EGL-13/SoxD Specifies Distinct O₂ and CO₂ Sensory Neuron Fates in Caenorhabditis elegans

Jakob Gramstrup Petersen¹, Teresa Rojo Romanos¹, Vaida Juozaityte¹, Alba Redo Riveiro¹, Ingrid Hums², Lisa Traunmüller², Manuel Zimmer², Roger Pocock¹*

1 Biotech Research and Innovation Centre, University of Copenhagen, Copenhagen, Denmark, 2 Research Institute of Molecular Pathology, Vienna, Austria

Abstract

Animals harbor specialized neuronal systems that are used for sensing and coordinating responses to changes in oxygen (O₂) and carbon dioxide (CO₂). In Caenorhabditis elegans, the O₂/CO₂ sensory system comprises functionally and morphologically distinct sensory neurons that mediate rapid behavioral responses to exquisite changes in O₂ or CO₂ levels via different sensory receptors. How the diversification of the O₂- and CO₂-sensing neurons is established is poorly understood. We show here that the molecular identity of both the BAG (O₂-sensing) and the URX (CO₂-sensing) neurons is controlled by the phylogenetically conserved SoxD transcription factor homolog EGL-13. egl-13 mutant animals fail to fully express the distinct terminal gene batteries of the BAG and URX neurons and, as such, are unable to mount behavioral responses to changes in O₂ and CO₂. We found that the expression of egl-13 is regulated in the BAG and URX neurons by two conserved transcription factors—ETS-5(Ets factor) in the BAG neurons and AHR-1(bHLH factor) in the URX neurons. In addition, we found that EGL-13 acts in partially parallel pathways with both ETS-5 and AHR-1 to direct BAG and URX neuronal fate respectively. Finally, we found that EGL-13 is sufficient to induce O₂- and CO₂-sensing cell fates in some cellular contexts. Thus, the same core regulatory factor, egl-13, is required and sufficient to specify the distinct fates of O₂- and CO₂-sensing neurons in C. elegans. These findings extend our understanding of mechanisms of neuronal diversification and the regulation of molecular factors that may be conserved in higher organisms.

Introduction

The capacity of the nervous system to sense and respond to fluctuations in the external and internal environment is essential for homeostasis and survival. Neuronally controlled homeostatic buffering is delivered through cellular and systemic physiological adjustments and by seeking optimal environmental conditions through behavioral strategies [1–4]. A crucial homeostatic capacity of animals is the ability to sense and respond to changes in concentration of the respiratory gases oxygen (O₂) and carbon dioxide (CO₂) [5,6]. O₂ is essential for the generation of energy in the form of adenosine triphosphate (ATP); however, O₂ also exerts toxicity through the production of reactive oxygen species (ROS) [1–4,7]. CO₂ is a by-product of oxidative metabolism and prolonged exposure leads to acidosis [5,6,8]. CO₂ is also an environmental cue used in host- and mate-finding and can initiate both aversive or attractive behaviors [9–11]. The evolution of mechanisms required to sense and respond to O₂ and CO₂ is therefore paramount for survival.

In Drosophila, specific sensory systems respond to external O₂ levels [1,7,12]. In addition, Drosophila uses specialized olfactory and gustatory neurons to detect CO₂ changes via specialized chemo-sensory receptors called Gr21a/Gr63a [9,13]. In humans, O₂, CO₂ and pH levels are monitored by specific regions of the brainstem and by specialized neurosecretory glomus cells of the carotid body [14], whereas in non-human mammals CO₂ is also sensed by specific olfactory neurons that target the nasal glomeruli in the olfactory bulb via the guanylyl cyclase GC-D [15]. It is poorly understood how the specification of such specialized sensory neurons is regulated. However, recent work in Drosophila has shown that epigenetic mechanisms play an important role [16].

Respiratory gas sensing is a crucial modality for Caenorhabditis elegans whose natural environment, such as rotting fruit and compost, can have wide ranges of O₂ and CO₂ levels [17]. Previous work has shown that in the laboratory, worms have a behavioral preference for 5%–10% O₂ and are exquisitely sensitive to minor changes in O₂ concentration [18,19]. In addition, worms mount avoidance responses to CO₂ levels above 0.5% [4,11]. Of the 302 neurons in the C. elegans nervous system, at least six neurons are specifically dedicated to the detection and response to changes in O₂ and CO₂ levels. These include the BAGL/R, URXL/R, AQR and PQR neurons. The BAG neurons are the primary O₂ sensors and they also respond to decrease in O₂ concentration [11,20–22]. The URX, AQR and PQR neurons are specialized for responding to increasing O₂ concentrations [20]. In C. elegans, members of the guanylyl cyclase family of proteins are crucial factors required for O₂ and CO₂ sensing. Pioneering work revealed that the soluble guanylyl cyclases (sGCs) GCY-35 and GCY-36 mediate high O₂ avoidance...
behavior via the URX, AQR and PQR neurons and that GCY-35 directly binds to molecular O$_2$ [18]. In contrast, the sGCs, GCY-31 and GCY-33 function in the BAG neurons to sense decreases in O$_2$ [20]. Recent work found that the membrane-bound receptor-type guanylyl cyclase GGY-9 acts specifically in the BAG neurons to mediate CO$_2$ avoidance behavior [21]. Other molecules such as the Phe-Met-Arg-Phe-NH$_2$ (FMRF-amide)-related peptides (FLP-8, FLP-13, FLP-17 and FLP-19) are expressed in either a subset or all of the O$_2$- and CO$_2$-sensing neurons; however their precise molecular functions in O$_2$ and CO$_2$ sensing are not known [23,24].

Neuronal specialization within the O$_2$/CO$_2$-sensing system in *C. elegans* is an excellent model to study the control of neuronal diversity. The O$_2$-sensing (URX, AQR and PQR) and O$_2$/CO$_2$-sensing (BAG) neurons have overlapping and non-overlapping patterns of guanylyl cyclase and neuropeptide expression, which are reflected in their related, albeit distinct functionalities [20]. At present, it is unclear how the expression of these molecules is restricted to certain parts of the O$_2$/CO$_2$-sensing nervous system, and how such restrictions coordinate neuronal fate and function.

Here, we have identified the Sox transcription factor EGL-13 as an important regulator of the O$_2$ and CO$_2$-sensing neuron cell fate decision. EGL-13 is required for the expression of distinct proteins required for sensing both O$_2$ and CO$_2$ and as such, *egl-13* mutant animals are unable to mount behavioral responses to changes in O$_2$ and CO$_2$. We found that the expression of EGL-13 is controlled by ETS-5 in the BAG neurons and by AHR-1 in the URX neurons, and acts in partially parallel pathways with these factors to drive neuronal fate. Finally, we found that EGL-13 is sufficient to drive O$_2$- and CO$_2$-sensing cell fates in certain cellular contexts. Therefore, EGL-13 is a core regulatory factor that is both required and sufficient to drive O$_2$- and CO$_2$-sensing neuron specification in *C. elegans*. As EGL-13 is a member of the SoxD family of transcription factors, we anticipate that the regulatory relationships described here will provide a paradigm for the control of neuronal fate specification by Sox proteins in other cellular contexts.

**Author Summary**

During the development of an organism, certain neurons are programmed to perform specific tasks. For example, motor neurons coordinate locomotion and sensory neurons recognize specific environmental cues. The molecular mechanisms that generate specific neuronal classes are not fully understood. We investigated mechanisms that control the development of two distinct classes of neurons that are required for the nematode *Caenorhabditis elegans* to sense the respiratory gases O$_2$ or CO$_2$. In this study, we identified and characterized a conserved transcription factor, *egl-13*, that is required for the development of both of these classes of neurons. *egl-13* is related to the SoxD family of transcription factor proteins in vertebrates. We found that *egl-13* controls the production of specific proteins that provide these cells with the ability to sense both O$_2$ and CO$_2$. Further, we found that *egl-13* works in conjunction with two additional factors, *ahr-1* and *ets-5*, to regulate this developmental decision. This work provides new insight into how transcriptional regulatory networks specify different but related neuronal identities and provides a platform for future studies to understand how neuronal diversity is generated.

**Results/Discussion**

The SoxD transcription factor *egl-13* specifies distinct O$_2$- and CO$_2$-sensing neurons

In order to identify molecules and pathways important for O$_2$- and CO$_2$-sensing neuron specification, we have taken a forward genetics approach in *C. elegans*. We isolated four independent allelic mutations (*q14*, *q22*, *q23* and *q26*) that affect the expression of terminal differentiation markers in the O$_2$ and/or CO$_2$-sensing neurons (Figure 1 and Table S1). Mutant hermaphrodites of each of these alleles are severely egg-laying defective (Egl) and form a bag-of-worms where embryos hatch inside the mother (Figure S1). We investigated their vulval phenotype and found that the anchor cell fails to fuse with the uterine seam cell, causing a blockage of the uterus and the resultant Egl phenotype (Figure S1). This anchor cell fusion defect is reminiscent of that observed in *egl-13(ku194)* mutant animals [25] which we found to also exhibit defects in O$_2$ and CO$_2$ reporter expression (Figure 1 and Figure S1). Subsequent Sanger sequencing of *q14*, *q22*, *q23* and *q26* revealed genetic lesions in the *egl-13* locus (Figure 1A). *egl-13* encodes the *C. elegans* ortholog of the HMG-domain-containing SoxD family of transcription factors that has no previously reported role in the worm nervous system.

**Loss of *egl-13* affects terminal fate of O$_2$- and CO$_2$-sensing neurons**

The BAG, URX, AQR and PQR neurons in *C. elegans* are required for sensing and responding to fluctuations of O$_2$ and CO$_2$ levels in the environment [11,18,19]. Distinct batteries of genes are expressed in these neurons that are predicted to provide the optimal functionality required for O$_2$ and CO$_2$ sensing, however the role of only a subset of these genes has been analyzed in detail [20,21,26]. We used fluorescent reporter constructs to monitor expression of these gene batteries to understand how *egl-13* controls O$_2$ and CO$_2$-sensing neuron cell fate (Figure 1). We analyzed the expression of guanylyl cyclases (gyr-9, gyi-31, gyi-33, gyr-35 and gyr-36) and Phe-Met-Arg-Phe-NH$_2$ (FMRF-amide)-related peptides (flp-8, flp-13, flp-17 and flp-19) that are all terminal differentiation genes expressed in all or a subset of O$_2$ and CO$_2$-sensing neurons [18,19,21,24]. We crossed these reporter transgenes into *egl-13* mutant animals (*ku194* allele) and found that none of the reporters were properly expressed in *egl-13* mutants (Figure 1 and Table S1). We also found similar effects in the four *egl-13* mutant alleles we isolated (*q14*, *q22*, *q23* and *q26*) (Figure S1C). We noticed that some of the reporters were exclusively sensitive to *egl-13* loss whereas others exhibited partially penetrant defects (Table S1). This suggests that the expression of some terminal differentiation factors are under the collaborative control of additional factors that are able to compensate for the loss of *egl-13*.

The BAG and URX neurons are derived from the AB lineage and are posterior sisters of other neurons that have distinct fates [27] (Figure S2). We therefore asked whether *egl-13* is also required for the specification of the sister cells of BAG or URX. We crossed *egl-13(ku194)* mutant animals into fluorescent reporter strains for the SMDV, *zfhl2* (lge-35::mCherry) and CEPD, *zfh1* (dat-1::gfp), sister cells for BAG and URX neurons respectively. We found that the expression of these reporters were unaffected by loss of *egl-13* suggesting a specific role for *egl-13* in the posterior branch of these lineages (Table S1 and Figure S2). Taken together, we conclude that *egl-13* controls the expression of the distinct O$_2$- and CO$_2$-sensing neuron terminal gene batteries that distinguish them from lineage-related neurons.
Figure 1. *egl-13* is required for O$_2$- and CO$_2$-sensing neuron specification. (A) Molecular identity of mutant alleles obtained from the forward genetic screens. *egl-13* alleles first described in this article are shown in black (*rp14, 22, 23*, and *26*) and the previously described *ku194* [25] is shown in
eegl-13 acts cell-autonomously in O$_2$- and CO$_2$-sensing neurons

To monitor eegl-13 expression, we generated two promoter-driven fluorescent reporters (eegl-13*prom3::mCherry and eegl-13*prom2::gfp) that contain 3.5 kb of eegl-13 upstream sequence (Figure 2A and Figure S5). Expression is first detected in 4 neuronal cells at around 350 min post-fertilization, which is the time at which the BAG and URX neurons are born (Figure S3). Expression is restricted to these 4 neurons during embryogenesis (Figure S3). At the first larval stage, eegl-13 expression is observed in the BAG and URX neurons plus occasionally in a small number of unidentified cells in the head and tail (including the AQR and PQR neurons) (Figure S3). Later during larval development, eegl-13 expression is observed in body wall muscle and vulval cells (data not shown). Neuronal expression is restricted to the O$_2$ and CO$_2$-sensing neurons in the adult (Figure 2A). Using the 3.5 kb eegl-13 promoter (eegl-13*prom1) we transgenically expressed eegl-13*isoformC cDNA in eegl-13(ku194) mutant animals and were able to rescue both the defect in O$_2$ and CO$_2$-sensing neuron fate marker expression and the Egl phenotype (Figure 2B–2C, Figure S4 and data not shown). To confirm that eegl-13 acts cell autonomously to control O$_2$ and CO$_2$-sensing neuron fate, we used neuron-specific promoters to drive eegl-13*isoformA cDNA expression in the BAG or URX neurons (Figure 2D). We found that indeed neuron-specific expression of eegl-13 rescued the O$_2$ and CO$_2$-sensing neuron fate defect of eegl-13(ku194) mutant animals (Figure 2D). Therefore, we conclude that eegl-13 acts autonomously in the BAG and URX neurons to direct their fate.

The eegl-13 gene has 4 predicted isoforms, all of which contain the same HMG DNA/protein binding domain, however they each have varying lengths of amino terminal tail. Such tails in SoxD proteins can cooperate with other factors to control gene expression [28,29]. We therefore tested whether the long N-terminal region of EGL-13*isoformA is required for its rescuing ability. We used eegl-13*promB to drive EGL-13*promB (lacking 157 amino acids of the N-terminal tail of isoformA) in eegl-13(ku194) animals and found that it fully rescued the defect in O$_2$ and CO$_2$-sensing neuron fate marker expression and the Egl phenotype (Figure S4 and data not shown). Thus, the EGL-13 N-terminal region is not required for its roles in O$_2$ and CO$_2$-sensing neuron specification, as was recently found in chickens [30].

eegl-13 is required and sufficient to induce O$_2$-sensing neuron fate

We have shown that eegl-13 is expressed throughout the life of the worm in the O$_2$ and CO$_2$-sensing neurons; and is required to induce terminal differentiation features. To ask whether eegl-13 is required continuously to maintain the expression of the terminal gene battery of these neurons, we sought to postdevelopmentally remove eegl-13 gene activity, eegl-13 gene activity could not be removed by RNA-mediated interference in an RNAi sensitized background (data not shown) and there are no temperature-sensitive alleles of eegl-13 available. Instead, we generated animals that lack endogenous EGL-13 protein but express heat-shock inducible eegl-13 cDNA from an extrachromosomal array under the control of the hsp-16.2 promoter (Figure 3). We focused our analysis on the URX neurons and found that the loss of ggy-3*prom3::gcy-33::gfp reporter expression in eegl-13(ku194) worms could be rescued through heat-shock induction of eegl-13 during mid-larval stages (Figure 3A). This indicates that O$_2$-sensing neurons generated during embryogenesis persist in an eegl-13-responsive state. These neurons are, therefore, not converted into another fate when eegl-13 is lost; however, they do not acquire the terminal O$_2$-sensing neuron differentiation program. When eegl-13 activity was supplied transiently, through removal of heat-shock stimulus, we observed a gradual loss of reporter expression during adulthood in the URX neurons (Figure 3A). Therefore, eegl-13 gene activity is continuously required to maintain URX cell fate. To ask whether misexpression of eegl-13 in other neurons is sufficient to induce O$_2$ and CO$_2$ terminal fate we expressed eegl-13 under the control of an early neuronal promoter (Figure 3B). We found that eegl-13 is indeed sufficient to induce expression of O$_2$ and CO$_2$ terminal fate markers in some cellular contexts (Figure 3B). This suggests that eegl-13 is not only required but also sufficient to induce O$_2$ and CO$_2$-sensing neuron fate in specific contexts, which is similar to previous studies of terminal selector genes [29–31]. The restricted induction we observed may be dependent on the embryonic time-point of induction or the expression of other unknown co-factors that are required for induction of O$_2$ and CO$_2$-sensing neuron fate.

eegl-13 mutants are defective in O$_2$- and CO$_2$-sensing

The crucial role for eegl-13 in O$_2$ and CO$_2$-sensing neuron fate determination suggested that eegl-13 mutant animals would be defective in O$_2$ and CO$_2$ sensing. We applied three behavioral paradigms that have been previously reported to be specific to either one of these neuron classes: BAG neurons modulate the animals’ locomotion speed in response to an oxygen downshift from 21% O$_2$ towards 10% O$_2$ (Figure 4A, 4E) [20]. In addition, BAG neurons detect increases in CO$_2$ concentrations, which trigger reorientation movements (omega turns) (Figure 4G) [11,21]. URX neurons modulate the animals’ locomotion speed in response to O$_2$ upshifts towards 21% O$_2$ (Figure 4A, 4F) [20]. We applied these behavioral assays to test how BAG and URX neurons are functionally affected in eegl-13 mutants. We tracked animals in a chamber without food, in an air-flow that switched between 21% O$_2$ and 10% O$_2$, or between 0% CO$_2$ and 1% CO$_2$. In contrast to wild-type animals, eegl-13(ku194) mutant animals do not slow their locomotion in response to O$_2$ upshift or downshift (Figure 4A, 4C, 4E, 4F). We found that eegl-13(ku194) mutants are...
Figure 2. *egl-13* functions cell autonomously to drive O$_2$- and CO$_2$-sensing neuron cell fate. (A) Dorsal view of a young adult hermaphrodite expressing an *egl-13*::mCherry transcriptional reporter transgene (top panels), a gcy-33::gfp translational reporter
also defective in CO2 sensing since they fail to slow or perform omega turns in response to CO2 (Figure 4G). O2 and CO2 behavioral defects of egl-13(ku194) mutants are fully rescued when egl-13 cDNA is re-supplied under the control of egl-13 promoter, which is immediately upstream of the ATG codon, which is sufficient to drive expression in the BAG neurons. Intriguingly, we found two conserved ETS-5/Pet1 binding sites in this region (Figure 5). Previous work identified ETS-5 as a crucial factor required for the specification of the BAG neurons, suggesting that ETS-5 may regulate egl-13 expression in these neurons [26,33]. We used site-directed mutagenesis to eliminate the ETS-5/Pet1 binding sites individually and in combination, and found that when both ETS-5/Pet1 binding sites are mutated the expression of egl-13 is abrogated in the BAG neurons (Figure 5A). This suggests that ETS-5 directly regulates the expression of egl-13 in the BAG neurons via conserved binding sites. We crossed the egl-13(tm1734) mutant into the egl-13prom::mCherry strain and indeed found that BAG expression was affected (Figure 5B). In addition, we found that egl-13 can regulate its own expression in the BAG neurons independently of egl-5 via an, as yet, unidentified mechanism (Figure 5B).

Taken together, these data indicate that control of egl-13 expression is coordinated by two independent regulatory mechanisms. In the O2-sensing URX neurons, egl-13 expression is predominantly regulated by AHR-1 (Figure 5C). In contrast, an independent promoter module controlled by ETS-5 regulates egl-13 expression in the O2/CO2-sensing BAG neurons (Figure 5C).

**egl-13 acts in partially parallel pathways with egl-5 and ahr-1**

Previous work has identified egl-5 and ahr-1 as regulators of BAG and URX specification respectively [26,32,33] and we have shown that these factors are predominantly required to drive egl-13 expression in these cells. To understand how these factors function together to coordinate BAG and URX specification, we analyzed the expression of terminal fate markers in single and double mutant combinations, where appropriate. We analyzed three URX markers (flp-19::gfp, flp-19::mCherry, and gcy-33::gfp) and six BAG markers (flp-13::gfp, flp-17::gfp, flp-19::gfp, gcy-9::gfp, gcy-31::gfp, and gcy-33::gfp) and compared the effect of the individual loss of egl-13, egl-5 and ahr-1 (Figure S5).

The first observation from this analysis was that the expression of a subset of terminal differentiation markers is completely dependent on egl-13 and one of the other factors acting in a linear pathway. For example, we find that BAG expression of flp-13::gfp and flp-19::gfp is almost 100% affected in both the egl-13 and egl-5 single mutants (Figure S5). This suggests that for these markers egl-13 and egl-5 act in the same pathway to drive marker expression. In contrast, expression of gcy-9::mCherry is completely dependent on egl-5 with egl-13 playing a minor role in its regulation (Figure S5).
Figure 3. egl-13 is required and sufficient to induce O$_2$- and CO$_2$-sensing neuron fate. (A) egl-13 is required to maintain O$_2$/CO$_2$-sensing neuron fate. Induction of egl-13 expression via transient heat-shock at the L2/L3 stage restores gcy-33::gfp expression in the URX neurons of.
At the other end of the spectrum, ets-5 and egl-13 are minimally required to drive gcy-31::mCherry expression in the BAG neurons suggesting other factor(s) control the expression of this terminal fate marker (Figure S5). Taken together, these data indicate that egl-13 and ets-5 act in partially parallel pathways to drive BAG cell fate and that other unknown factors possibly act in a combinatorial manner to drive specific aspects of BAG fate.

We also observed differential effects of egl-13 loss with URX terminal fate markers. Expression of the flp-8::gfp reporter is partially affected by single loss of egl-13 and abr-1, whereas loss of both genes totally abrogates expression, suggesting that egl-13 and abr-1 act in parallel pathways to regulate flp-8::gfp expression (Figure S5). However, in the case of flp-19::gfp, loss of egl-13 causes complete loss of expression and abr-1 plays a minor role in its regulation (Figure S5).

To further investigate the regulatory relationship between egl-13, ets-5 and abr-1 we analyzed how they affect the expression of each other. We have already shown that ets-5 positively regulates the expression of egl-13 in the BAG neurons (Figure 5B). In a reciprocal experiment, we found that ets-5::gfp expression is unaffected in egl-13(ku194) mutant animals (Figure 5B). These data and other work [26,33] suggest that ets-5 acts upstream and in parallel to egl-13 to direct BAG cell fate (Figure 5C). In addition, we found that egl-13 is able to regulate its own expression in the BAG neurons, in parallel to ets-5; however, the mechanistic basis of this regulation is unclear (Figure 5B). In the URX neurons, we found that egl-13 and abr-1 regulate the expression of each other in addition to having autoregulatory capabilities (Figure 5B,C and Figure S5).

Our studies have elucidated a novel function for egl-13, the SoxD homolog, in the specification of distinct classes of O₂ and CO₂ sensory neurons in C. elegans. We show that egl-13 is expressed in the O₂ and CO₂-sensing neurons and acts cell-autonomously to regulate their distinct cell fates. We further show that egl-13 is continuously expressed in the O₂ and CO₂-sensing system to maintain the expression of terminal features of these neurons. In certain cellular contexts, egl-13 is also sufficient to induce O₂- and CO₂-sensing neuron cell fate. We found that the regulatory inputs controlling the expression of egl-13 in the O₂ and CO₂-sensing system are mechanistically distinct. Independent regulatory modules control egl-13 expression in the BAG neurons (CO₂ and O₂ downshift sensors) versus the URX neurons (O₂ upshift sensors). Interestingly, we found that egl-13 expression in the BAG neurons is controlled by the ETS-5 transcription factor via conserved ETS binding sites. In contrast, in the URX neurons, egl-13 expression is controlled by the bHLH transcription factor AHR-1 via a conserved AHR1 binding site.

The influence EGL-13 exerts on the expression of the terminal gene batteries of the distinct O₂- and CO₂-sensing neurons is diverse. Particular factors are exquisitely sensitive to loss of egl-13, whereas others are only partially affected. These findings suggest that alternative unknown modes of regulation are in place to ensure that particular molecules are faithfully expressed in the O₂ and CO₂ sensory neurons, which work in conjunction with and/or in parallel to egl-13.

Sox transcription factors have diverse functions during development and play crucial roles in regulating neuronal fate [34–37]. In addition, Sox proteins act at different levels to preselect neuronal genes in embryonic stem cells and to direct the activation of these genes in neuronal precursors and fully differentiated neurons [38]. Here we describe a novel role for EGL-13, the SoxD transcription factor in C. elegans, in driving the specification of different but related sensory neuron identities. Closely related orthologs of EGL-13 are found in vertebrates, some of which are expressed in sensory neurons [39], therefore; SoxD proteins may have a previously unrecognized conserved function in the specification of gas-sensing neurons in higher organisms.

Materials and Methods

Strains used in this study

Strains were grown using standard growth conditions on NGM agar at 20°C on Escherichia coli OP50 [8,40]. Transgenic animals were created according to [41]. Strain information is detailed in Table S2.

Forward genetic screening approaches

In all screens, animals were mutagenized with EMS (ethyl methanesulfonate) according to standard protocols [42]. Worms were incubated at 25°C at all times. In the manual screens, 5 parental (P0) mutagenized animals were placed in each of 10 founder plates. Three days later, 400 F1 progeny of the mutagenized P0 animals were singled. Their ensuing F2 progeny were screened under a fluorescence stereomicroscope.

In the automated worm sorter screen, around 100,000 synchronized larval stage L4 animals were mutagenized with EMS, the following day the P0 young adult animals were bleached and their F1 progeny synchronized at larval stage L1 by starvation (approximately 1,000,000 animals). F1 animals were grown to the young adult stage, bleached and their F2 progeny synchronized at larval stage L1 by starvation (approximately 10,000,000 animals). The F2 progeny were grown until larval stage L4 and 10% of the population (approximately 1,000,000) was passed through a COPAS biosorter (Karolinska Institute, Stockholm, Sweden).

The influence EGL-13 exerts on the expression of the terminal gene batteries of the distinct O₂- and CO₂-sensing neurons is diverse. Particular factors are exquisitely sensitive to loss of egl-13, whereas others are only partially affected. These findings suggest that alternative unknown modes of regulation are in place to ensure that particular molecules are faithfully expressed in the O₂ and CO₂ sensory neurons, which work in conjunction with and/or in parallel to egl-13.

Sox transcription factors have diverse functions during development and play crucial roles in regulating neuronal fate [34–37]. In addition, Sox proteins act at different levels to preselect neuronal genes in embryonic stem cells and to direct the activation of these genes in neuronal precursors and fully differentiated neurons [38]. Here we describe a novel role for EGL-13, the SoxD transcription factor in C. elegans, in driving the specification of different but related sensory neuron identities. Closely related orthologs of EGL-13 are found in vertebrates, some of which are expressed in sensory neurons [39], therefore; SoxD proteins may have a previously unrecognized conserved function in the specification of gas-sensing neurons in higher organisms.

Materials and Methods

Strains used in this study

Strains were grown using standard growth conditions on NGM agar at 20°C on Escherichia coli OP50 [8,40]. Transgenic animals were created according to [41]. Strain information is detailed in Table S2.

Forward genetic screening approaches

In all screens, animals were mutagenized with EMS (ethyl methanesulfonate) according to standard protocols [42]. Worms were incubated at 25°C at all times. In the manual screens, 5 parental (P0) mutagenized animals were placed in each of 10 founder plates. Three days later, 400 F1 progeny of the mutagenized P0 animals were singled. Their ensuing F2 progeny were screened under a fluorescence stereomicroscope.

In the automated worm sorter screen, around 100,000 synchronized larval stage L4 animals were mutagenized with EMS, the following day the P0 young adult animals were bleached and their F1 progeny synchronized at larval stage L1 by starvation (approximately 1,000,000 animals). F1 animals were grown to the young adult stage, bleached and their F2 progeny synchronized at larval stage L1 by starvation (approximately 10,000,000 animals). The F2 progeny were grown until larval stage L4 and 10% of the population (approximately 1,000,000) was passed through a COPAS biosorter (Karolinska Institute, Stockholm, Sweden).
Figure 4. *egl-13* is essential for locomotion responses to O$_2$ and CO$_2$ concentration shifts. (A-D) Locomotion speed of *C. elegans* during O$_2$ concentration shifts. Traces show average forward speed and dark shading indicates standard error of the mean (SEM). O$_2$ concentrations were switched between 21% and 10%. Shading represents intervals at 21%. (A) Wild-type animals. (B) *egl-13(pk23)* mutants. (C) *egl-13(ku194)* mutants. (D) Rescue of *egl-13(ku194)* mutant phenotype by transgenic expression of *egl-13* cDNA under control of its own promoter. Transgene *rpEx401*. Note the respective reduction of speed levels in N2 after up- and downshift, which are abolished in *egl-13(ku194)* and restored in the transgenic line. *egl-13(pk23)* animals are affected mostly in their response to O$_2$ upshift. (E, F) Quantification of data in A–D. Average speed changes in percent from basal speed to O$_2$ downshift (E) and upshift (F) of animals with indicated genotypes. Transgenic rescue lines are significantly different from *egl-13(ku194)* mutant animals. (G) Average changes in omega turn frequency of animals with indicated genotypes, in response to 1% CO$_2$. The defect in omega turn responses seen in *egl-13(ku194)* animals is restored in the transgenic lines (*rpEx399* and *rpEx401*). *egl-13(pk23)* animals only exhibit a partial defect. Error bars = SEM. Symbols indicate all significant differences one-way ANOVA with Bonferroni’s Multiple Comparison Test (**/**** p = 0.01–0.05, ***/***** p < 0.001). Asterisks indicate significant difference compared to wild-type, while diamonds indicate significant difference compared to *egl-13(ku194)* mutants. Data were calculated from n = 3 independent experiments for each mutant and transgenic rescue strain, and n = 6 independent experiments for wild-type. Each individual experiment was performed on 60–70 animals.

doi:10.1371/journal.pgen.1003511.g004

**C. elegans** expression constructs and generation of transgenic worms

Reporter gene constructs were generated by PCR amplifying promoter elements and cloning into the pPD95.75-mCherry and gfp vectors (Fire Vector Kit). Mutagenesis was performed using the QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene). Rescue constructs were generated by cloning promoter and cDNA sequences into the pPD49.26 expression vector (Fire Vector Kit). Constructs were injected into young adult hermaphrodites as either simple arrays (gcy-33::gfp/gcy-33::gfy (50 ng ul$^{-1}$) and pRF4 (50 ng ul$^{-1}$) as injection markers) or as complex arrays using 1–10 ng ul$^{-1}$ of linearized plasmid, 150 ng ul$^{-1}$ of *ParII*-digested bacterial genomic DNA and myo-2::gcy-33::dsRed (3–5 ng ul$^{-1}$), elt-2::gcy-33::gfy (3–15 ng ul$^{-1}$) as injection markers.

**Behavioral assays**

Animals were transferred without food to 14 cm NGM assay plates containing a cut out arena of Whatman filter paper soaked in 20 mM CuCl$_2$ to prevent them from leaving a 36 mm x 36 mm center area. Sixty to seventy animals were used in a single experiment and starved for one hour prior to examination. Each experiment was carried out three times, except for wild-type, which was performed six times. A custom-made transparent plexiglass device with a flow arena of 60 mm x 60 mm x 0.7 mm was placed onto the assay arena and animals were accustomed to a gas flow of 100 ml/min containing 21% (%v/v) oxygen for 5 minutes. During the assays animals were exposed for 6 minutes to 21% O$_2$ before and after a 6 minute stimulus interval of either 10% O$_2$ or 1% CO$_2$ (+21% O$_2$). All gas mixtures were balanced with N$_2$. Gases were mixed with a static mixing element connected to mass flow controllers (Vogtlin Instruments) that were operated by LabView software. Recordings of freely behaving animals illuminated with flat red LED lights were made at 3 fps on a megapixel CCD camera (Jai) using Streampix software (Norpix). Movies were analyzed by MatLab-based image processing and tracking scripts as previously described [43,44]. The resulting trajectories were used to calculate instantaneous speed during continuous forward movements (1 second binning). Omega turns were detected based on characteristic changes in object eccentricity and their frequency was calculated in 15 second bins. For quantifications, relative speed changes were calculated between representative intervals of 120 seconds before (basal level) and 4 seconds after the stimulus, capturing the minimum speed levels (4–8 seconds post stimulus). Data were normalized to the basal level. Changes in omega turn frequency were calculated between representate intervals of 180 seconds before (basal level) and 60 seconds after the stimulus, to capture the maximum rise phase (55–115 seconds post stimulus).

**Heat-shock experiments**

Two transgenic lines for *hsp-16.2::egl-13* were used for the heat-shock experiments. For the rescue and maintenance experiments, third larval stage (L3) worms were heat shocked at 37°C two times for 30 min. After heat shock, worms were kept at 25°C overnight and then transferred to 15°C for 2 days.

**Microscopy**

Worms were mounted on 5% agarose on glass slides and images were taken using an automated fluorescence microscope (Zeiss, AXIO Imager M2) and MicroManager software (version 3.1).

**Neuronal scoring**

Neurons were given a numerical value according to their expression levels. Wild-type expression scored 1, decreased expression scored 0.5 and abolished expression scored 0. Percentage of GFP expressing animals was then correlated to the theoretical maximum score using the equation below.

\[
\% \text{ of GFP expressing animals} = \frac{\text{observed score}(n_1 + 1) + (n_2 + 0.5) + (n_3 + 0)}{\text{theoretical score}(n_1 + n_2 + n_3) + 1} \times 100%
\]

**Bioinformatic analysis**

The Jaspar program (http://jaspar.genereg.net/) was used to predict the transcription factor binding sites in the *egl-13* upstream regulatory sequence.

**Statistical analysis**

Statistical analysis was performed in GraphPad Prism 5 using one-way ANOVA with Newman-Keuls Multiple Comparison Test. Values are expressed as mean ± s.d. Differences with a *P* value < 0.05 were considered significant. For the behavioral assays statistical significance was determined using one-way ANOVA with Bonferroni’s Multiple Comparison Test.

**Supporting Information**

Figure S1 *egl-13* mutants have anchor cell fusion defects. (A) *egl-13(ku194)* mutant hermaphrodites are severely egg-laying defective due to a defect in anchor cell fusion (right, compared to the wild-type adult on the left). (B) *egl-13* mutant alleles isolated from our forward genetic screens (*rp14, 22, 23 and 26*) all have anchor cell fusion defects comparable to the *ku194* allele. In the wild-type vulva, the anchor cell fuses to the utse cells to form the mature uterine-vulval connection (green arrowhead, top left). In *egl-13* mutant animals, the anchor cell fails to fuse to the utse (red
Figure 5. Independent regulatory modules drive egl-13 expression in O\textsubscript{2} versus O\textsubscript{2}/CO\textsubscript{2}-sensing neurons. (A) egl-13 promoter analysis. Schematic representation of the egl-13 locus with its 3.5 kb upstream region. The ATG codon is marked with an arrow and the exons are represented.
as black blocks. The upstream region deleted in the rp23 allele is indicated with a blue horizontal line and a grey dashed vertical line. Below is a representation of cloned and injected constructs, and their expression pattern in the URX and BAG neurons. Black lines denote the promoter fragment placed in front of mCherry fluorescent protein (red boxes). Orange, blue and green crosses represent mutated Sox5, AHR-1 and ETS binding sites respectively. “+” indicates consistent reporter expression in at least 50% of animals in all lines. “−” indicates less than 50% of animals expressing the reporter in all lines. “−−” indicates loss of expression. Three independent transgenic lines were analyzed for each promoter as indicated. (B) Expression of egl-13 and ets-5 reporter transgenes in the BAG neurons in the reciprocal mutant backgrounds (left). egl-13(ku194) and ets-5(tm1734) alleles were used. Expression of egl-13 and ahr-1 reporter transgenes in the URX neurons in the reciprocal mutant backgrounds (right). egl-13(ku194) and ahr-1(ku194) alleles were used. n>50 per strain. Error bars represent the standard error of the mean (SEM) **P<0.05, ***P<0.005. (n.s.) indicates no significant difference from wild-type. See materials and methods for neuronal scoring criteria used. (C) In the BAG neurons (left), the ETS transcription factor, ETS-5 regulates the expression of egl-13 in the BAG neurons via conserved ETS binding sites located in the egl-13 promoter. In addition, egl-13 and ets-5 autoregulate in the BAG neurons to generate BAG fate. In the URX neurons (right), egl-13 and ahr-1 autoregulate, in addition to regulating the expression of each other. Other unknown factors act in parallel to and in combination with these factors to drive BAG and URX neuronal fate.

doi:10.1371/journal.pgen.1003511.g005

Table S1 Expression of the O2− and CO2-sensing neuron terminal gene battery is affected in egl-13 mutant animals. (A) Effects in the expression of O2− and CO2-sensing neuron reporters in egl-13(ku194) mutant animals differ from gene to gene and from cell to cell. For example, flp-19 expression is strongly affected in both URX and BAG neurons, whereas gcy-33 expression is more strongly affected in the URX neurons than BAG neurons. Fluorescent reporters for the sister cells of the URX (CEPD) and BAG (SMDV) neurons are unaffected in egl-13(ku194) mutant animals. (B) The deletion in the egl-13 locus of rp23 leaves the BAG-element intact (see Figure 5). As a result, BAG terminal fate marker expression is not significantly affected in egl-13(rp23) mutant animals. In both tables, quantification indicates the percentage of animals in which fluorescence was observed (ON), not observed (OFF) or asymmetrically affected left expressed (L ON) or right expressed (R ON). “−” indicates that the reporter is not expressed in those neurons. Animals were scored at the young adult stage. See materials and methods for neuronal scoring criteria used.

Table S2 Strains used in this study.

Acknowledgments

We thank members of the Pocock Laboratory and Oliver Hobert, Baris Tursun, and Iva Greenwald for comments on the manuscript; we thank the following people for strains and reagents: Oliver Hobert (Columbia University), Cornelia Bargmann (Rockefeller University), Mark Alkema (University of Massachusetts Medical School). Some strains were provided by the CGC and by Shohei Mitani at the National Bioresource Project (Japan). We also thank Thomas Burglin and Peter Swoboda at the Karolinska Institute for use of their COPAS Biosorter facility, under the C. elegans Nordforsk Researcher Network of Shared Technology Platforms.

Author Contributions

Conceived and designed the experiments: JGP TRR VJ ARR IH LT MZ RP. Performed the experiments: JGP TRR VJ ARR IH LT MZ RP. Analyzed the data: JGP TRR VJ ARR IH LT MZ RP. Wrote the paper: RP.
Neurons