LETTERS

Natural variation in a neural globin tunes oxygen sensing in wild *Caenorhabditis elegans*

Annelie Persson[1,2]*, Einav Gross[1]*, Patrick Laurent[1], Karl Emanuel Busch[1], Hugo Bretes[1] & Mario de Bono[1]

Behaviours evolve by iterations of natural selection, but we have few insights into the molecular and neural mechanisms involved. Here we show that some *Caenorhabditis elegans* wild strains switch between two foraging behaviours in response to subtle changes in ambient oxygen. This finely tuned switch is conferred by a naturally variable hexacoordinated globin, GLB-5. GLB-5 acts with the atypical soluble guanylate cyclases[1–3], which are a different type of oxygen binding protein, to tune the dynamic range of oxygen-sensing neurons close to atmospheric (21%) concentrations. Calcium imaging indicates that one group of these neurons is activated when oxygen rises towards 21%, and is inhibited as oxygen drops below 21%. The soluble guanylate cyclase GCY-35 is required for high oxygen to activate the neurons; GLB-5 provides inhibitory input when oxygen decreases below 21%. Together, these oxygen binding proteins tune neuronal and behavioural responses to a narrow oxygen concentration range close to atmospheric levels. The effect of the *glb-5* gene on oxygen sensing and foraging is modified by the naturally variable neuropeptide receptor *npr-1* (refs 4, 5), providing insights into how polygenic variation reshapes neural circuit function.

Behaviour can evolve rapidly. Even within species different populations can show marked behavioural divergence[6]. Such variation is typically genetically complex[7,8]. The human genetic pool contains millions of polymorphisms[9], and twin studies suggest that human behavioural variation usually has a genetic component[10]. However, unambiguously linking behavioural variation to DNA polymorphisms remains a challenge in any animal.

C. elegans strains collected worldwide can show heritable foraging differences[4,11]. This is partly due to variation in a neuropeptide Y-like receptor, NPR-1 (ref. 4). Some strains such as N2 Bristol encode NPR-1 215V, with valine at position 215; others, such as CB4856 Hawaii, encode the less potent NPR-1 215F isoform. NPR-1 modulates the AQR, PQR and URX neurons, which mediate avoidance of high O_2 (21–11%): *C. elegans* bearing *npr-1 215V* do not avoid high O_2 environments while feeding, whereas strains bearing *npr-1 215F* do[2,3].

Air contains 21% O_2 at the Earth's surface, but less in buried spaces and rotting material owing to respiration[12]. We speculated that 21% O_2 is a salient reference for *C. elegans*, enabling surface avoidance and accumulation on bacteria. To test this we examined how wild strains feeding on *Escherichia coli* responded to subtle O_2 shifts, from 21% to 19.2% or 17.4%. As expected[3], N2 Bristol animals bearing *npr-1 215V* moved slowly on food at all three O_2 concentrations (Fig. 1a). In contrast, 20 wild strains bearing *npr-1 215F* moved rapidly in 21% O_2 but markedly decelerated at 19.2% or 17.4%. Interestingly, this slowing was not observed in AX613, a strain bearing *npr-1 215F* in a Bristol background, (Fig. 1a), indicating that natural variation in genes other than *npr-1* tunes O_2 responses. The slowing response was weaker when food was absent (Fig. 1b), suggesting that it is part of a foraging strategy.

To investigate this variation, we crossed CB4856 Hawaii and AX613 animals. The F_1 cross-progeny showed Hawaiian-like O_2 responses, indicating that this trait was dominant (Fig. 1c). We next created recombinant inbred lines from crosses between CB4856 and AX613, examined their responses to shifts from 21% to 17.4% O_2, and genotyped them for single nucleotide polymorphisms (SNPs) between the Hawaiian and Bristol strains[13]. These experiments highlighted an interval on chromosome V that, when derived from CB4856 Hawaii, conferred Hawaiian-like responses. By selecting recombinants in this interval we mapped the trait to 8 kilobases (kb) of Hawaiian DNA (Fig. 2a, b and Supplementary Fig. 2). This interval contained two genes, one of which, *glb-5*, had similarity to globins. Comparison of
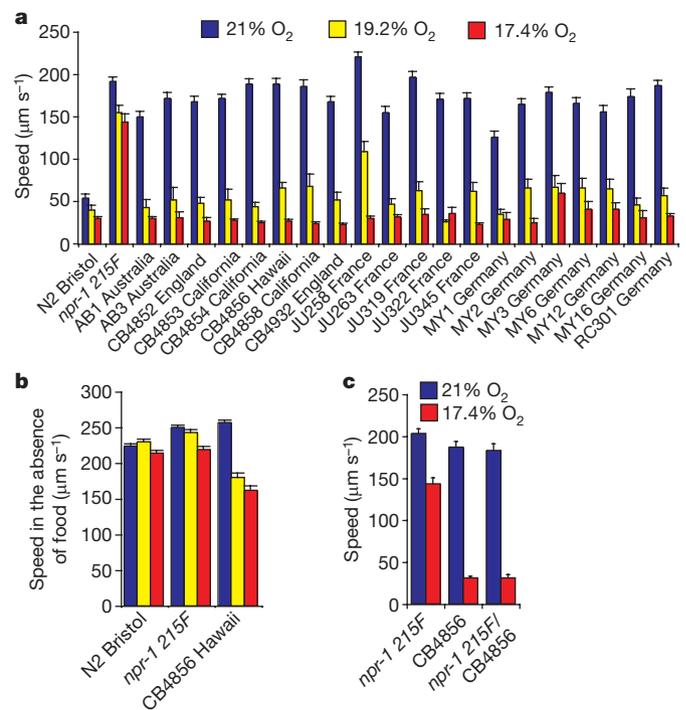

Figure 1 | Polygenic natural variation tunes O_2 responses in wild *C. elegans*. a, Feeding *C. elegans* wild strains bearing *npr-1 215F* switch from roaming to dwelling when O_2 drops just below 21%. This response was absent in AX613, a laboratory strain bearing *npr-1 215F* in an N2 Bristol background. Apart from N2 and AX613, all strains shown are wild isolates bearing *npr-1 215F*. In this and subsequent figures 'speed' is the average speed during a 2-min period in which O_2 levels have stabilized (see Supplementary Fig. 1). Error bars indicate s.e.m. **b**, When food is absent, CB4856 Hawaii animals only weakly reduce movement in response to small drops in O_2. The colour key is as in **a**. **c**, F_1 progeny of a cross between CB4856 Hawaii males and AX613 display Hawaiian-like responses.

[1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. [2]Department of Cell and Molecular Biology, Göteborg University, 405 30 Göteborg, Sweden.
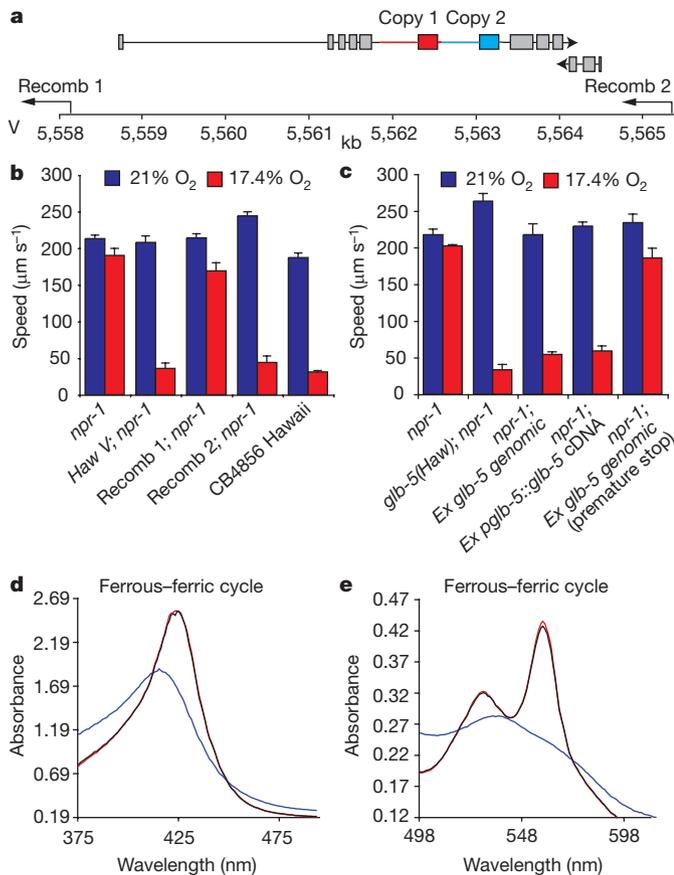*These authors contributed equally to this work.

Figure 2 | Natural variation in a hexacoordinated globin sculpts *C. elegans* O$_2$ responses. a, Recombinants (recomb) 1 and 2 map variation in the O$_2$ response to an 8-kb interval containing *glb-5*. Arrows indicate CB4856 sequences in each recombinant chromosome. Bristol *glb-5* is partially duplicated (two copies shown in blue and red) compared to the Hawaiian allele. Numbering refers to the position on chromosome V. **b**, Recombinant 1 does not display Hawaiian-like O$_2$ responses whereas recombinant 2 does. **c**, Transgenes of *glb-5*(*Haw*) confer Hawaiian-like responses to *npr-1*(*ad609*) mutants. Error bars indicate s.e.m. **d, e**, GLB-5 absorbance spectra. The blue line represents ferric GLB-5; red and black lines represent ferrous GLB-5 at the beginning and end of a ferrous–ferric–ferrous oxidation–reduction cycle.

N2 and CB4856 sequences in the 8 kb revealed 11 polymorphisms, all in *glb-5*. Ten SNPs altered introns; the remaining polymorphism partially duplicated *glb-5* in N2 but not in CB4856 (Fig. 2a). We repeatedly found N2 versions of the ten intronic SNPs in wild strains exhibiting Hawaiian-like O$_2$ responses (Fig. 1a and Supplementary Table 2); by contrast, all 20 wild strains with Hawaiian-like O$_2$ responses carried unduplicated *glb-5* (Fig. 1a and Supplementary Table 1). These results indicate that variation in *glb-5* reconfigures O$_2$ responses in wild *C. elegans*, and that the causal polymorphism is the duplication. We confirmed that the Hawaiian *glb-5* allele, *glb-5*(*Haw*), was sufficient to confer Hawaiian-like O$_2$ responses by expressing this variant from its endogenous promoter in N2 animals defective in *npr-1* (Fig. 2c).

The N2 allele of *glb-5*, *glb-5*(*Bri*), was recessive to *glb-5*(*Haw*) (Fig. 1c). To investigate why, we isolated complementary DNAs from both alleles. Whereas PCR with reverse transcription (RT–PCR) from CB4856 RNA yielded one *glb-5* cDNA product, N2 RNA yielded two additional products encoding truncated proteins (Supplementary Fig. 3). This suggests that the duplication in *glb-5*(*Bri*) reduces gene function by interfering with correct splicing.

Globins vary considerably in sequence but typically conserve the proximal and distal histidines in helices F and E, the phenylalanine in the CD region, and the proline in helix C[14]. GLB-5 conserves all these residues (Supplementary Fig. 4). To establish GLB-5 as a haem-binding globin, we expressed it in *E. coli* tagged with maltose binding protein (MBP) or polyhistidine. MBP–GLB-5 could be purified as a soluble

red-brown protein (Supplementary Fig. 5a); His-tagged GLB-5 localized to inclusion bodies but could be reconstituted with haem *in vitro* (Supplementary Fig. 5b, c). For both proteins, the absorption spectrum of the deoxygenated ferrous form resembled that of hexa-coordinated globins such as human neuroglobin and cytoglobin[15], with absorption peaks at 423, 559 and 529 nm (Soret, α and β band peaks, respectively) (Fig. 2d, e and Supplementary Fig. 5c). Like mammalian neuroglobins, GLB-5 rapidly oxidized to a ferric form after binding O$_2$, but could be reduced to its ferrous state by NADH enzymatic reduction[16]. These data suggest GLB-5 is a hexacoordinated globin that reversibly binds O$_2$.

To determine where *glb-5* is expressed, we fused DNA encoding mCherry fluorescent protein upstream of the *glb-5* stop codon in a genomic fragment, and introduced the construct into *npr-1* mutants. The transgene conferred Hawaiian-like O$_2$ responses to these animals (Figs 2c and 3e, f), and gave strong fluorescence in six neurons: URXL/R, AQR/PQR and BAGL/R (Fig. 3a–d). Like AQR, PQR and URX, the BAG neurons may be involved in O$_2$ sensing as they express the atypical soluble guanylate cyclases GCY-31 and GCY-33 (ref. 17)
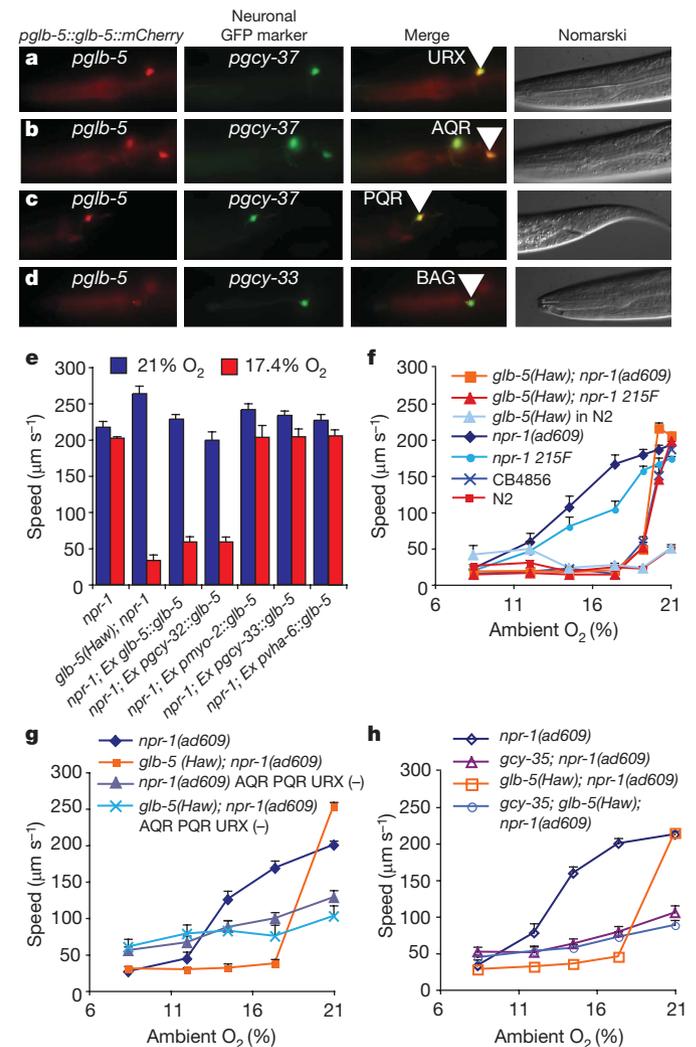

Figure 3 | *glb-5* acts in O$_2$ sensing neurons. a–d, *glb-5* is expressed in the URX (**a**), AQR (**b**), PQR (**c**) and BAG (**d**) neurons. Original magnification for all images, ×100. **e**, Selective expression of *glb-5*(*Haw*) in the AQR, PQR and URX neurons sensitizes *npr-1* animals to O$_2$ variation around 21%. *pgcy-32* drives expression in the AQR, PQR and URX neurons; *pgcy-33* in BAG, *pmyo-2* in the pharynx, and *pvha-6* in the intestine. **f**, *npr-1* and *glb-5* natural alleles interact to sculpt O$_2$ responses. All strains except CB4856 have an N2 genetic background. **g, h**, Ablating the AQR, PQR and URX neurons or mutating *gcy-35* disrupts the affect of *glb-5*(*Haw*) on O$_2$ responses. All error bars indicate s.e.m.

(Wormbase). We also observed fluorescence in the pharynx, intestine, coelomocytes, uv2 vulval cells and the excretory cell (data not shown).

To establish where GLB-5 functions, we expressed its cDNA specifically in the AQR, PQR and URX neurons, in the BAG neurons, in the pharynx, or in the intestine. We monitored transgene expression by co-expressing green fluorescent protein (GFP) or mCherry in a polycistronic message. Expression in the AQR, PQR and URX neurons conferred Hawaiian-like slowing responses to *npr-1* mutants (Fig. 3e). Expression in other cells did not have significant effects in our assay. These data suggest *glb-5* acts in the AQR, PQR and URX neurons to tune O$_2$ responses.

We next investigated how different allelic combinations of *glb-5* and *npr-1* alter locomotory behaviour across a broad range of O$_2$ tensions, from 21% to 8.4% (Fig. 3f). To maintain a constant genetic background we backcrossed all Hawaiian alleles ten times into N2. For comparison we included the strong loss-of-function allele *npr-1(ad609)*. Animals bearing *npr-1 215V* settled on food, irrespective of O$_2$ concentration and *glb-5* genotype. In contrast animals bearing *npr-1 215F* roamed in 21% O$_2$. However whereas *glb-5(Haw); npr-1 215F* and *glb-5(Haw); npr-1(ad609)* animals suppressed movement at 19.2% O$_2$, *glb-5(Bri); npr-1 215F* animals only reduced movement gradually as O$_2$ dropped. Thus *npr-1 215F* allows feeding animals to increase movement when O$_2$ rises, whereas *glb-5(Haw)* permits sharp suppression of movement when O$_2$ falls just below 21%.

To probe the role of the AQR, PQR and URX neurons in regulating locomotion we selectively ablated them using the cell death gene *egl-1* (ref. 18). Ablated *npr-1(ad609)* animals displayed intermediate locomotory activity on food at 21% O$_2$, irrespective of *glb-5* genotype, and weak slowing as O$_2$ fell to 8.4% (Fig. 3g). These data confirm that the AQR, PQR and URX neurons regulate locomotory activity, but suggest that other O$_2$-sensing neurons also contribute.

Apart from expressing GLB-5, the AQR, PQR and URX neurons express putative atypical soluble guanylate cyclases composed of GCY-35 and GCY-36 subunits that seem to be O$_2$ sensors[1–3,17,19]. We investigated their relationship to *glb-5* by studying single and multiple mutants. Irrespective of *glb-5* genotype, *npr-1* animals lacking *gcy-35*, *gcy-36* or both behaved like animals lacking AQR, PQR and URX neurons (Fig. 3h and data not shown). Thus, *gcy-35* and *gcy-36* seem to be required for the *glb-5(Haw)* allele to exert its affect on locomotory behaviour in our model.

To study the physiological responses of the AQR, PQR and URX neurons to O$_2$ we developed a microfluidic device that allowing gas stimuli to be switched in less than 3 s (Fig. 4a, b and Supplementary Fig. 6). We placed the device over immobilized *npr-1(ad609)* animals expressing the ratiometric Ca^{2+} sensor cameleon YC3.60 in AQR, PQR and URX, and measured yellow fluorescent protein (YFP) to cyan fluorescent protein (CFP) emission ratios while delivering different O$_2$ stimuli. All three neurons responded to a rise from 7 to 21% O$_2$ with an increase in the YFP/CFP ratio, indicative of Ca^{2+} influx (Fig. 4c–e). In URX neurons cameleon reported a Ca^{2+} spike followed by a smaller Ca^{2+} plateau that perdured while O$_2$ was at 21%. In AQR and PQR neurons the strong Ca^{2+} rise plateaued and remained high for the 2 min O$_2$ remained at 21% (Fig. 4d, e). O$_2$ responses were graded: higher O$_2$ elicited higher ratio changes (Fig. 4g). Furthermore, when O$_2$ levels decreased all three neurons responded with a drop in Ca^{2+} (Fig. 4c–e).

Loss of the *gcy-35* soluble guanylate cyclase or of *tax-4*, which encodes a cyclic-GMP-gated ion channel subunit expressed in AQR, PQR and URX[2,5], disrupted Ca^{2+} responses to both rises and falls in O$_2$ (Fig. 4f and data not shown). These results indicate that rising O$_2$ activates soluble guanylate cyclases, leading to increased cGMP and depolarization through the gating of a cGMP channel that includes TAX-4.

To study how natural variation in *glb-5* alters neuronal O$_2$ responses we gave *glb-5(Haw); npr-1* and *glb-5(Bri); npr-1* animals expressing the same cameleon transgene a series of O$_2$ stimuli and compared ratio changes in the PQR neurons between the two strains. Responses to switches from 21% to 11% or 7% O$_2$ were not significantly different
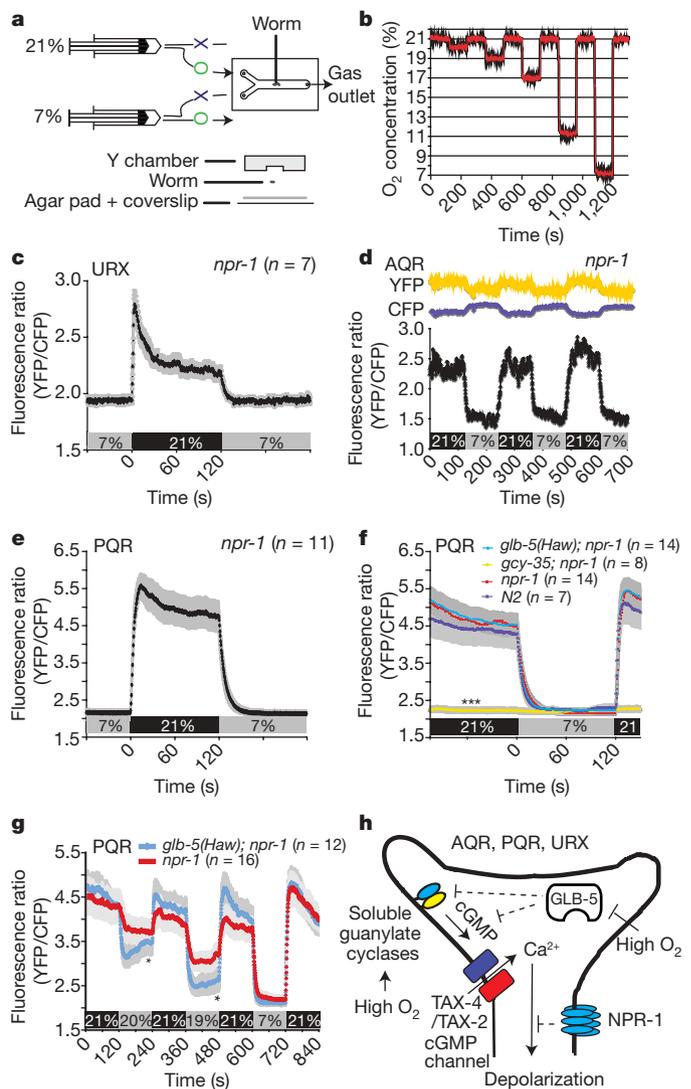

Figure 4 | Responses of O$_2$ sensing neurons. a, Microfluidic chamber used to deliver gas stimuli. O, valve open; X, valve closed. **b**, O$_2$ stimuli measured in the chamber using an optode. Two traces, red and black, are overlayed to indicate stimulus reproducibility. **c–g**, Ca^{2+} transients in response to O$_2$ stimuli measured using cameleon YC3.60. Grey error bars represent s.e.m. **c**, Averaged Ca^{2+} transients in URX neurons. **d**, Representative trace showing Ca^{2+} transients in AQR and indicating reciprocal changes in YFP and CFP fluorescence intensity. **e**, Averaged Ca^{2+} transients in PQR neurons. **f**, Mutations in *gcy-35* disrupt Ca^{2+} transients in PQR. **g**, *glb-5(Haw); npr-1* animals respond more sharply when O$_2$ drops slightly below 21% than control *npr-1* animals bearing *glb-5(Bri)*. *$P < 0.05$; ***$P < 0.001$. **h**, Model for O$_2$ sensing in AQR, PQR and URX neurons.

(Fig. 4f, g and data not shown). However, Ca^{2+} fell significantly more in *glb-1(Haw); npr-1* animals than in *glb-5(Bri); npr-1* animals when they experienced a switch from 21–20% or 21–19% O$_2$. These data mirror our behavioural results and suggest that high GLB-5 activity inhibits neuronal activation when O$_2$ falls below 21%, tuning the dynamic range of PQR to a narrow interval just below 21% O$_2$.

To investigate the evolutionary history of *glb-5* we examined its orthologues in *Caenorhabditis brenneri*, *Caenorhabditis briggsae* and *Caenorhabditis remanei* (Supplementary Table 3). All encoded unduplicated *glb-5* genes, suggesting that *glb-5(Bri)* arose in *C. elegans* from *glb-5(Haw)* as a reduction-of-function mutation. These *Caenorhabditis* species, as well as the as yet unnamed species *C. sp3* and *C. spn.* (see Wormbase), also reduced movement when shifted from 21% to 17.4% O$_2$ (Supplementary Fig. 7a).

To determine the distribution of *glb-5* and *npr-1* allelic combinations in natural *C. elegans* populations, we genotyped 98 wild isolates

(Supplementary Fig. 7b and Supplementary Table 1). Ninety had the Hawaiian-like combination, 7 were Bristol-like and one encoded *glb-5(Haw)*; *npr-1 215V*. None bore the *glb-5(Bri)*; *npr-1 215F* combination. Because *glb-5* and *npr-1* are on different chromosomes, these co-inheritance patterns could reflect low interbreeding between Bristol-like and Hawaiian-like foraging types, co-selection of *glb-5* and *npr-1* alleles, or both.

To conclude, vertebrate and invertebrate genomes encode hexacoordinated globins, but their function is mysterious. GLB-5, a *C. elegans* hexacoordinated globin, participates in O_2 sensing. We propose that high O_2 stimulates cGMP production from atypical soluble guanylate cyclases expressed in AQR, PQR and URX neurons, activating cGMP-gated ion channels and depolarizing these neurons (Fig. 4h). A different type of O_2 sensor, GLB-5, inhibits these neurons when O_2 falls below 21%. A similar system of multiple sensors may generate the sharply tuned O_2 responses observed in different mammalian tissues[20]. It will be interesting to examine whether mammalian neuroglobin or cytoglobin act as O_2 sensors and regulators of cGMP signalling.

Globins could regulate O_2 sensing as O_2 buffers, O_2 sinks or as signalling molecules. It seems unlikely that GLB-5 acts as an O_2 buffer in AQR, PQR and URX neurons as this would dull, not sharpen, O_2 responsiveness. GLB-5 is also unlikely to act as a passive O_2 sink, because diffusion from the atmosphere should render the sink ineffective. We favour a signalling model in which varying O_2 triggers conformational changes that modulate downstream signalling molecules. A precedent exists in O_2-sensing bacterial globins[21]. Interestingly, vertebrate neuroglobin binds to G proteins in its ferric state[22,23]. GLB-5 has an extra domain apart from the globin domain that could act as an adaptor, specify subcellular localization, or transduce conformational changes. Notably, the *C. elegans* genome encodes 33 putative globins[24]. Some of these are neurally expressed, potentially adding further complexity to the O_2 responses of *C. elegans*.

Natural variation in *glb-5* and *npr-1* modulate the same neurons in different ways. Iterative modification of the same neural circuit may be common in evolution. By reconfiguring responses to food and O_2 these loci may alter habitat preferences, resulting in balancing selection maintaining different feeding types.

METHODS SUMMARY

Strains. Except were indicated, nematodes were grown under standard conditions[25]. Strains used are listed in Methods. We refer to the Hawaiian allele of *glb-5* as *glb-5(Haw)* for clarity, but its formal allele designation is *glb-5(db200)*.

Behaviour. Assays of locomotory activity were carried out as described[3,19]. Oxygen was monitored using a microsensor or an optode (Presens). Behavioural data represent the mean of >three assays, carried out on three different days. Significance was determined using the two-tailed *t*-test.

Biochemistry and spectroscopy. Expression and purification of GLB-5 are described in Methods. In brief, His$_6$- or MBP-tagged GLB-5 was expressed in *E. coli* and purified to >95% purity and absorption spectra measured using a Varian Cary 50 Bio spectrophotometer.

Calcium imaging. A *pgcy-32::yellow cameleon 3.60* transgene was used for ratiometric imaging of relative $[Ca^{2+}]$ in the URX, AQR and PQR cell bodies. General guidelines for image acquisition were as described[26]. Specific guidelines for image acquisition and data analysis are described in Methods. Adults (24 h post-L4 stage) were immobilized using surgical glue (Nexaband S/C, Abbott Labs). Animals were covered by the long arm of the microfluidic Y-chamber (Fig. 4a) and the desired $O_2:N_2$ mixture was pumped in at a constant flow of $1.45\,ml\,min^{-1}$. All recordings started within 5 min of animals being glued, and 1 min after the start of pumping.

Molecular biology. General molecular manipulations followed standard protocols[27]. Details for plasmid construction are in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Cheung, B. H., Arellano-Carbajal, F., Rybicki, I. & De Bono, M. Soluble guanylate cyclases act in neurons exposed to the body fluid to promote *C. elegans* aggregation behavior. *Curr. Biol.* **14**, 1105–1111 (2004).
2. Gray, J. M. *et al.* Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. *Nature* **430**, 317–322 (2004).
3. Cheung, B. H., Cohen, M., Rogers, C., Albayram, O. & de Bono, M. Experience-dependent modulation of *C. elegans* behavior by ambient oxygen. *Curr. Biol.* **15**, 905–917 (2005).
4. de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**, 679–689 (1998).
5. Coates, J. C. & de Bono, M. Antagonistic pathways in neurons exposed to body fluid regulate social feeding in *Caenorhabditis elegans*. *Nature* **419**, 925–929 (2002).
6. Korol, A., Rashkovetsky, E., Iliadi, K. & Nevo, E. *Drosophila* flies in "Evolution Canyon" as a model for incipient sympatric speciation. *Proc. Natl Acad. Sci. USA* **103**, 18184–18189 (2006).
7. Valdar, W. *et al.* Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nature Genet.* **38**, 879–887 (2006).
8. Anholt, R. R. & Mackay, T. F. Quantitative genetic analyses of complex behaviours in *Drosophila*. *Nature Rev. Genet.* **5**, 838–849 (2004).
9. Sachidanandam, R. *et al.* A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).
10. Bouchard, T. J. J. & McGue, M. Genetic and environmental influences on human psychological differences. *J. Neurobiol.* **54**, 4–45 (2003).
11. Hodgkin, J. & Doniach, T. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**, 149–164 (1997).
12. Baumgartl, H., Kritzler, K., Zimelka, W. & Zinkler, D. Local PO2 measurements in the environment of submerged soil microarthropods. *Acta Oecol.* **15**, 781–789 (1994).
13. Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H. & Plasterk, R. H. Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nature Genet.* **28**, 160–164 (2001).
14. Ota, M., Isogai, Y. & Nishikawa, K. Structural requirement of highly-conserved residues in globins. *FEBS Lett.* **415**, 129–133 (1997).
15. Dewilde, S. *et al.* Biochemical characterization and ligand binding properties of neuroglobin, a novel member of the globin family. *J. Biol. Chem.* **276**, 38949–38955 (2001).
16. Trandafir, F. *et al.* Neuroglobin and cytoglobin as potential enzyme or substrate. *Gene* **398**, 103–113 (2007).
17. Yu, S., Avery, L., Baude, E. & Garbers, D. L. Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. *Proc. Natl Acad. Sci. USA* **94**, 3384–3387 (1997).
18. Chang, A. J., Chronis, N., Karow, D. S., Marletta, M. A. & Bargmann, C. I. A distributed chemosensory circuit for oxygen preference in *C. elegans*. *PLoS Biol.* **4**, e274 (2006).
19. Rogers, C., Persson, A., Cheung, B. & de Bono, M. Behavioral motifs and neural pathways coordinating O_2 responses and aggregation in *C. elegans*. *Curr. Biol.* **16**, 649–659 (2006).
20. Ward, J. P. Oxygen sensors in context. *Biochim. Biophys. Acta* **1777**, 1–14 (2008).
21. Hou, S. *et al.* Myoglobin-like aerotaxis transducers in Archaea and Bacteria. *Nature* **403**, 540–544 (2000).
22. Wakasugi, K., Nakano, T. & Morishima, I. Oxidized human neuroglobin acts as a heterotrimeric Gα protein guanine nucleotide dissociation inhibitor. *J. Biol. Chem.* **278**, 36505–36512 (2003).
23. Kitatsuji, C., Kurogochi, M., Nishimura, S., Ishimori, K. & Wakasugi, K. Molecular basis of guanine nucleotide dissociation inhibitor activity of human neuroglobin by chemical cross-linking and mass spectrometry. *J. Mol. Biol.* **368**, 150–160 (2007).
24. Hoogewijs, D. *et al.* Wide diversity in structure and expression profiles among members of the *Caenorhabditis elegans* globin protein family. *BMC Genomics* **8**, 356 (2007).
25. Sulston, J. & Hodgkin, J. in *The Nematode Caenorhabditis Elegans* (ed. Wood, W. B.) 587–606 (Cold Spring Harbour Laboratory Press, 1988).
26. Kerr, R. A. & Schafer, W. R. Intracellular Ca^{2+} imaging in *C. elegans*. *Methods Mol. Biol.* **351**, 253–264 (2006).
27. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbour Laboratory Press, 1989).

METHODS

Strains. Double-mutant strains were created by following visible phenotypes or by using PCR to confirm genotypes. Strains made or used in this study include: *C. elegans* wild isolates: See Supplementary Table 1. Non-*C. elegans* wild isolates: AF16, BW288, PS1185, PS1186, VT847, EM464, PB227, PB228, PB229, SB129, PS1010, RGD1, RGD2, JU272 and CB5161. Mutant strains: All strains are in the N2 Bristol background unless otherwise indicated. AX204, *npr-1(ad609) X*; AX215, *npr-1(ad609) lin-15(n765ts) X*; AX266, *unc-46(e177) dpy-11(e224) V, npr-1(ad609) X*; AX296, *dpy-11(e224) unc-76(e911) V, npr-1(ad609) X*; AX613, *npr-1(g320) X*; AX1891, *glb-5(Haw) V; npr-1(ad609) X*; AX1796, *glb-5(Haw) V; npr-1(g320) X*; AX1797, *glb-5(Haw) V; npr-1(ad609) gcy-36(db42) X*; AX1198, *gcy-35(ok769) I; npr-1(ad609) X*; AX1848, *gcy-35(ok765) I; glb-5 (Haw) V; npr-1(ad609) X*; AX1904, *npr-1(ad609) qaIs2241[pgcy-36::egl-1 +pgcy-35::gfp+lin-15(+)] X*; AX1918, *glb-5(Haw) V; npr-1(ad609) qaIs2241[pgcy-36::egl-1 +p gcy-35::gfp+lin-15(+)] X*. Transgenic strains: AX1911, *npr-1(ad609) lin-15(n765) X; dbEx [myo-2:glb-5(Haw)::polycis GFP + lin-15(+)]*; AX1912, *npr-1(ad609) lin-15(n765) X; dbEx[glb-5:glb-5(Haw)::polycis GFP + lin-15(+)]*; AX1913, *npr-1(ad609) lin-15(n765) X; dbEx[gcy-32:cglb-5(HAW)::polycis GFP, lin-15(+)]*; AX1914, *npr-1(ad609) lin-15(n765) X; dbEx[vha-6:genomic glb-5(Haw)::polycis mCherry + lin-15(+)]*; AX1916, *npr-1(ad609) lin-15(n765) X; dbEx[gcy-33:genomic glb-5(Haw)::polycis mCherry + lin-15(+)]*; AX1917, *npr-1(ad609) lin-15(n765) X; dbEx[glb-5:genomic glb-5(Haw) H144Stop::polycis mCherry + lin-15(+)]*; AX1846, *npr-1(ad609) lin-15(n765) X, dbEx[glb-5:genomic glb-5(Haw)::polycis mCherry + lin-15(+)]*. Imaging strains: AX1864, *npr-1(ad609) lin-15(n765ts) X; dbEx[pgcy-32::YC3.60 + lin-15(+)]*; AX1907, *dbEx[pgcy-32::YC3.60 + lin-15(+)]*; AX1908, *glb-5(Haw) V; npr-1(ad609) X; dbEx[pgcy-32::YC3.60 +lin-15(+)]*; AX1909, *gcy-35(ok765) I; npr-1(ad609) lin-15(n765ts) X; dbEx[pgcy-32::YC3.60 + lin-15(+)*; AX1910, *tax-4(p678) III; npr-1(ad609) lin-15(n765ts) X, dbEx[pgcy-32::YC3.60 + lin-15(+)]*.

Genetics. To map *glb-5(Haw)* finely, we picked 265 Unc-non-Dpy and Dpy-non-Unc recombinant progeny from *unc-46 glb-5(Bri) dpy-11/+(Haw) glb-5(Haw) +(Haw); npr-1(ad609)* animals and in the next generation selected animals homozygous for the recombinant chromosome. To determine whether the recombinants bore the dominant *glb-5(Haw)* allele we mated them with *npr-1(ad609)* males and assayed the responses of their progeny to a switch from 21% to 17.4% O_2. Recombinants were genotyped using SNPs between the N2 and CB4856 strains that are described in Wormbase (Fig. 2a, b and Supplementary Fig. 2).

Molecular biology. Genotyping natural isolates: The presence or absence of a duplication in *glb-5* was determined by examining the size of PCR fragments amplified with primers that flank the duplication. The genotype at *npr-1* was determined by amplifying the interval flanking the polymorphic *npr-1* codon 215, and sequencing.

glb-5(Bri) and *glb-5(Haw)* cDNA: We extracted total RNA from N2 and CB4856 animals by vortexing with Trizol and acid-washed glass beads. RNA was further purified using an RNeasy Midi Kit (Qiagen). cDNAs were generated by RT–PCR (OneStep RT–PCR kit, Qiagen). Oligonucleotides were designed on the basis of the predicted C18C4.1 gene sequences in Wormbase. Attempts to amplify the predicted full-length C18C4.1a gene product using 5′ oligonucleotides that anneal to various starting points in exon 1, failed. However, the full-length C18C4.1b gene product could be amplified readily. The resulting cDNAs were cloned into pDrive cloning vector (Qiagen) and sequenced.

Transgenes: Genomic DNA from CB4856 was used to amplify the *glb-5(Haw)* gene, including 3 kb of upstream sequences. The genomic fragment was cloned into pPD95.75 (A. Fire, personal communication), and then modified by adding an outron-mCherry fragment to make a polycistronic expression vector, as described[5]. This expression construct was modified to create other expression constructs. To create a *glb-5(Haw)* cDNA expression construct, the genomic part of the *glb-5(Haw)* gene was replaced with cDNA. To disrupt the *glb-5(Haw)* open reading frame we changed the conserved His in exon 6 into a stop codon, using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene). To create expression constructs that express *glb-5(Haw)* in the intestine or in BAG neurons, the promoter region of *glb-5(Haw)* was replaced with the 2.8-kb or the 1.0-kb promoter of *vha-6* (ref. 28) and *gcy-33*, respectively[17]. To drive expression of *glb-5(Haw)* in the AQR, PQR and URX neurons or in the pharynx, the coding region of *glb-5(Haw)* was inserted using Gateway into a GFP polycistronic expression vector that contained a 0.7-kb or 1.2-kb promoter region of *gcy-32* (ref. 17) or *myo-2* (ref. 29), respectively. Unless otherwise mentioned, 25–50 ng µl^{-1} of the above expression constructs was microinjected into the gonads of *npr-1(ad609) lin-15(n765ts) X* hermaphrodites, together with 50 ng µl^{-1} pJMZ-*lin-15*, following standard methods[30].

Biochemistry of GLB-5. cDNA for *glb-5(Haw)* was generated by RT–PCR from total RNA isolated from the CB4856 strain and cloned into the bacterial expression plasmids pET-28a (Novagen) and pMAL-c2X (NEB). The resulting constructs lacked the first two amino acids of GLB-5 (see Wormbase open reading frame C18C4.1a for the sequence) and had either a carboxy-terminal His$_6$ affinity tag or an amino-terminal MBP tag. The expression constructs were transformed into BL21-CodonPlus(DE3)-RIL (Stratagene) bacteria and fresh colonies used to inoculate M9-ZB medium containing 0.4% (w/v) glucose, 50 mg l^{-1} kanamycin, or 50 mg l^{-1} ampicillin and 30 mg l^{-1} chloramphenicol. Cultures were grown at 37 °C to an A_{600} optical density of 0.3–0.4, after which the temperature was reduced to 23 °C. At this stage, when expressing MBP-GLB-5, haemin chloride (Sigma) was added to a final concentration of 30 mg l^{-1}. When the A_{600} reached 0.7–0.8 the cultures were brought up to 0.1 mM IPTG. After a further 16–18 h, cells were collected by centrifugation and stored at −80 °C. To purify MBP-GLB-5, cells were thawed on ice, resuspended in lysis buffer (25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.5 mM EDTA, pH 7.4, and complete, EDTA-free protease inhibitor cocktail (Roche)), and treated twice with an Emulsiflex-C5 high-pressure homogenizer at 15,000–20,000 p.s.i. (pounds per square inch). The resulted cell lysate was centrifuged for 30 min at 33,000g and the soluble fraction was mixed with amylose resin (NEB) at 4 °C. After 2 h the resin was transferred into a column, washed with ten column volumes of lysis buffer, and eluted with the same buffer supplemented with 10 mM maltose. Cells expressing GLB-5 (HA)-His$_6$ were lysed, and centrifuged as described earlier. The cell pellet was resuspended in wash buffer (50 mM Tris-HCl, pH 8.3, 1 % (w/v) Triton X-100). After 20-min incubation on ice, the cell lysate was centrifuged again under the same conditions and the washing step was repeated twice more. Cell pellets were resuspended in Milli-Q water and centrifuged under the same conditions. This step was repeated twice. The resulting pellet was resuspended in solubilizing buffer (50 mM CAPS (pH 11.0), 300 mM NaCl, 0.3 % (w/v) Sarkosyl, and 2.5 mM dithiothreitol (DTT)). After a 2-h incubation at room temperature on a roller platform, insoluble debris was removed by centrifugation (30 min at 16,500 r.p.m.). The soluble fraction was applied to Ni-NTA beads (Qiagen) and the protein:bead mixture incubated at 4 °C as described earlier. After 4 h the beads were collected by centrifugation, loaded into a column, and washed with five column volumes of solubilizing buffer supplemented with 5 mM imidazol (pH 8.0). Protein was eluted with 300 mM imidazol in the same buffer. The eluted protein was dialysed extensively against reconstitution buffer (200 mM potassium phosphate, pH 8.0, 10 % (v/v) glycerol, 0.5 mM EDTA and 10 mM DTT) at 4 °C. Fresh haem solution was made by dissolving haemin (Sigma) in a minimum amount of 0.1 M of potassium hydroxide and diluting it with 25 volumes of reconstitution buffer. Insoluble haemin was removed by filtration (0.2 µm) and the concentration of the haem solution estimated using an extinction coefficient of $\varepsilon_{390} = 50$ mM^{-1} cm^{-1} in 2% borate buffer[31]. 0.3 M equivalents of haem were added to the protein solution while stirring on ice in a hypoxia chamber (Coy Laboratories) to a final 1.2 M excess of haem. Removal of non-incorporated haem and buffer exchange (to spectroscopy buffer, 50 mM potassium phosphate buffer (pH 8.0), 150 mM sodium chloride and 0.5 mM EDTA) were done by gel filtration chromatography. The monomeric fraction (as determined by analytical gel filtration using Superose 6 and broad range protein standards (Sigma)) was concentrated and used for spectroscopic measurements. The recombinant GLB-5 had an apparent purity of 95% as judged by SDS–PAGE (4–12% gradient, Novagene).

Spectroscopy. All experiments were done in spectroscopy buffer (see earlier). Absorption spectra were measured using a Varian Cary 50 Bio spectrophotometer. MBP-GLB-5 or GLB-5-His$_6$ protein was added to an anaerobic quartz cuvette (Helma) in its oxidized ferric form and spectra recorded. Dithionite was then added under argon to a final concentration of 1 mM and the deoxy ferrous spectra recorded. To demonstrate the ferric–ferrous cycle under physiological conditions, 5 µM ferric GLB-5 was incubated with an enzymatic reduction system from *E.coli*[16] in the presence of 0.5 mM NADH (Sigma), 200 U ml^{-1} glucose oxidase (Sigma), and 4 mM glucose. An absorption spectrum (ferrous) was recorded after a 20-min incubation. The cuvette was opened to air for 5 min, and another spectrum (ferric) was recorded. The cuvette was then closed, and after 15 min a further spectrum (ferrous) was recorded.

Calcium imaging. General guidelines for image acquisition and data analysis were described previously[26]. In brief, time-lapse image data were acquired through an agar pad using a × 40 C-Apochromat lens on an inverted compound microscope (Axiovert, Zeiss), with the Metamorph recording software (Molecular Devices). Photobleaching was limited by using a 2.0 optical-density filter and a shutter to limit exposure time to 100 ms per frame. An excitation filter (Chroma) restricted illumination to the cyan channel. A beam splitter (Optical insights) was used to separate the cyan and yellow emission light. The ratio of the background-subtracted fluorescence in the CFP and YFP channels emitted from the two cameleon fluorophores was calculated with Jmalyze. Results were processed in Matlab and Excel. All movies were captured at 2 frames per s.

To monitor O_2 changes in the microfluidic chamber (Supplementary Fig. 7) we used an oxygen optode (Presens), which measures absolute O_2 concentrations using lifetime-based luminescence quenching. This method does not consume O_2. The optode spot was cut to fit into the Y-chamber channel.

Quantification: PQR cell bodies were recorded for 2 min at each O_2 concentration in two sets of experiments. In the first set we used a 21–7–21–7–21% O_2 serial change, in the second set we used a 21–20–21–19–21–7–21% O_2 serial change. The first set was used to characterize responses of neurons in N2, *npr-1*, *gcy-35; npr-1* and *tax-4; npr-1* animals; the second was used to compare the dynamic range of *npr-1* and *npr-1; glb-5* animals. Error bars represent s.e.m.; statistical significance was calculated using *t*-tests.

Cell identification. Microscopy was carried out on a Zeiss Axioskop using a Neo-Fluar × 100 NA1.4 lens. The position of the cell body was used to identify neurons. GFP markers were used to confirm neuronal identities: animals carrying a functional *glb-5(HA)::mCherry* operon were crossed with strains expressing GFP from the *gcy-37* (AQR, PQR, URX) or *gcy-33* (BAG) promoter, respectively.

Modelling GLB-5 structure. The 3D-Pssm and Phyre programs[32] were used to find proteins with a fold similar to GLB-5. The best match was the rice non-symbiotic haemoglobin. This was used as a template to generate a three-dimensional model of the *glb-5* globin domain in the SWISS-MODEL program[33].

28. Oka, T., Toyomura, T., Honjo, K., Wada, Y. & Futai, M. Four subunit aisoforms of *Caenorhabditis elegans* vacuolar H^+-ATPase. Cell-specific expression during development. *J. Biol. Chem.* **276**, 33079–33085 (2001).

29. Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A. & Fire, A. Sequencerequirements for myosin gene expression and regulation in *Caenorhabditis elegans*. *Genetics* **135**, 385–404 (1993).

30. Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).

31. Antonini, E. & Brunori, M. *Frontiers of Biology* (Elsevier, 1971).

32. Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J. & Kelley, L. A. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **70**, 611–625 (2008).

33. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201 (2006).