previously published: (a) Croset et al., PLOS Genetics (2010); (b) Croset et al., Scientific Reports (2016)
- <sup>4</sup>: *Ir52e* is not present in the reference *D. melanogaster* genome, so no driver line was constructed for the locus in this study (see also Koh et al., Neuron 2014)
- <sup>5</sup>: *Ir56a-Gal4* displays expression in several olfactory sensory neuron populations; this is likely to be non-specific as there is no evidence for antennal expression of *Ir56a* (Menuz et al., PLOS Genetics 2014); as *Ir56a* is located within the intron of another gene (*5-HT1A*) this may reflect overlapping regulatory elements of these genes
- <sup>6</sup>: *Ir56d-Gal4* was also detected in some larval head chemosensory neurons, but this expression weak and was not confirmed in the *Ir56d<sup>56d</sup>* reporter allele.
- <sup>7</sup>: *Ir60a-Gal4* displays extensive and variable non-neuronal expression (data not shown), which is likely to be non-specific; as *Ir60a* is located within the intron of another gene (*nord*), this may reflect overlapping regulatory elements of these genes

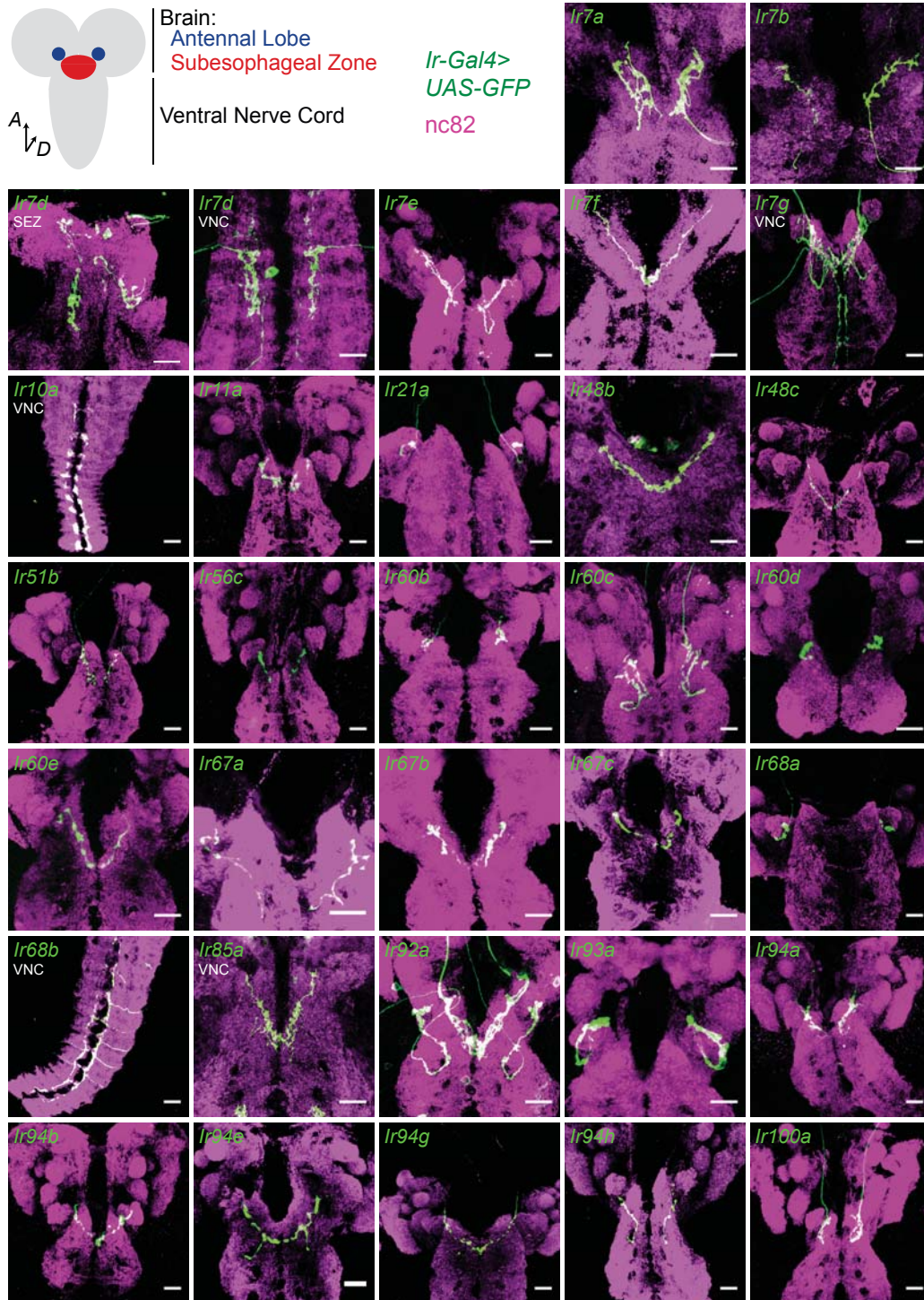
Supplementary Table 1. Oligonucleotide sequences and construction of *Ir-Gal4* lines

Tastant	CAS	Calcium imaging Figure 5c		Notes
		Concentration		
water	-	-	-	
glycerol	56-81-5	-	10% (v/v)	
fructose	57-48-7	-	1 M	
trehalose	6138-23-4	-	1 M	
sucrose	57-50-1	-	1 M	
caffeine	58-08-2	-	15 mg/ml	
denatonium	3734-33-6	-	10 mM	
arginine	74-79-3	-	100 mM	
histidine	71-00-1	-	100 mM	
lysine	56-87-1	-	100 mM	
aspartic acid	56-84-8	-	100 mM	
glutamic acid	56-86-0	-	100 mM	
acetic acid	64-19-7	-	1% (v/v)	
NaCl (high)	7647-14-5	-	1 M	
NaCl (low)	7647-14-5	-	10 mM	
NaHCO <sub>3</sub> (pH 5)	144-55-8	-	200 mM	0.5 ml of 200 mM NaHCO <sub>3</sub> (pH 6.5) + 50 µl of 5 M H <sub>2</sub> PO <sub>4</sub>
CsHCO <sub>3</sub> (pH 7)	29703-01-3	-	200 mM	pH set with HCl immediately before use
NaH <sub>2</sub> PO <sub>4</sub>	13472-35-0	-	500 mM	
PBS pH 4	-	-	100%	
PBS pH 7	-	-	100%	
PBS pH 10	-	-	100%	7.8 mM NaH <sub>2</sub> PO <sub>4</sub> + 12.2 mM Na <sub>2</sub> HPO <sub>4</sub> + 153.8 mM NaCl (pH set with HCl or NaOH)
carbonated water	-	-	-	- Aproz ® (in mg/ml: Ca <sup>2+</sup> 360; Mg <sup>2+</sup> 70; Na <sup>+</sup> 6; K <sup>+</sup> 2.5; HCO <sub>3</sub> <sup>-</sup> 250; NO <sub>3</sub> <sup>-</sup> 1.5; SO <sub>4</sub> <sup>2-</sup> 930; SiO <sub>2</sub> 7) + gaseous CO <sub>2</sub>
non-carbonated water	-	-	-	- Aproz ® (in mg/ml: Ca <sup>2+</sup> 360; Mg <sup>2+</sup> 70; Na <sup>+</sup> 6; K <sup>+</sup> 2.5; HCO <sub>3</sub> <sup>-</sup> 250; NO <sub>3</sub> <sup>-</sup> 1.5; SO <sub>4</sub> <sup>2-</sup> 930; SiO <sub>2</sub> 7)
hexanoic acid	142-62-1	-	-	-

## Supplementary Table 2. Gustatory stimuli

Oligonucleotide	Sequence (5'-3')	Notes
CRISPRsgR	AAAGCACCGACTCGGTGCCACTTTTTCAGATTGATAACGGACTAGCCCTTATTTAACTTGGCTATTTCTAGCTCTAAAAC	
CRISPRsg F- <i>lr56d</i> <sup>1</sup>	GAAATTAATACGACTCACATAAGGTCATCACGAGCGCCATGTGTTTATAGCTAGAAAATAGC	T7 promoter sequence is italicised; <i>lr56d</i> target sequence is underlined
CRISPRsg F- <i>lr56d</i> <sup>2</sup>	GAAATTAATACGACTCACATAAGGTCATCACGACTAGCTATCCCATGTTTATAGCTAGAAAATAGC	T7 promoter sequence is italicised; <i>lr56d</i> target sequence is underlined
<i>lr56d</i> sgRNAs-fwd	GCGCCCGGGTTGGATTCGCCCGCGATGCCAAATATGGATAATCGTCTGGTTTTAGAGCTAGAAAATAGCAAG	sgRNA cloning into <i>pCFD5</i>
<i>lr56d</i> sgRNAs-rev	ATTTTAACTTGGCTATTCTAGCTCTAAAAGCAAGCCAGATGGTTCTCTGCACCGCCGGGAATCGAACCC	sgRNA cloning into <i>pCFD5</i>
<i>lr56d</i> Gal4-HA1-fwd	GATCCACCTGCCGATCTCGCCCGCCACGCACTGTGCATCCCTTGAAGTGC	Homology Arm 1-Gal4 ORF fusion
<i>lr56d</i> Gal4-HA1-rev	GATCCACCTGCCGATCTCGCCCGCCACGCACTGTGCATCCCTTGAAGTGC	Homology Arm 1-Gal4 ORF fusion
<i>lr56d</i> Gal4-HA1-internal	ACTGGCAGTCGCCCGTACAAAATATGAAGCTACTGTCTTCTATCGAACCAAGC	Homology Arm 1-Gal4 ORF fusion
<i>lr56d</i> Gal4-HA1-internal	CGATAGAACACAGTAGCTTCATATTTGTACGGCGACTGCCAGTGGGTAAAC	Homology Arm 1-Gal4 ORF fusion
<i>lr56d</i> -HA2-fwd	GATCCGCTCTCGTATAGCCGAGATCGTTTCTCAGCGCTTCATG	Homology Arm 2
<i>lr56d</i> -HA2-rev	GATCCGCTCTCGGACGATGCCCTTCCAAATTGATACGTGAACG	Homology Arm 2

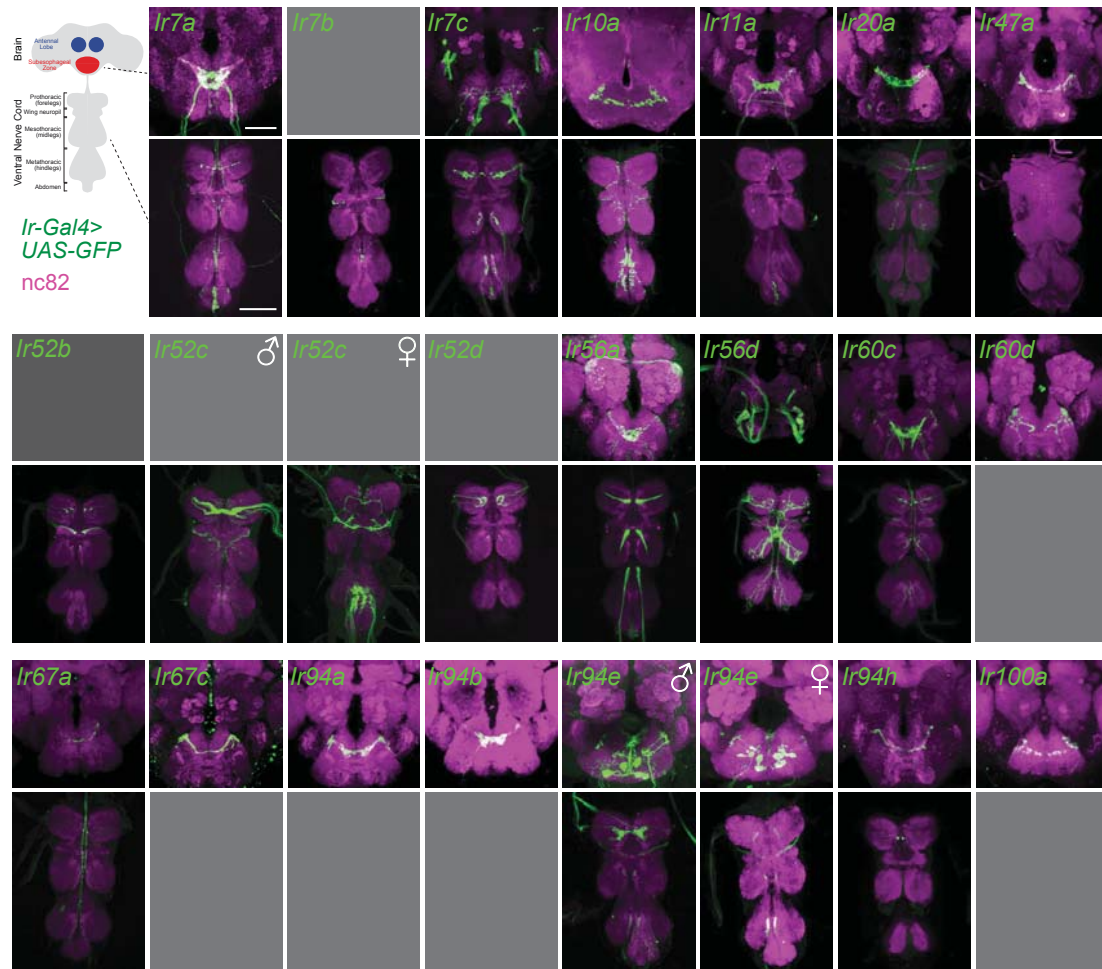
Supplementary Table 3. Oligonucleotide sequences used for CRISPR/Cas9-mediated editing of the *lr56d* locus



**Supplementary Figure 1. Projections of *Ir-Gal4* expressing sensory neurons in the larval central nervous system**

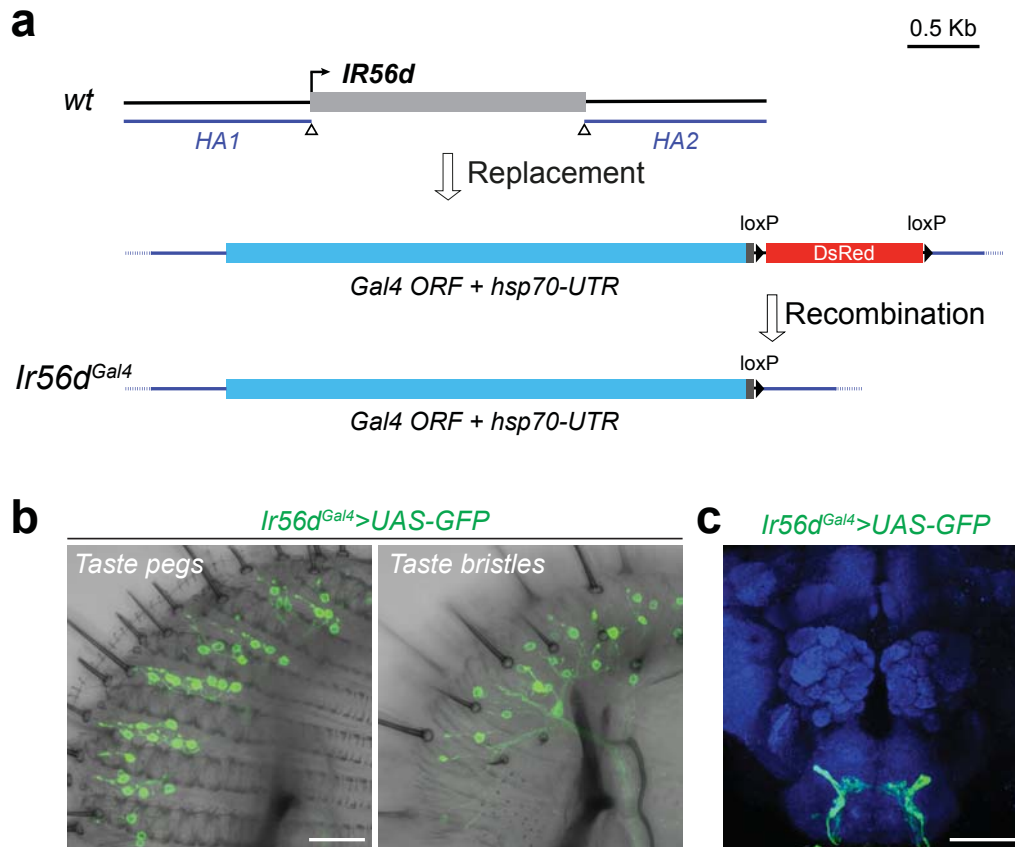
Immunofluorescence with anti-GFP (green) and nc82 (magenta) on whole-mount brains of third instar larvae, revealing the projection patterns of *Ir-Gal4*-expressing neuron populations in the brain and ventral nerve cord (as schematised in the cartoon at the top left). Images for *Ir7a*, *Ir7b*, *Ir7g*, *Ir94e* and *Ir94h* drivers have been adapted from Croset et al., 2016<sup>28</sup>. Genotypes are of the form: *w;UAS-mCD8:GFP;IrX-Gal4*. SEZ: Subesophageal Zone; VNC: Ventral Nerve Cord. Scale bars: 20  $\mu$ m.





**Supplementary Figure 2. Projections of *Ir-Gal4* expressing sensory neurons in the adult central nervous system**

Immunofluorescence with anti-GFP (green) and nc82 (magenta) on whole-mount adult brains and ventral nerve cords (schematised top left), revealing the projection patterns of *Ir-Gal4*-expressing neuron populations. Grey panels indicate no expression was detected for that Gal4 driver. Genotypes are of the form: *UAS-mCD8:GFP;IrX-Gal4*. Scale bars: 50  $\mu\text{m}$  (brain), 100  $\mu\text{m}$  (ventral nerve cord).

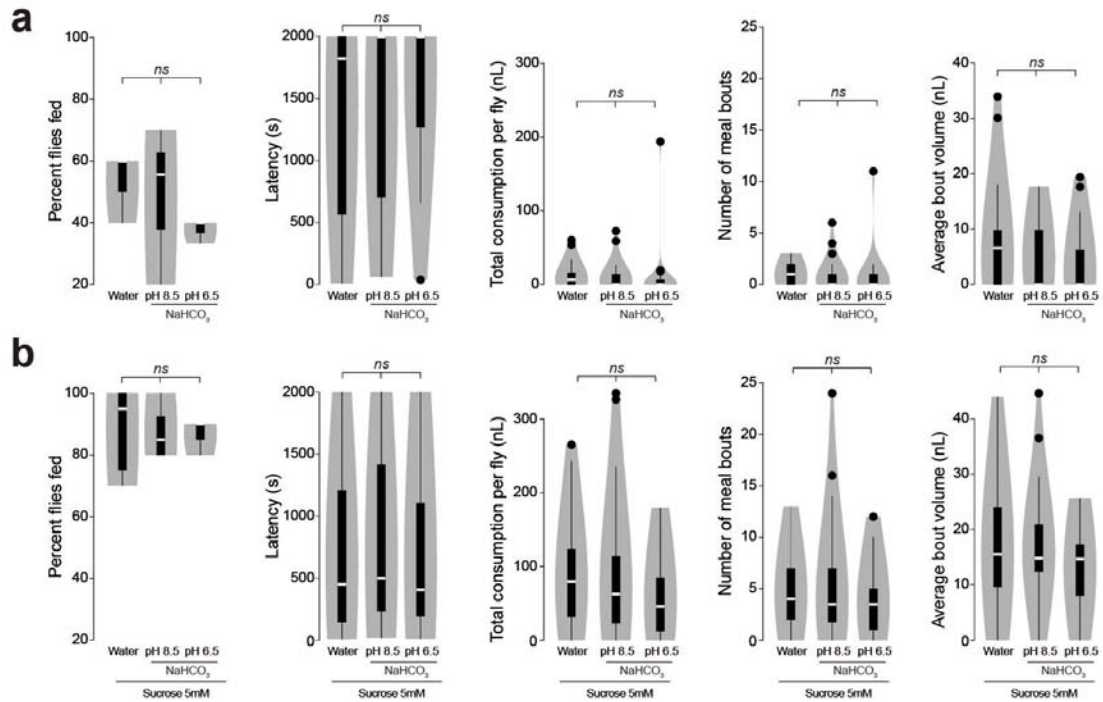


**Supplementary Figure 3. An *Ir56d<sup>Gal4</sup>* allele shows a similar expression pattern to *Ir56d-Gal4***

(a) Schematic representing the generation of the *Ir56d<sup>Gal4</sup>* allele. The entire *Ir56d* exon was substituted with the *Gal4* sequence. In an additional step, the *DsRed* marker used for screening of positive recombination events was removed with Cre recombinase. HA: Homology Arm. Unfilled arrowheads point to the sgRNA targets.

(b) Immunofluorescence with anti-GFP (green), overlaid on a bright-field image, of a whole mount proboscis of a *w;Ir56d<sup>Gal4</sup>;UAS-mCD8:GFP* animal. Scale bar: 25  $\mu$ m.

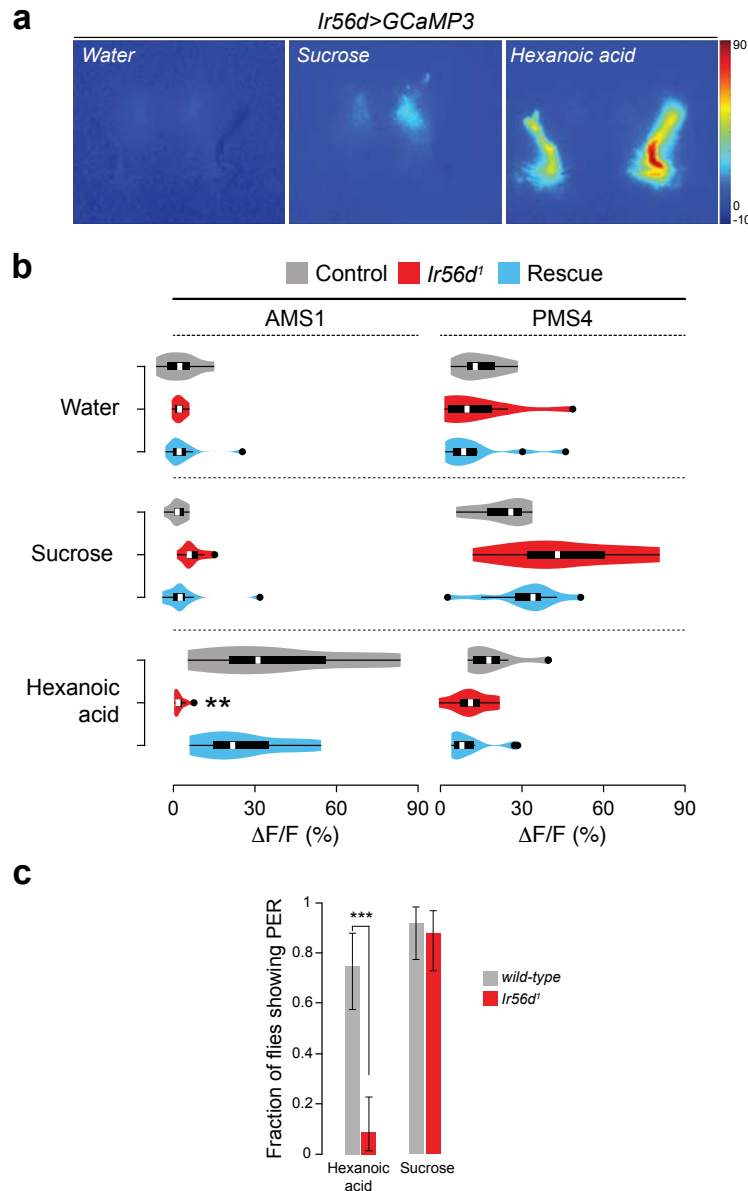
(c) Immunofluorescence with anti-GFP (green) and anti-nc82 (blue) on a whole mount brain of a *w;Ir56d<sup>Gal4</sup>;UAS-mCD8:GFP* animal. Scale bar: 50  $\mu$ m (right).



#### Supplementary Figure 4. Analysis of carbonation-evoked feeding by Expresso

(a) Feeding success, latency to first bout, total consumption per fly, number of meal bouts and average bout volume for male  $w^{1118}$  flies ( $n=30$  per tastant) feeding from water, 100 mM NaHCO<sub>3</sub> pH 8.5 or 100 mM NaHCO<sub>3</sub> pH 6.5 solutions.

(b) Feeding success, latency to first bout, total consumption per fly, number of meal bouts and average bout volume for male  $w^{1118}$  flies ( $n=30$  per tastant) feeding from solutions containing 5 mM sucrose, 5 mM sucrose + 100 mM NaHCO<sub>3</sub> pH 8.5 or 5 mM sucrose + 100 mM NaHCO<sub>3</sub> pH 6.5 solutions. *ns*: non-significant (pairwise comparisons using Tukey and Kramer (Nemenyi) test with Tukey-Dist approximation for independent samples).



### Supplementary Figure 5. IR56d is required for physiological and behavioural responses to hexanoic acid

(a) Colour-coded images (reflecting the maximal GCaMP3 fluorescence intensity changes) in *w;Bl/+;UAS-GCaMP3,Ir56d-Gal4/+* animals stimulated with distilled water, 1 M sucrose and 1% (v/v) hexanoic acid.

(b) Quantification of changes in GCaMP3 fluorescence ( $\Delta F/F$ ) upon stimulation with the indicated chemicals (as in (a)) to the proboscis of the flies. Genotypes: Control: *w;Bl/+;UAS-GCaMP3,Ir56d-Gal4/+* (n=7); Mutant: *w;Ir56d<sup>1</sup>/Ir56d<sup>1</sup>;UAS-GCaMP3,Ir56d-Gal4/+* (n=8); Rescue: *w;Ir56d<sup>1</sup>,UAS-Ir56d/Ir56d<sup>1</sup>;UAS-GCaMP3,Ir56d-Gal4/+* (n=11). \*P<0.05, \*\*P<0.01, \*\*\*P<0.001 (Wilcoxon rank sum test with Bonferroni correction for multiple comparisons).

(c) Fraction of *w<sup>1118</sup>* (n=36) or *Ir56d<sup>1</sup>* mutant (n=33) flies showing proboscis extension reflex (PER) to 1% (v/v) hexanoic acid and 100 mM sucrose. Error bars represent the  $\pm 95\%$  binomial confidence intervals \*\*\*P<0.001 (Fisher exact test).