www.pnas.org

Supplementary Information for
Dual oxidase enables insect gut symbiosis by mediating respiratory network formation

Seonghan Jang, Peter Mergaert, Tsubasa Ohbayashi, Kota Ishigami, Shuji Shigenobu, Hideomi Itoh and Yoshitomo Kikuchi

***Correspondence**
Email: y-kikuchi@aist.go.jp or peter.mergaert@i2bc.paris-saclay.fr

This PDF file includes:

- Supplementary text
- Figures S1 to S7
- Tables S1 to S2
- SI References

Supplementary text

Supporting Materials and Methods

Measurement of the insect survival rate

To estimate the pathogenicity of entomopathogenic bacterium *S. marcescens*, third instar nymphs of Sym and *RpDuox*-RNAi bean bugs were orally infected with 10^7 cells/ml of *S. marcescens* and the number of dead insects was counted every day. The number of carcasses of *RpBnl*-, *RpTrh*-, *RpSima*-, and *RpDuox*-RNAi insects was counted every day to measure the mortality of insects upon the silencing of these genes. For each experiment, 20 insects were used.

RNA-sequencing analysis

Total RNA was extracted from the dissected symbiotic organ of Sym and *RpDuox*-RNAi insects using the RNAiso Plus and RNeasy Mini Kits. The cDNA libraries were prepared by TruSeq RNA Sample Preparation Kit v2 (Illumina) and sequenced by Illumina HiSeq-2000 (Illumina). The adapter sequences of raw reads (Accession number DRR215696-DRR215701) were trimmed by Trimmomatic (1) and the reference contigs were constructed by *de novo* assembly using Trinity (2). To identify the genes used in this study, assembled sequences were compared with the protein database of *Drosophila* by blastx. Transcripts of each sample were quantified by RSEM with Bowtie2 Aligner and differential expression of genes between Sym and *RpDuox*-RNAi insects was determined with DESeq2 (3). To visualize relatively up- or down-regulated genes, heatmap presentations were produced with the gplots package in program R ver. 3.6.2.

Protein domain and phylogenetic analysis

The transmembrane and functional domains of RpDuox were predicted by Phobius and

InterPro. The amino acid sequences of Duox proteins of selected organisms were retrieved from DDBJ/ENA/GenBank database and aligned using the MAFFT version 7 (4). Phylogenies were inferred based on the maximum likelihood methods with the Jones-Taylor-Thornton model (5) and 1,000 bootstrap replications, using the MEGA version 7.0.26 (6).

Statistical analyses

All statistical analyses were performed using the program R ver. 3.6.2 and GraphPad Prism ver. 8.3.1.

Data availability

RNA-seq data of the M4 of $\mathrm{Sym}^{\mathrm{control}}$ and $\mathrm{Sym}^{\mathrm{RpDuox\text{-}RNAi}}$ are available under National Center for Biotechnology Information (NCBI) BioProject PRJDB9456 with DRA accession numbers DRR215696-DRR215701. The gene sequences of *RpDuox*, *RpBnl*, *RpTrh*, *RpSima*, and *RpRelish* are deposited in GenBank with accession number MT270146-MT270150. The sequences of Duox of *Gryllus bimaculatus*, *Bombyx mori*, and *Tribolium castaneum* are deposited in GenBank with accession number MT270151, NW_004582026, and NC_007418, respectively.

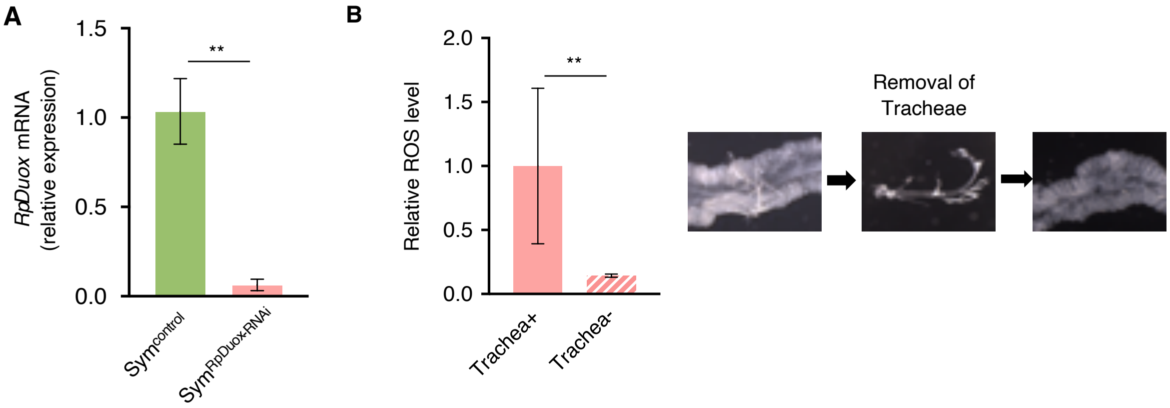

Fig. S1. Duox-mediated generation of ROS from the tracheal network in *R. pedestris.* (A) The efficiency of *RpDuox*-RNAi. The expression level of *RpDuox* was quantified by RT-qPCR after the injection of *RpDuox* dsRNA. **(B)** Relative ROS levels in the M4 of Sym$^{\mathrm{control}}$ insects before and after removing tracheae (left); photographs showing the removal of the tracheae (right). Data shown in **(A and B)** are mean ± s.d. $n = 5$ insects. Asterisks indicate statistically significant differences (**, $P < 0.01$). The statistical significance among samples was analyzed by the Mann-Whitney U test.

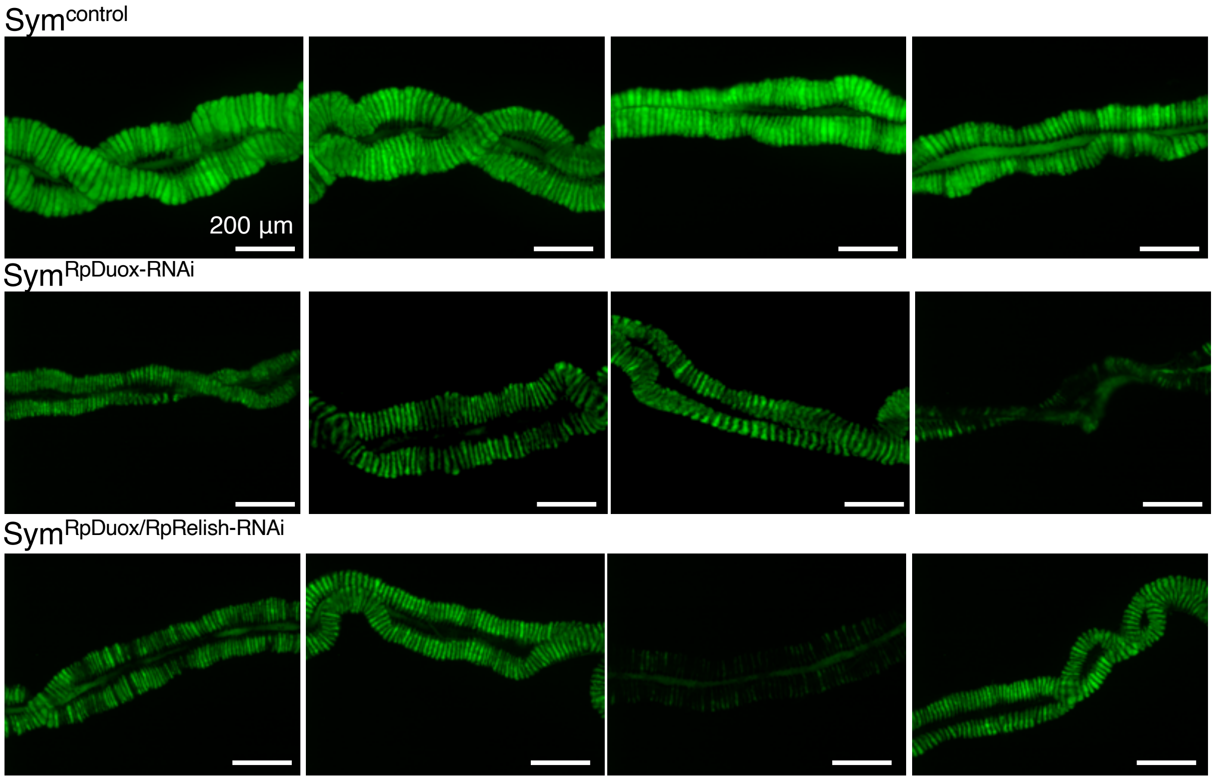

Fig. S2. Effect of *RpDuox/RpRelish* double RNAi on gut symbiosis. Effect of *RpDuox/RpRelish*-double RNAi on symbiosis. Green color indicates the GFP signal from gut-colonizing *Burkholderia* symbionts.

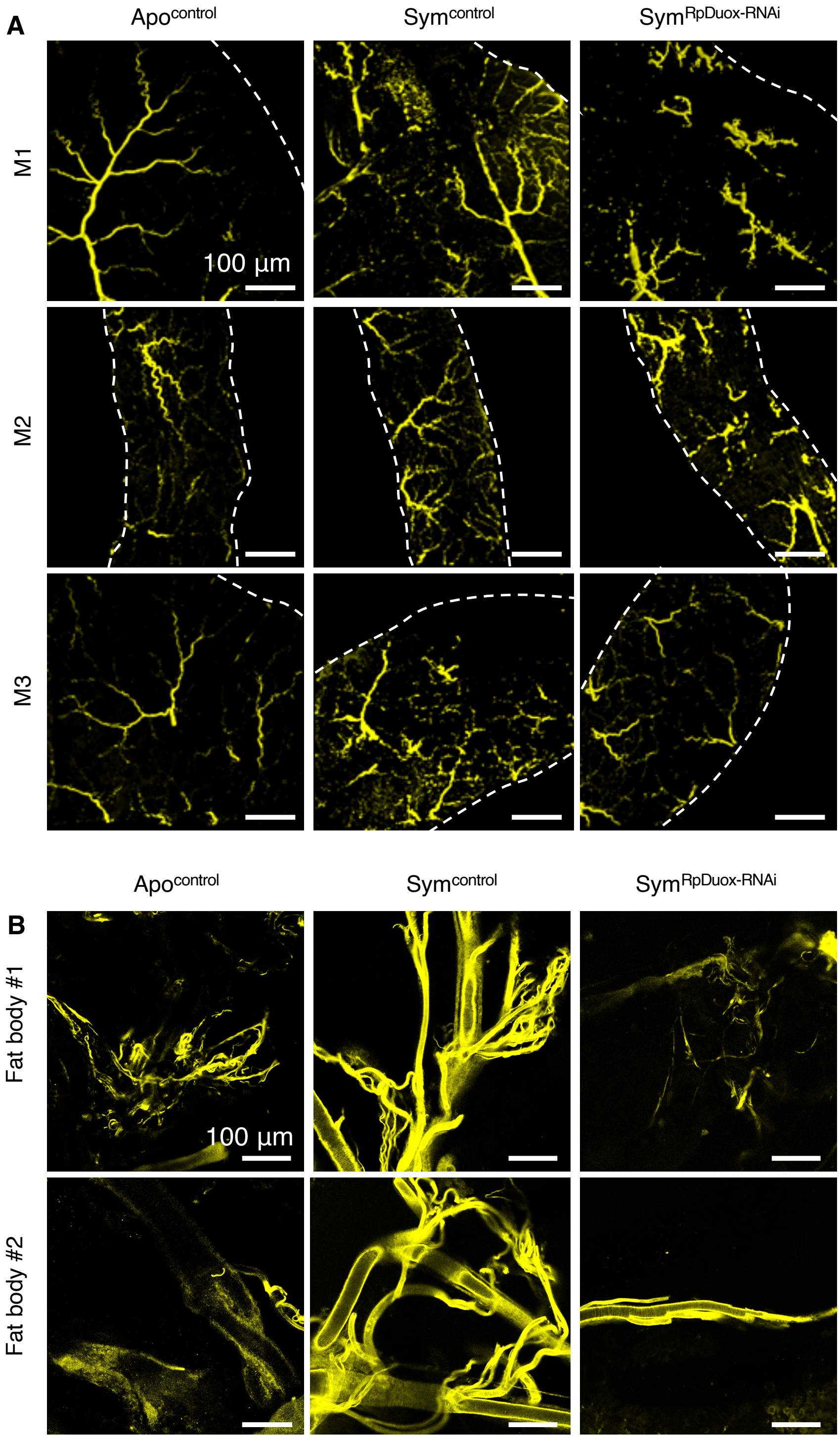

Fig.S3. Dityrosine network (DTN) in the tracheae of the midgut regions and the fat body. (A) Dissected M1, M2, M3, and **(B)** fat body of Apo (Apo$^{\text{control}}$), Sym (Sym$^{\text{control}}$) and *RpDuox*-RNAi (Sym$^{\text{RpDuox-RNAi}}$) insects were stained by anti-dityrosine antibody. The tracheoles were much more ramified in Sym insects, compared to Apo or *RpDuox*-RNAi insects. Yellow color indicates the DTN of tracheae.

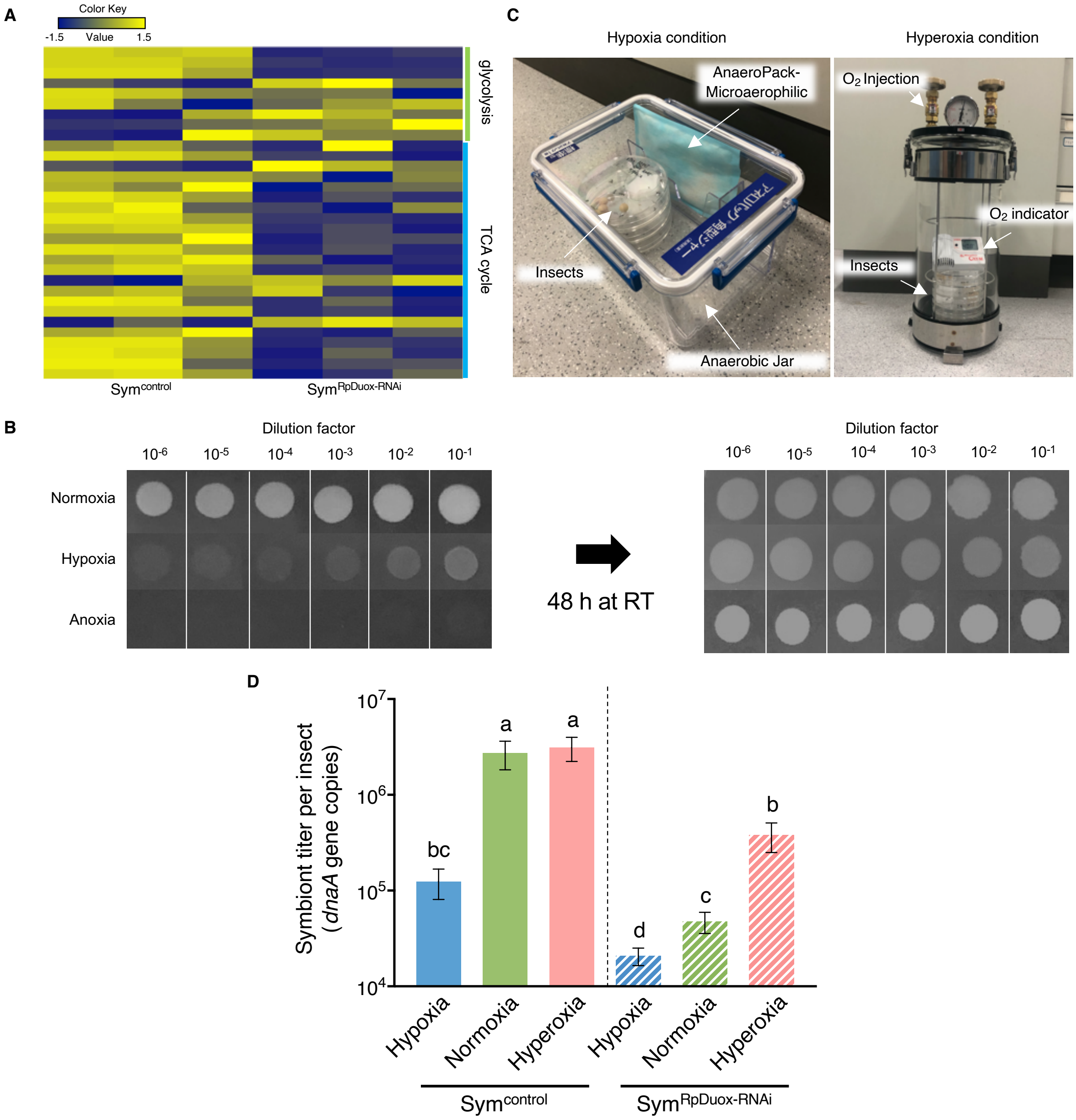

Fig. S4. Effect of oxygen level on host metabolism and *Burkholderia* symbiont. (A) The expression level of the genes related to glycolysis and TCA cycle in Sym^{control} and Sym$^{RpDuox-RNAi}$ insects. Each column is a replicate experiment. **(B)** Growth of the wild-type *Burkholderia* gut symbiont on agar plate in aerobic (21% O$_2$), microaerophilic (6-12% O$_2$), and anaerobic (<0.1% O$_2$) environment for 48 h (left panel) and growth after transfer of plates to normoxia and further incubation for 48 h (right panel). **(C)** Images of the hypoxia (6-12% O$_2$) and hyperoxia (39-42% O$_2$) set ups. **(D)** The number of gut symbionts in the M4 of Sym^{control} and Sym$^{RpDuox-RNAi}$ insects reared in hypoxia, normoxia, and hyperoxia condition. Data shown in **(D)** are mean ± s.d. n = 10 insects. Different letters indicate statistically significant differences (P < 0.05). The statistical significance of differences between samples was analyzed by a Kruskal-Wallis test with Bonferroni correction.

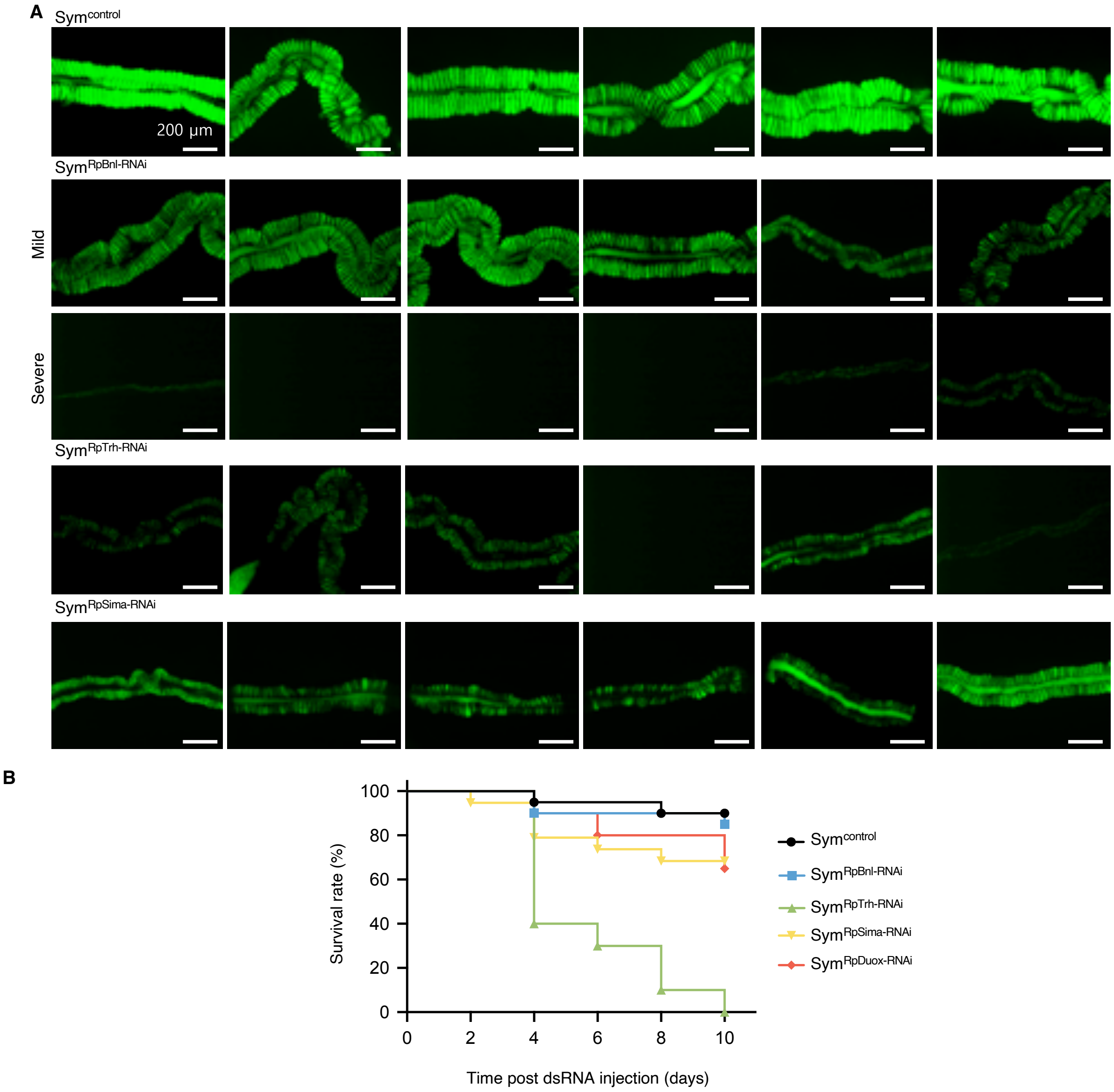

Fig. S5. Effect of *RpBnl*-, *RpTrh-,* and *RpSima*-RNAi on symbiosis and survival of insects. (A) Fluorescence microscopy images of the symbiotic organ of Sym (Sym$^{\text{control}}$), *RpBnl*-RNAi (Sym$^{\text{RpBnl-RNAi}}$), *RpTrh*-RNAi (Sym$^{\text{RpTrh-RNAi}}$), and *RpSima*-RNAi (Sym$^{\text{RpSima-RNAi}}$) insects. In Sym$^{\text{RpBnl-RNAi}}$ insects, “Mild” are examples in which symbiosis is moderately affected by RNAi and “Severeare examples in which symbiosis is seriously affected by RNAi. **(B)** The survival rate of Sym$^{\text{control}}$, Sym$^{\text{RpBnl-RNAi}}$, Sym$^{\text{RpTrh-RNAi}}$, and Sym$^{\text{RpSima-RNAi}}$ insects in days after the silencing of the target genes. *n* = 20 insects.

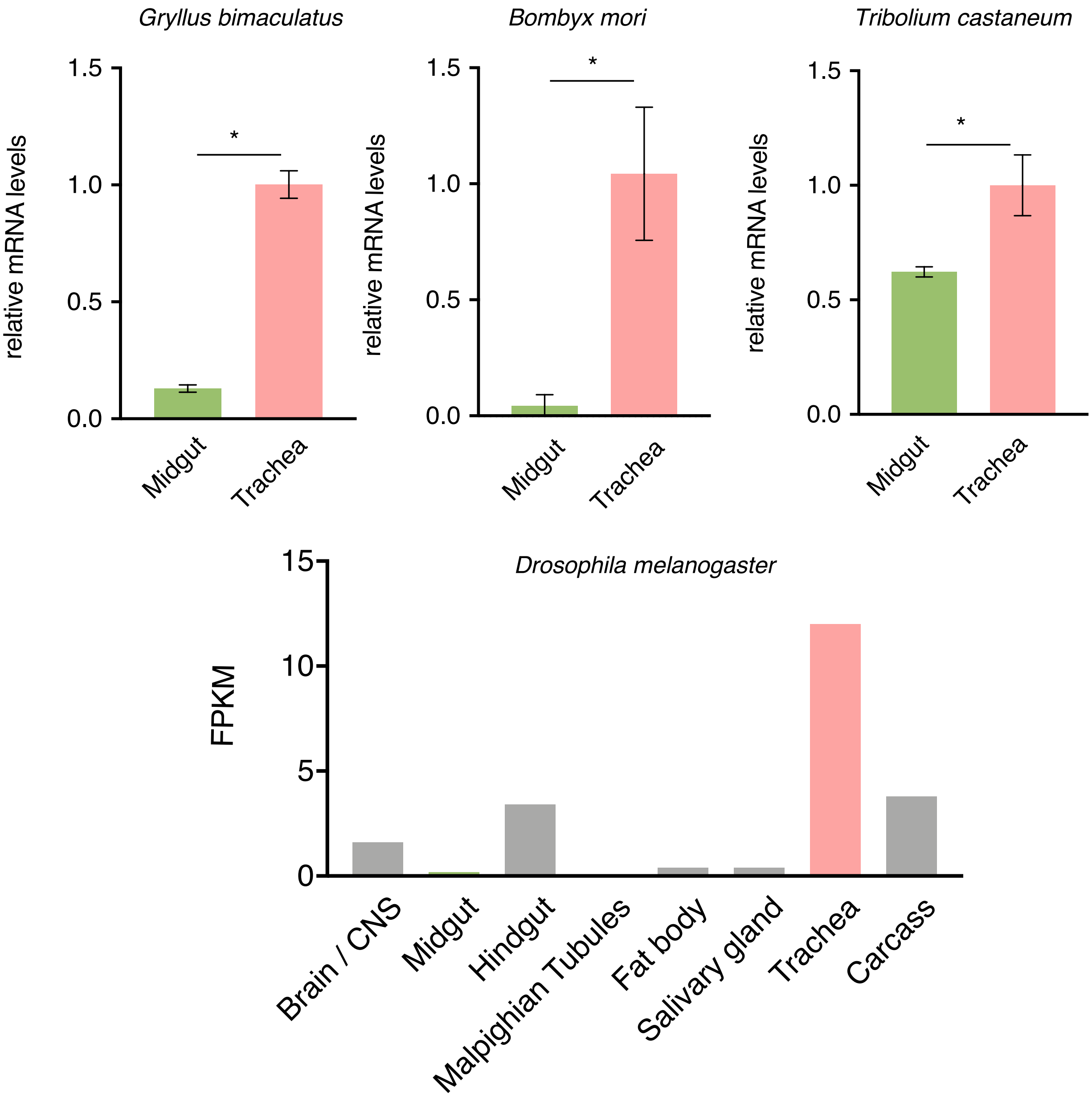

Fig. S6. The expression level of *Duox* in phylogenetically different orders of insects. The mRNA levels of *Duox* were measured from extracted midgut and tracheae of nymphal *G. bimaculatus* (Orthoptera), larval *B. mori* (Lepidoptera), and adult *T. castaneum* (Coleoptera), by RT-qPCR (top). Data shown in a,b are mean ± s.d. n = 4 insects. The statistical significance among samples was analyzed by the Mann-Whitney U test. Asterisks indicate statistically significant differences ($P < 0.05$). The FPKM (fragments per kilobase of exon model per million reads mapped) values of the *Duox* gene in larval *D. melanogaster* (Diptera) were extracted from the FlyAtlas 2 database (http://flyatlas.gla.ac.uk/FlyAtlas2/index.html) (bottom). CNS, central nerval system.

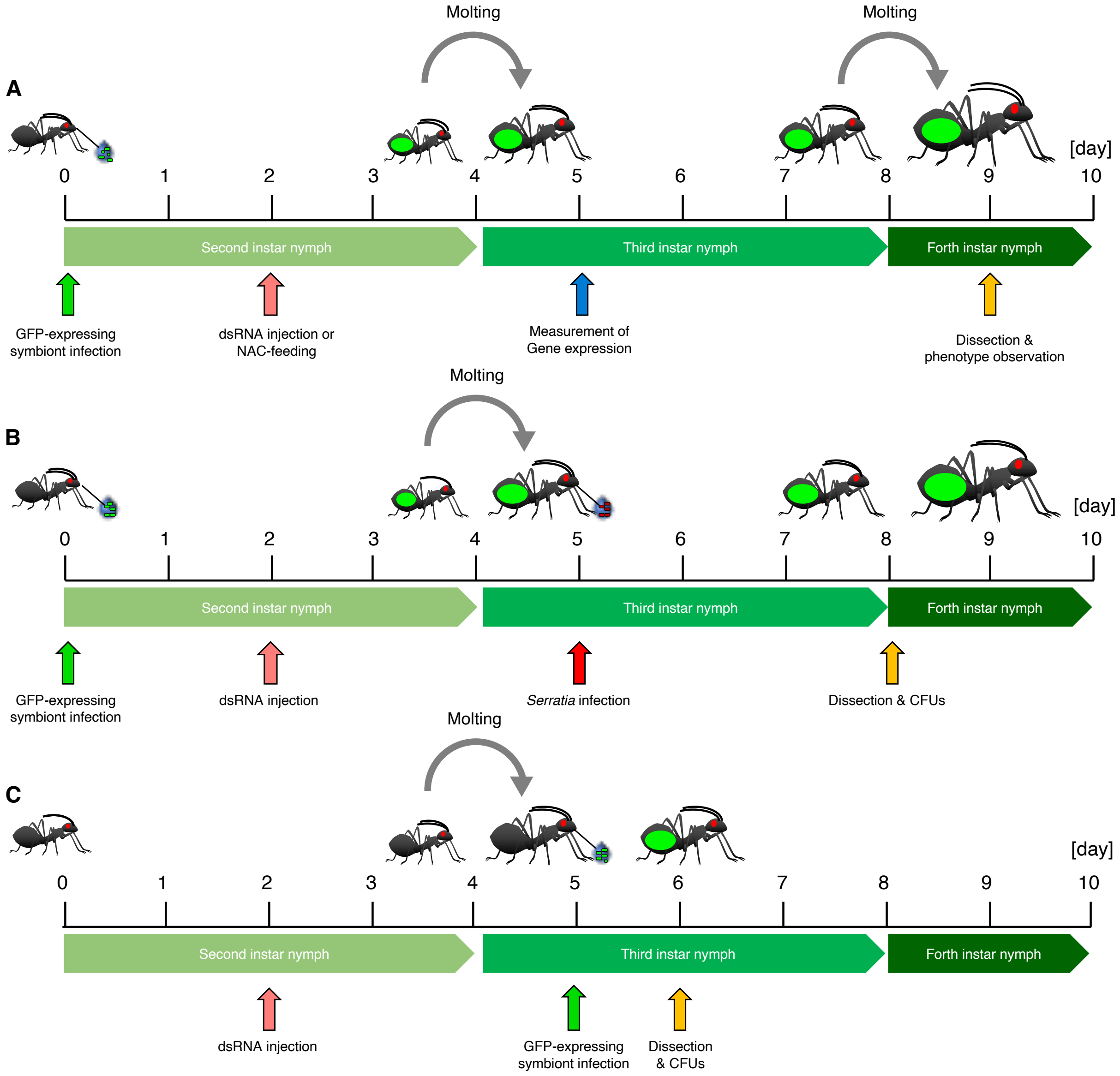

Fig. S7. The timeline of experiments used in this study. (A) The timeline of the experiments for measuring gene expression and phenotypic observation. **(B)** The timeline of the experiment for *Serratia* CFU. **(C)** The timeline of the experiment for *Burkholderia* CFU; CFU, colony-forming units.

Table 1. Bacterial strains

Bacteria	Description	Fluorescence	Reference[c]
Burkholderia insecticola symbiont			
RPE75	Spontaneous Rf^{r} mutant of wild-type *Burkholderia insecticola* symbiont RPE64	-	(1)
RPE225	GFP-labeled mutant of *Burkholderia* symbiont RPE75	GFP[a]	(2)
resB::pVO-GFP	Cytochrome c biogenesis mutant	GFP[b]	This study
ΔuppP	Cell wall synthesis mutant	GFP[a]	(3)
purL::Tn5	Purine biosynthesis mutant	GFP[a]	(4)
fliC::Tn5	Flagella mutant	GFP[a]	(5)
Escherichia coli			
DH5α	F- Φ80d*lac*ZΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1* *hsdR17* (r_K–, m_K+) *phoA supE44* λ- *thi*-1 *gyrA96 relA1*		Toyobo
Serratia marcescens			
Db11	Spontaneous Rf^{r} mutant of *S. marcescens* Db10		(6)

[a] GFP was labelled by labelled by the Tn7 minitransposon system

[b] GFP was labelled by insertion of GFP-expressing plasmid pVO-GFP into *resB* gene

[c] references: **(1)** Kikuchi Y, Meng XY, Fukatsu T. Gut symbiotic bacteria of the genus *Burkholderia* in the broad-headed bugs *Riptortus clavatus* and *Leptocorisa chinensis* (Heteroptera: Alydidae). *Appl Environ Microbiol* 2005; 71: 4035-4043; **(2)** Kikuchi Y, Fukatsu T. Live imaging of symbiosis: spatiotemporal infection dynamics of a GFP-labelled *Burkholderia* symbiont in the bean bug *Riptortus pedestris*. *Mol Ecol* 2014; 23: 1445-1456; **(3)** Kim JK, Lee HJ, Kikuchi Y, Kitagawa W, Nikoh N, Fukatsu T et al. Bacterial cell wall synthesis gene *uppP* is required for *Burkholderia* colonization of the stinkbug gut. *Appl Environ Microbiol* 2013; 79(16): 4879-4886; **(4)** Kim JK, Jang HA, Won YJ, Kikuchi Y , Han SH , Kim CH et al . Purine biosynthesis-deficient *Burkholderia* mutants are incapable of symbiotic accommodation in the stinkbug. *ISME J* 2014; 8: 552-63; **(5)** Ohbayashi T, Takeshita K, Kitagawa W, Nikoh N, Koga R, Meng XY et al. Insect's intestinal organ for symbiont sorting. *Proc Natl Acad Sci USA* 2015; 112: E5179-E5188; **(6)** Kurz CL, Chauvet S, Andrès E, Aurouze M, Vallet I, Michel GP et al. Virulence factors of the human opportunistic pathogen *Serratia marcescens* identified by *in vivo* screening. *EMBO J.* 2003 1;22(7):1451-60.

Table 2. Primers

Primers	Sequence (5'-3')	Tm (°C)	Description	Product size (bp)
bsDnaA_F	AGCGCGAGATCAGACGGTCGTCGAT	70.7	Measurement of symbiont *dnaA* gene copies by qPCR	150
bsDnaA_R	TCCGGCAAGTCGCGCACGCA	66.6		
T7_pT7bR20	TAATACGACTCACTATAGGGGATCTACTAGTCATATGGAT	74	Used for dsRNA synthesis	Depend on the insert size
T7_pT7bU19_2	TAATACGACTCACTATAGGGGACGGCCAGTGAAT	74.1		
Duox_dsF	TGCCCTGCTGACATTCTCTA	58.4	Used for dsRNA synthesis targeting *RpDuox*	451
Duox_dsR	AGTTGTATCCTCCCGCTCTG	60.5		
Duox_qF	CATTGGCGGGATGTTAGAGT	51.8	Used for qPCR targeting *RpDuox*	139
Duox_qR	TTCTTCGGGTGTGAAAATCC	58.4		
Bnl_dsF	TCCCCTCAGCTAGCTCACAT	60.5	Used for dsRNA synthesis targeting *RpBnl*	490
Bnl_dsR	TGGGTCTAGGTGGAGCTGAG	62.5		
Bnl_qF	AACAGAAGGATCAGGCACCG	60.5	Used for qPCR targeting *RpBnl*	117
Bnl_qR	TGTGTCGGTATGGTCCTCCT	60.5		
Trh_dsF	CCATCACCTCCCAACTGGAC	62.5	Used for dsRNA synthesis targeting *RpTrh*	581
Trh_dsR	GAGGGTTGACTTCATCCGCA	60.5		
Trh_qF	GACGCACACAGACTGAGGAA	60.5	Used for qPCR targeting *RpTrh*	151
Trh_qR	GCTCGTCTGCGTTTTTCGAA	58.4		
Sima_dsF	TGTCCACTGGAGCAGTCAAG	60.5	Used for dsRNA synthesis targeting *RpSima*	516
Sima_dsR	GGTTTCTCCCTTTGCTGGAT	58.4		
Sima_qF	AGGATCATTTCCGGTTTGTG	56.4	Used for qPCR targeting *RpSima*	134
Sima_qR	CATCCTTTCATCCGCATAGG	58.4		
Ef1a_F	CCTGCATCCGTTGCTTTTGT	58.4	Used for qPCR targeting *rEf1α*	150
Ef1a_R	GGCATCGAGGGCTTCAATAA	58.4		
Relish_dsF	GCTGGCACGAAAACTACCTAA	59.5	Used for dsRNA synthesis targeting *RpRelish*	525
Relish_dsR	TGTACATTCTGAGGGCCAAAG	59.5		
Relish_qF	GCCAGATGAATATTGTCGAATG	58.4	Used for qPCR targeting *RpRelish*	107
Relish_qR	TAACAGCCGACTGCGAGAG	59.5		
Riptocin_F	TCCGAAGCTGAGGGTCTTCCCG	67.9	Used for qPCR targeting *riptocin*	128
Riptocin_R	TCCGCATCCAAGTTCGCGTCC	65.3		
Defensin_F	TCGGTCGGACTGAGACTGAA	60.5	Used for qPCR targeting *rDefensin*	108
Defensin_R	TTGCCGCCTTTGTATCCCTT	58.4		
Thanatin_F	GTCTGCCTTCGTTGAAGACG	60.5	Used for qPCR targeting *rThanatin*	108

Thanatin_R	ATTCGCTTGCAAACGCCG	56.3		
BmRp49_F	CAGGCATCAATCGGATCGCT	60.5	Used for qPCR targeting *BmRp49*	150
BmRp49_R	GGGAGCATATGACGGGTCTT	60.5		
BmDuox_F	GGTGGGATGCTGGAGTCTAA	60.5	Used for qPCR targeting *BmDuox*	103
BmDuox_R	TTTCGAACCAAAATCGATCA	52.3		
GbEf1a_F	ATTGATGCCCCTGGTCACAG	60.5	Used for qPCR targeting *GbEf1α*	104
GbEf1a_R	TCAAACTCACCAGTACCCGC	60.5		
GbDuox_F	CCAAAGATTCGTCTCGAAGG	58.4	Used for qPCR targeting *GbDuox*	149
GbDuox_R	TACCGATTGGTTGAGGTTCC	58.4		
TcRps18_F	CTGCTGAAGTGAGGTCTGGG	62.5	Used for qPCR targeting *TcRps18*	141
TcRps18_R	CGCAATTTGGAGTCAAGCGT	58.4		
TcDuox_F	GTATTCGCTCACCCGACAAT	58.4	Used for qPCR targeting *TcDuox*	147
TcDuox_R	GCCGATGAAAAATAGCCAAA	54.3		

SI references

1. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
2. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494-1512 (2013).
3. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
4. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780 (2013).
5. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* **8**, 275-282 (1992).
6. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870-1874 (2016).