ISCI, Volume *19*

Supplemental Information

Commensal Gut Bacteria Buffer the Impact of Host Genetic Variants on *Drosophila* Developmental Traits under Nutritional Stress

Dali Ma, Maroun Bou-Sleiman, Pauline Joncour, Claire-Emmanuelle Indelicato, Michael Frochaux, Virginie Braman, Maria Litovchenko, Gilles Storelli, Bart Deplancke, and François Leulier

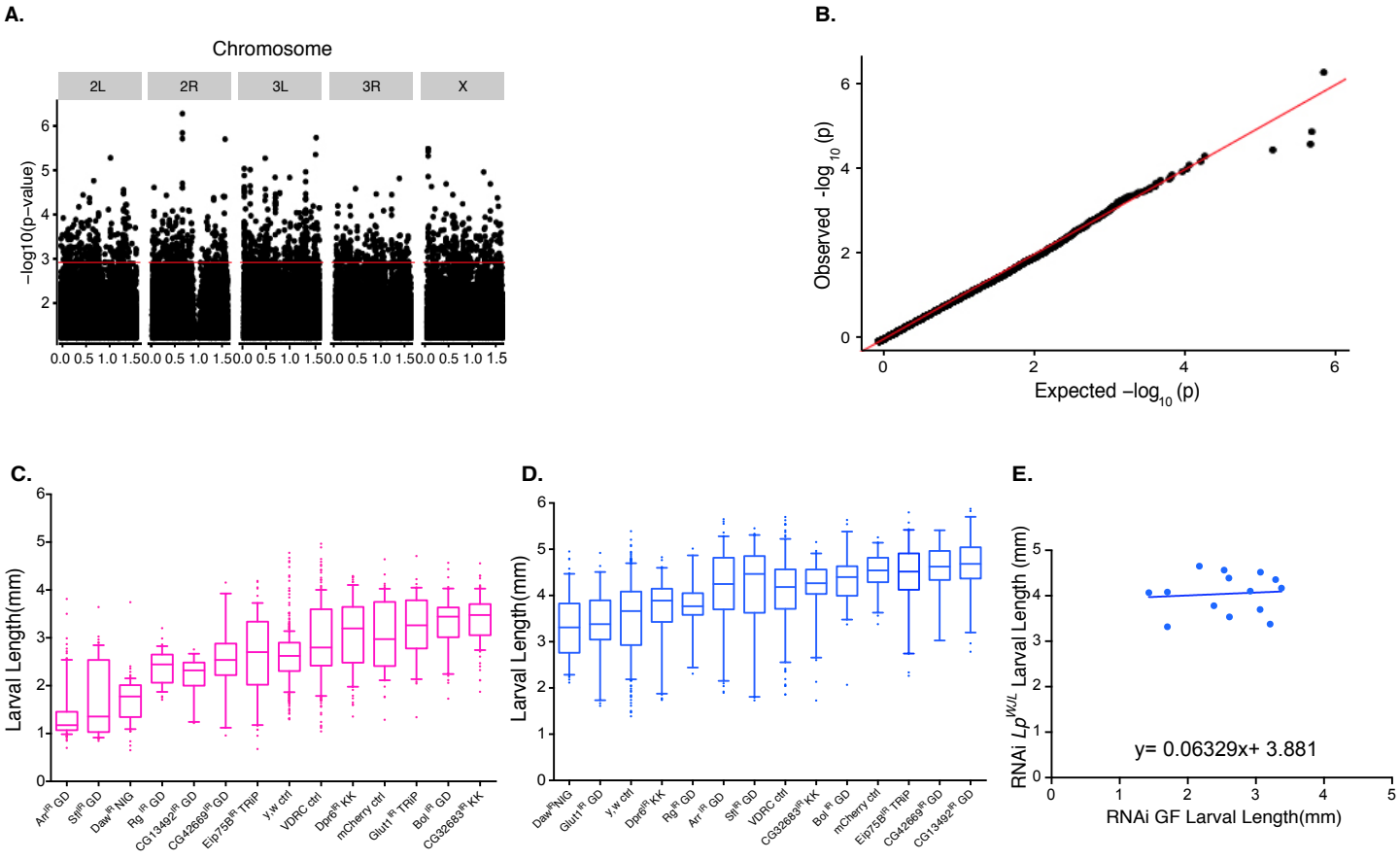

A.
Chromosome
2L
2R
3L
3R
X
−log10(p−value)
B.
Observed -log10 (p)
Expected –log10 (p)
C.
Larval Length(mm)
D.
Larval Length(mm)
E.
RNAi Lpwk/L Larval Length (mm)
y= 0.06329x+ 3.881
RNAi GF Larval Length(mm)

Figure S1

Figure S1. GWAS discovery of the growth promotion effect by *Lp*^{WJL} unexpectedly unveils the microbial buffering capacity in different host genetic backgrounds. Related to Figure 1

A). Manhattan plot of the GWAS performed on the average larval length fold change per DGRP line. We used the DGRP2 website for the association analysis. (http://dgrp2.gnets.ncsu.edu/)(Huang et al., 2014; Mackay et al., 2012).

B). Quantile-Quantile plot of the GWAS results.

C). and D). Box and whiskers plots illustrating the effect of RNAi knockdown on larval length on day 7 AEL. Each bar represents the average length from pooled 3-5 biological replicates from either condition, with 15-40 larvae in each replicate. **C:** GF. **D:** *Lp*^{WJL}. Three different control knockdowns were used: one control fly strain recommended by VDRC for RNAi constructs obtained from VDRC, one control strain (against mCherry) recommended by the Harvard TRiP collection, and the *y,w* strain from Bloomington. All control and RNAi strains were crossed to *y,w;; tubulin*-GAL80^{ts} ,*daugtherless*-GAL4. "GD" refers to the VDRC RNAi GD collection. "KK" refers to the VDRC RNAi KK collection. For specific genotypes, refer to Material and Methods.

E). *Lp*^{WJL} also buffers growth differences in the RNAi knock-down experiments for each of the candidate genes. Each data point represents the intercept of the average GF length and its corresponding mono-associated average larval length on Day 7 for each RNAi knockdown experiment. (Null hypothesis: Slope =1. *P=0.0008* , the null hypothesis is therefore rejected). These data points were fitted into an unconstraint model. For specific genotypes, we refer to Table 2 and Methods. Data are represented as mean and 10-90 percentile in all panels.

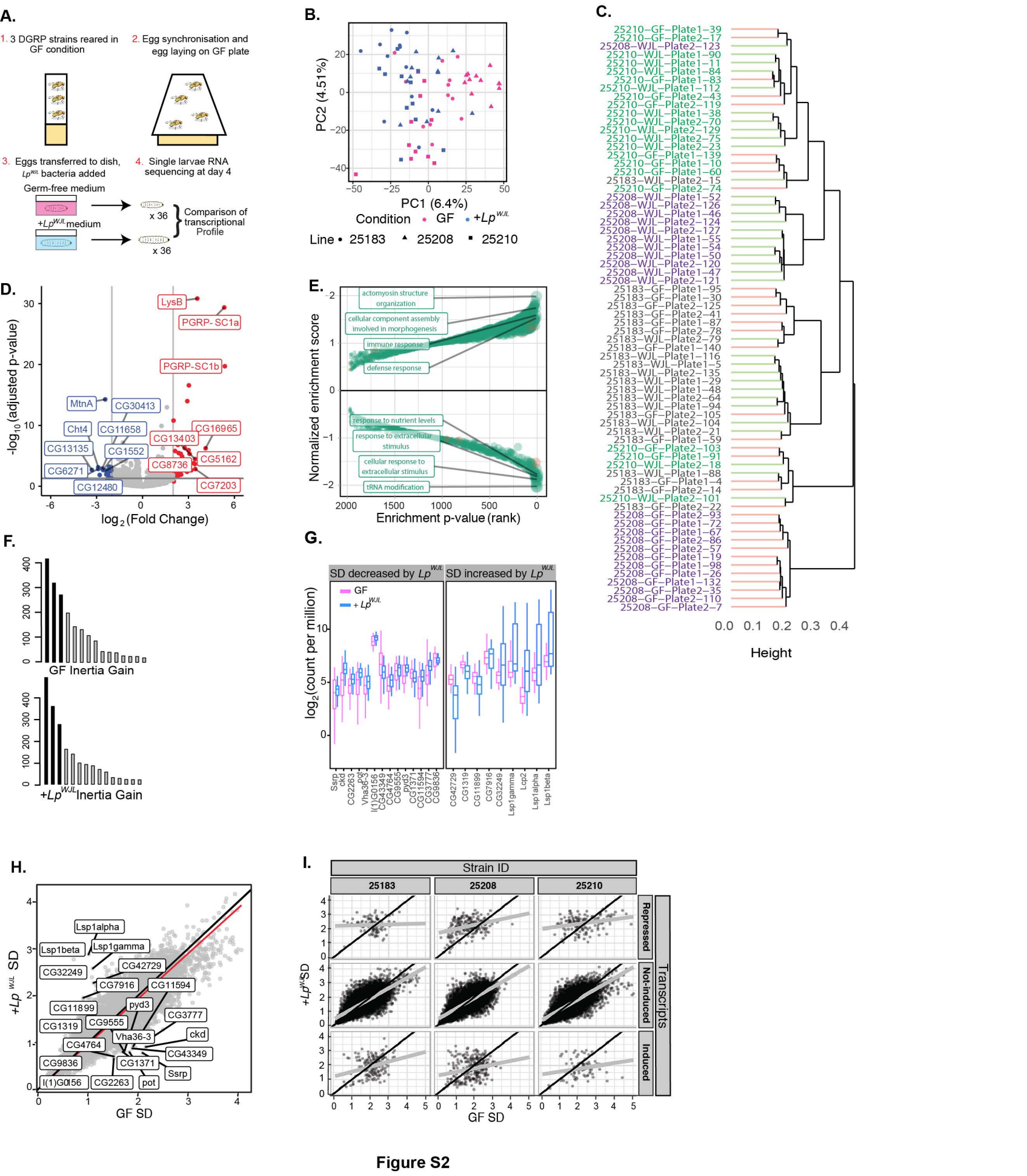

A.
1. 3 DGRP strains reared in GF condition
2. Egg synchronisation and egg laying on GF plate
3. Eggs transferred to dish, Lp^WJL bacteria added
4. Single larvae RNA sequencing at day 4
Germ-free medium
+Lp^WJL medium
x 36
x 36
Comparison of transcriptional Profile
B.
PC2 (4.51%)
PC1 (6.4%)
Condition GF +Lp^WJL
Line 25183 25208 25210
C.
25210–GF–Plate1–39
25210–GF–Plate2–17
25208–WJL–Plate2–123
25210–WJL–Plate1–90
25210–WJL–Plate1–11
25210–WJL–Plate1–84
25210–GF–Plate1–83
25210–WJL–Plate1–112
25210–GF–Plate2–43
25210–GF–Plate2–119
25210–WJL–Plate1–38
25210–WJL–Plate2–70
25210–WJL–Plate2–129
25210–WJL–Plate2–75
25210–WJL–Plate2–23
25210–GF–Plate1–139
25210–GF–Plate1–10
25210–GF–Plate1–60
25183–WJL–Plate2–15
25210–GF–Plate2–74
25208–WJL–Plate1–52
25208–WJL–Plate2–126
25208–WJL–Plate1–46
25208–WJL–Plate2–124
25208–WJL–Plate2–127
25208–WJL–Plate1–55
25208–WJL–Plate1–54
25208–WJL–Plate1–50
25208–WJL–Plate2–120
25208–WJL–Plate1–47
25208–WJL–Plate2–121
25183–GF–Plate1–95
25183–GF–Plate1–30
25183–GF–Plate2–125
25183–GF–Plate2–41
25183–GF–Plate1–87
25183–GF–Plate2–78
25183–WJL–Plate2–79
25183–GF–Plate1–140
25183–WJL–Plate1–116
25183–WJL–Plate1–5
25183–WJL–Plate2–135
25183–WJL–Plate1–29
25183–WJL–Plate1–48
25183–WJL–Plate2–64
25183–WJL–Plate1–94
25183–GF–Plate2–105
25183–WJL–Plate2–104
25183–WJL–Plate2–21
25183–GF–Plate1–59
25210–GF–Plate2–103
25210–GF–Plate1–91
25210–WJL–Plate2–18
25183–WJL–Plate1–88
25183–GF–Plate1–4
25183–GF–Plate2–14
25210–WJL–Plate2–101
25183–GF–Plate2–22
25208–GF–Plate2–93
25208–GF–Plate1–72
25208–GF–Plate1–67
25208–GF–Plate2–86
25208–GF–Plate2–57
25208–GF–Plate1–19
25208–GF–Plate1–98
25208–GF–Plate1–26
25208–GF–Plate1–132
25208–GF–Plate2–35
25208–GF–Plate2–110
25208–GF–Plate2–7
Height
D.
-log10 (adjusted p-value)
log2 (Fold Change)
LysB
PGRP- SC1a
PGRP-SC1b
MtnA
CG30413
Cht4
CG11658
CG16965
CG13403
CG13135
CG1552
CG5162
CG8736
CG6271
CG12480
CG7203
E.
Normalized enrichment score
Enrichment p-value (rank)
actomyosin structure organization
cellular component assembly involved in morphogenesis
immune response
defense response
response to nutrient levels
response to extracellular stimulus
cellular response to extracellular stimulus
tRNA modification
F.
GF Inertia Gain
+Lp^WJL Inertia Gain
G.
SD decreased by Lp^WJL
SD increased by Lp^WJL
GF
+Lp^WJL
log2(count per million)
Ssrp
ckd
CG2263
pot
Vha36-3
I(1)G0156
CG43349
CG4764
CG9555
pyd3
CG1371
CG11594
CG3777
CG9836
CG42729
CG1319
CG11899
CG7916
CG32249
Lsp1gamma
Lcp2
Lsp1alpha
Lsp1beta
H.
+Lp^WJL SD
GF SD
Lsp1alpha
Lsp1beta
Lsp1gamma
CG42729
CG32249
CG7916
CG11594
CG11899
pyd3
CG3777
CG1319
CG9555
Vha36-3
ckd
CG4764
CG43349
CG9836
CG1371
Ssrp
I(1)G0156
CG2263
pot
I.
Strain ID
25183
25208
25210
Repressed
Not-induced
Induced
Transcripts
+Lp^WJL SD
GF SD

Figure S2

Figure S2. The single-larva BRB-seq indicates transcriptomic buffering in developmental genes by *Lp*^{WJL}. Related to Figure 1

A). Experimental setup to perform BRB-seq-based transcriptomics on individual larvae. Flies from three DGRP strains were reared in GF conditions. Egg-laying was synchronized and embryos were transferred to food caps: three left germ-free (1X PBS) and three inoculated with *Lp*^{WJL} (OD 0.5 in 1x PBS). At day 4, single larvae were collected from all plates, RNA extraction and RNA sequencing were performed. 12 larvae were collected per line for each condition, totaling 72 single larval transcriptomes.

B). Principal component plot of the corrected expression data after batch correction.

C). Hierarchical clustering of the transcriptomic data using the Ward's method. A batch effect of plate was corrected prior to clustering. The genotypes are color-coded (Green: 25210, violet: 25208, black: 25183). The red "branches" of the cluster represent GF samples, and green ones represent mono-associated samples.

D). The observed effect of *Lp*^{WJL} mono-association on gene expression is consistent with our previous findings, thus validating our transcriptome approach on individual larvae. The horizontal grey line represents the 0.05 FDR-corrected p-value threshold. The vertical lines are the -2 and 2 log2 (Fold Change) thresholds. Genes in red are significantly up-regulated, genes in blue are significantly down-regulated. Several representative genes of the top differentially regulated genes from each category are highlighted.

E). Gene set enrichment analysis on biological process gene ontology (GO) terms based on the effect of *Lp*^{WJL} mono-association. Gene sets in orange were derived from GLAD(Hu et al., 2015), whereas green gene sets were extracted from GO2MSIG(Powell, 2014).

F). Inertia gain of the HCPC analysis from Figure 1G and 1H. the black bars represent the "optimal" level of division of the tree suggested by FactoMineR.

G). Scatterplot of the standard deviation in expression level of each gene in the GF and *Lp*^{WJL} mono-associated condition. The black line represents the theoretical slope of 1 and intercept 0. The red line is a linear fit of the points. Labelled genes show the highest relative change in their standard deviation, as determined by the absolute value of $\log_2(SD_{LpWJL}/SD_{GF})$.

H). Box and whiskers plots showing the expression levels of genes with high relative change in standard deviation, regardless whether the genes themselves were up- or down-regulated.

I). Scatterplots of standard deviations of each gene calculated by genotype. Genes were faceted by how their differential expression alters within each strain in both GF and *Lp*^{WJL} mono-associated conditions: repressed (top panel), non-induced (middle panel) and induced (bottom panel). The black lines represent the theoretical slope of 1 and intercepts 0, the grey lines are the linear fit to the data. Since transcripts specifically modulated by *Lp*^{WJL} tend to have incomparable SD, we assessed GO enrichment only on non-differentially expressed genes (**see Fig.1K**)

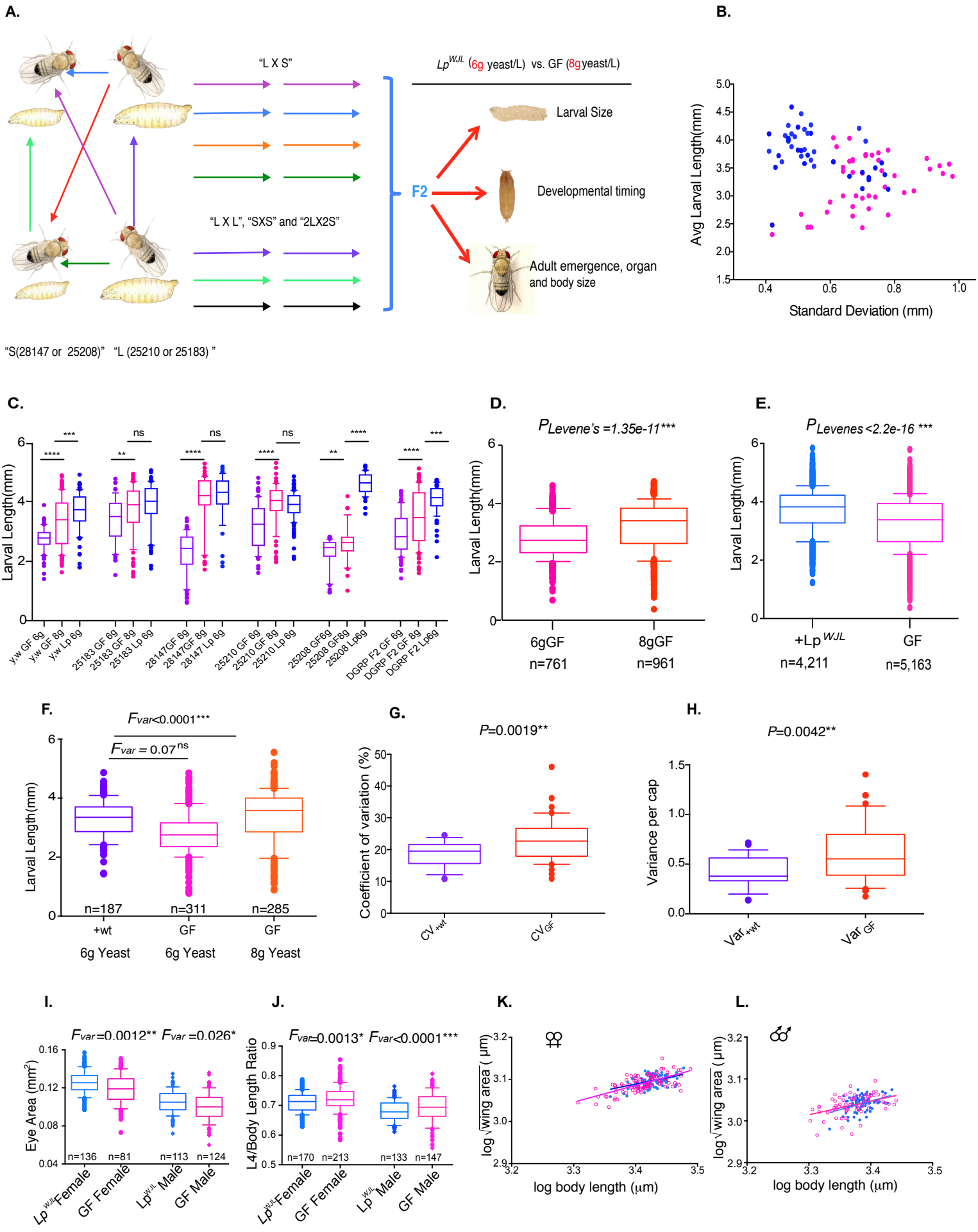

A.
"L X S"
"L X L", "SXS" and "2LX2S"
"S(28147 or 25208)" "L (25210 or 25183) "
Lp^WJL (6g yeast/L) vs. GF (8g yeast/L)
F2
Larval Size
Developmental timing
Adult emergence, organ and body size
B.
Avg Larval Length(mm)
Standard Deviation (mm)
C.
Larval Length(mm)
y,w GF 6g
y,w GF 8g
y,w Lp 6g
25183 GF 6g
25183 GF 8g
25183 Lp 6g
28147GF 6g
28147GF 8g
28147 Lp 6g
25210 GF 6g
25210 GF 8g
25210 Lp 6g
25208 GF6g
25208 GF8g
25208 Lp6g
DGRP F2 GF 6g
DGRP F2 GF 8g
DGRP F2 Lp6g
D.
P Levene's =1.35e-11***
Larval Length(mm)
6gGF
n=761
8gGF
n=961
E.
P Levenes <2.2e-16 ***
Larval Length(mm)
+Lp WJL
n=4,211
GF
n=5,163
F.
Fvar<0.0001***
Fvar = 0.07ns
Larval Length(mm)
n=187
n=311
n=285
+wt
GF
GF
6g Yeast
6g Yeast
8g Yeast
G.
P=0.0019**
Coefficient of variation (%)
CV+wt
CVGF
H.
P=0.0042**
Variance per cap
Var+wt
Var GF
I.
Fvar=0.0012** Fvar =0.026*
Eye Area (mm2)
n=136
n=81
n=113
n=124
Lp WJL Female
GF Female
Lp WJL Male
GF Male
J.
Fvar=0.0013* Fvar<0.0001***
L4/Body Length Ratio
n=170
n=213
n=133
n=147
Lp WJL Female
GF Female
Lp WJL Male
GF Male
K.
log √wing area (µm)
log body length (µm)
L.
log √wing area (µm)
log body length (µm)

Figure S3

Figure S3 In the genetically diverse DGRP F_2 population, *Lp*^{WJL} reduces variation in different physical fitness traits. Related to Figure 2

A). A diagram illustrating DGRP crosses to generate the F_2 generation for studying variation in larval size, pupariation and adult emergence. 25210 (RAL-859), 25183(RAL-335) are the lines with the "large" ("L") larvae as germ-free, and 25208(RAL-820) and 28147(RAL-158) are the lines with the "small" larvae as germ-free ("S"). Seven possible crosses are set up: 25210X25183 ("LXL"), 25208X28147("SXS"), 25210X25208, 25183X25208, 25210X28147, 25183X28147 are the four "LXS" crosses, and 25183 and 25210 X 25208 and 28147 is the "2L X 2S" cross.

B). A scatter plot showing how standard deviation (SD) changes as a function of larval length, and how such change differs in the DGRP F_2 GF (pink) and *Lp*^{WJL} mono-associated (blue) populations (see also Figure 2a and Methods for detailed schemes). Each data point represents the intercept of an SD value and its corresponding average larval length in a particular cross. Each SD and average length was derived from larvae measurements gathered from at least 3 biological replicates from either GF or *Lp*^{WJL} mono-associated conditions. Each replicate contains 10-40 larvae.

C). Larval lengths of axenic flies grown on media containing 6g (purple), 8g (pink) or 6g yeast with *Lp*^{WJL} inoculation (dark blue) on day 7 after egg-lay. Note that 2g extra yeast invariably boosts germ-free growth in different strains and genetic background. The asterisks indicate statistics differences when comparing average larval lengths between conditions.

D). Larval growth and variability comparison in DGRP F2 axenic larvae pooled from the parental strains (Figure S3C). For GF larvae raised on 6g/L yeast, average larval length =2.76mm, SD=0.66mm, CV=24.1%; for GF larvae raised on 8g/L yeast, average larval length =3.34mm, SD=0.85mm, CV=25.2%.

E). Box and Whisker graph illustrating the average length and standard deviation from pooled GF (pink) and *Lp*^{WJL} mono-associated DGRP (blue) F2 larvae, pooled from all the crosses in all three different repeats (Average GF larval length: 3.29mm; average Lp mono-associated larval length: 3.71mm; CV_{GF}=24.9%, CV_{Lp}=19.5%).

F). One representative experiment showing that re-associating the field-collected flies tends to buffer the variability in body length in size-matched larvae. The purple box represents body length from wild larvae grown on media contaminated with their untreated parents' fecal matter. Average GF larval length grown on 6g/L yeast media: 2.81mm; average GF larval length grown on 8g/L yeast media: 3.36mm: average re-associated larval length ("+wt"): 3.07 mm; P= 0.338. CV_{GF} (6g/L, pink) = 24.9%, CV_{GF} (8g/L, orange)= 27.0%, CVwt (purple)= 18.9%.

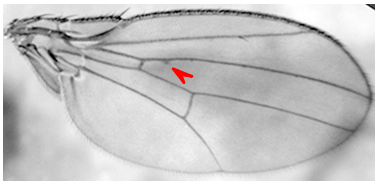
G). and H). The compiled CV values (**e**.) and variances (**f**.) derived from each low-yeast cap containing 40~50 field-collected larvae. The average CV and variance are lower in the population re-associated with its own microbes (purple) than in the GF population (orange)

I). In both male (lozenge) and female (circle) adults, the variances in eye size are greater in GF F_2 progeny. The difference in mean eye area, for females P<0.0001***; for males, P=0.0013**.

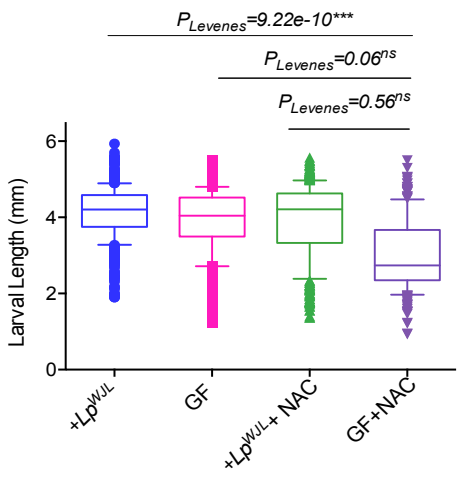
J). The length of the L4 vein in the wing is used as a proxy of the wing length. In the accumulated ratios of wing length over body length, the variances are greater in the GF flies (The difference in average L4/ body length, for females P<0.0028**; for males, P=0.02*).

K). and L). Scatter plots illustrating the allometric relationship between wing area and body size in female (i) and male (j) DGRP F_2 adults. Pink open circles: GF, blue filled circles: *Lp*^{WJL}. Each line represents the allometric slope of the data points shown by the same color. Either in males or females, there is no difference in allometric slope between the GF and mono-associated population. For GF females, Y_{GF} = 0.3963*X + 1.738, 95%C.I.= 0.3117 to 0.4810; for *Lp*^{WJL} females, Y_{Lp} = 0.2978*X + 2.076, 95%C.I.= 0,1785 to 0,4172, P=0.203, n.s ; for GF males, Y_{GF} = 0.3261*X + 1.939, 95%C.I.= 0.1725 to 0.4796 ; for *Lp*^{WJL} males, Y_{Lp}= 0.4141*X + 1.639, 95% C.I. =0.1842 to 0.6439, P=0.55, ns. Data are represented as Mean and 10-90 percentile in all panels.

A.



B.



C.

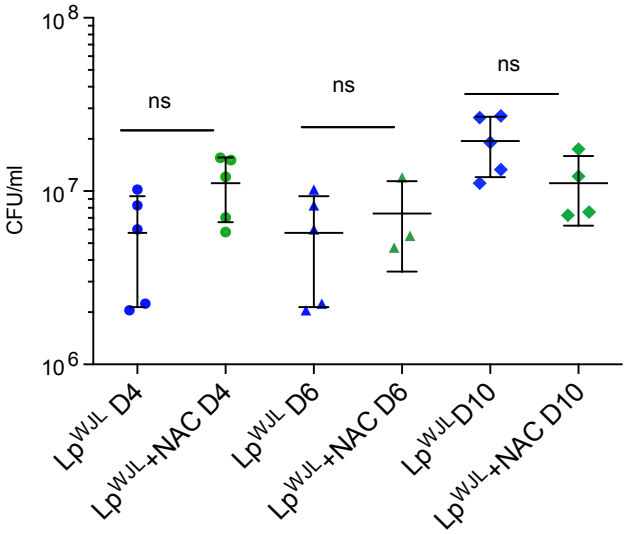


Figure S4

Figure S4 The LpWJL buffering in developmental trait and organ patterning robustness involves ROS signaling. Related to Figure 3 and 4.

A.) An image of a wing of an Lp^{WJL} adult is shown, as a representation of the most visible "defect" ever observed in mono-associated adults. Red arrow points to the subtle vein tissue thickening. We included these as "defects" in the Lp^{WJL} F_2 population in the analyses presented in Figure 3A, 3B, and 4F.

B). Germ-free larvae (light violet) that ingested NAC show comparable size variation to Lp^{WJL} larvae fed on NAC (McFall-Ngai et al.) or germ-free larvae who have not been exposed to NAC (pink).

C). Bacterial niche load (NL) evolution ("Niche" is defined as the substrate with both larvae and bacteria present) during the course of larval development with Lp^{WJL} with or without NAC treatment (Day 4, Day 6 and Day 10). Data are represented as mean ± SD.

TableS1. Average D7 larvae length for individual DGRP lines. Related to Figure 1

DGRP Lines	GF* Length(mm)	GF SD*(mm)	*Lp*^{WJL} * Length(mm)	*Lp*^{WJL} SD(mm)	*Lp*^{WJL}/GF Ratio
25174	2.193	0.584	3.637	0.895	1.658
25175	2.693	0.687	4.496	0.659	1.670
25176	1.443	0.536	3.903	0.648	2.704
25180	2.151	0.454	3.795	0.635	1.764
25181	2.374	0.824	4.224	0.946	1.779
25182	2.108	0.451	3.293	0.859	1.562
25183	2.961	0.657	4.066	0.548	1.373
25184	1.957	0.53	4.323	0.587	2.209
25185	2.459	0.681	3.93	0.722	1.598
25186	2.278	0.667	4.289	0.803	1.883
25187	2.109	0.479	3.798	0.744	1.801
25188	2.253	0.421	4.202	0.786	1.865
25189	2.586	0.393	3.448	0.876	1.333
25190	2.292	0.512	3.976	0.941	1.735
25191	2.348	0.428	3.953	0.797	1.684
25192	2.194	0.401	4.145	0.731	1.889
25193	2.414	0.582	4.05	0.782	1.678
25194	2.506	0.558	4.195	0.508	1.674
25195	2.07	0.402	3.635	0.867	1.756
25197	1.944	0.397	3.73	0.734	1.919
25198	2.051	0.394	3.936	0.673	1.919
25199	1.514	0.524	3.78	0.753	2.497
25200	2.869	0.752	4.227	0.605	1.473
25201	2.182	0.347	4.186	0.601	1.918
25202	2.273	0.639	3.85	0.792	1.694
25203	1.541	0.513	4.158	0.755	2.698
25204	1.686	0.678	4.088	0.774	2.425
25205	2.351	0.567	3.77	0.606	1.604
25206	2.5	0.643	4.173	0.619	1.669
25207	2.028	0.481	3.896	0.811	1.921
25208	1.649	0.443	4.103	0.947	2.488
25209	2.187	0.67	4.232	0.819	1.935
25210	2.772	0.633	4.03	0.466	1.454
25445	2.01	0.468	3.956	0.668	1.968
25744	2.097	0.34	4.235	0.666	2.020
25745	2.501	0.612	4.051	0.599	1.620
28132	2.828	0.684	4.485	0.534	1.586
28134	1.854	0.383	4.144	0.479	2.235
28136	1.707	0.415	4.204	0.548	2.463
28138	1.38	0.487	4.318	0.693	3.129
28142	2.938	0.836	4.487	0.489	1.527
28146	2.077	0.36	4.564	0.915	2.197
28147	1.575	0.552	4.061	0.728	2.578
28153	2.298	0.329	3.97	0.541	1.728
28154	2.256	0.339	4.365	0.482	1.935
28160	2.51	0.662	4.118	0.714	1.640
28164	2.394	0.448	4.207	0.584	1.757
28166	2.163	0.402	4.489	0.642	2.075
28173	2.039	0.309	4.122	0.697	2.022
28192	2.141	0.506	4.286	0.659	2.002
28194	2.269	0.565	4.424	0.72	1.950
28197	2.89	0.742	4.547	0.519	1.573
28208	2.339	0.438	4.14	0.705	1.767

*GF: germ-free

**Lp*^{WJL}: Lactobacillus plantarum, stain name: WJL

*SD: standard deviation

Table S2 . Variants associated with the growth benefits conferred by *Lactobacillus plantarum* (*Lp*^{WJL}). Related to Figure 1.

Variants	R^2	P-value	Minor allele	Major allele	Ref* allele	MAF*	Variant Class	Molecular and cellular functions
CG13492	46.46%	1.23E-06	C	T	C	0.245	intron	Unknown
	45.81%	4.526E-07	T	A	T	0.244		
	45.56%	1.65E-06	G	A	G	0.25		
CG32683	39.04%	2.76E-06	A	T	T	0.2453	Intron/ downstream	Unknown, arrestin-like
	39.04%	2.76E-06	A	C	C	0.2453		
	29.32%	4.03E-06	T	A	A	0.22		
	29.07%	3.19E-06	T	G	G	0.2245		
	29.80%	1.17E-05	CTGTTG	C	C	0.283		
CG33269	35.58%	8.21e-06	G	A	A	0.14	Intergenic	Unknown
dpr6	33.06%	2.94E-05	A	T	T	0.1224	Intron	Immunoglobulin-like domain; sensory perception of chemical stimulus
	21.34%	7.77E-06	A	G	G	0.08		
Eip75B	32.65%	1.22E-05	C	T	C	0.1176	Intron	Nuclear hormone receptor, ecdysone response, antimicrobial humoral response
rg	32.14%	9.25E-06	G	A	G	0.4	Intron	PKA-binding, cone cell differentiation, mushroom body development, olfactory learning
sfl	27.37%	9.18E-06	G	T	T	0.4706	Intron	heparan sulfate proteoglycans (HSPGs) biosynthesis/wg morphogen diffusion
CG42669	26.66%	1.23E-05	A	G	G	0.1373	Intron	Supervillin, actin-binding
bol	25.07%	3.76E-06	C	T	T	0.2	3'UTR	RNA binding protein. Role in meiotic entry and germline differentiation
CR43427, lncRNA566	23.7%	4,53E-06	G	T	T	0.3269	intergenic	Unknown, lncRNA
daw	15.1%	4.45E-06	T	C	C	0.1837	Synonymous substitution	TGF-β ligand: growth; regulation of insulin secretion
arr	14.68%	1.69E-06	G	C	C	0.1875	intron	wnt protein binding/canonical wnt pathway
glut1	11.14%	1.56E-06	G	T	T	0.2245	intron	General glucose/sugar transporter

*MAF: minor allele frequency in the 53 DGRP lines
*Ref allele: allele info derived from BDGP (Berkeley Drosophila Genome Project)
R^2 reflects effect size

Table S3. Individual larval transcriptome sample list. Related to Figure 1

SampleID	Genotype	Treatment	Plate	Individual	Well_Row	Well_Column	TotalReads	Timepoint
GF-d4-Plate1-25183-4	25183	GF	Plate1	4	D	1	3374679	d4
WJL-d4-Plate1-25183-5	25183	WJL	Plate1	5	E	2	4323699	d4
GF-d4-Plate2-25208-7	25208	GF	Plate2	7	E	9	1537636	d4
GF-d4-Plate1-25210-10	25210	GF	Plate1	10	D	5	3969828	d4
WJL-d4-Plate1-25210-11	25210	WJL	Plate1	11	E	6	5131500	d4
GF-d4-Plate2-25183-14	25183	GF	Plate2	14	E	1	3307084	d4
WJL-d4-Plate2-25183-15	25183	WJL	Plate2	15	D	2	2816461	d4
GF-d4-Plate2-25210-17	25210	GF	Plate2	17	E	5	5063082	d4
WJL-d4-Plate2-25210-18	25210	WJL	Plate2	18	D	6	4162852	d4
GF-d4-Plate1-25208-19	25208	GF	Plate1	19	D	9	2459570	d4
WJL-d4-Plate2-25183-21	25183	WJL	Plate2	21	E	2	2399808	d4
GF-d4-Plate2-25183-22	25183	GF	Plate2	22	D	1	4448517	d4
WJL-d4-Plate2-25210-23	25210	WJL	Plate2	23	E	6	4508569	d4
GF-d4-Plate1-25208-26	25208	GF	Plate1	26	E	9	2085683	d4
WJL-d4-Plate1-25183-29	25183	WJL	Plate1	29	D	2	1843092	d4
GF-d4-Plate1-25183-30	25183	GF	Plate1	30	E	1	3678838	d4
GF-d4-Plate2-25208-35	25208	GF	Plate2	35	D	9	3470625	d4
WJL-d4-Plate1-25210-38	25210	WJL	Plate1	38	D	6	3828526	d4
GF-d4-Plate1-25210-39	25210	GF	Plate1	39	E	5	4247231	d4
GF-d4-Plate2-25183-41	25183	GF	Plate2	41	F	1	1761823	d4
GF-d4-Plate2-25210-43	25210	GF	Plate2	43	F	5	3169382	d4
WJL-d4-Plate1-25208-46	25208	WJL	Plate1	46	C	10	2892171	d4
WJL-d4-Plate1-25208-47	25208	WJL	Plate1	47	B	10	3387926	d4
WJL-d4-Plate1-25183-48	25183	WJL	Plate1	48	F	2	3595814	d4
WJL-d4-Plate1-25208-50	25208	WJL	Plate1	50	A	10	5708076	d4
WJL-d4-Plate1-25208-52	25208	WJL	Plate1	52	E	10	3305828	d4
WJL-d4-Plate1-25208-54	25208	WJL	Plate1	54	D	10	2980174	d4
WJL-d4-Plate1-25208-55	25208	WJL	Plate1	55	F	10	2648893	d4
GF-d4-Plate2-25208-57	25208	GF	Plate2	57	F	9	1789505	d4
GF-d4-Plate1-25183-59	25183	GF	Plate1	59	F	1	3461758	d4
GF-d4-Plate1-25210-60	25210	GF	Plate1	60	F	5	3205718	d4
WJL-d4-Plate2-25183-64	25183	WJL	Plate2	64	F	2	3165014	d4
GF-d4-Plate1-25208-67	25208	GF	Plate1	67	F	9	1551867	d4
WJL-d4-Plate2-25210-70	25210	WJL	Plate2	70	F	6	8073425	d4
GF-d4-Plate1-25208-72	25208	GF	Plate1	72	C	9	2668655	d4
GF-d4-Plate2-25210-74	25210	GF	Plate2	74	B	5	947737	d4
WJL-d4-Plate2-25210-75	25210	WJL	Plate2	75	C	6	4812520	d4
GF-d4-Plate2-25183-78	25183	GF	Plate2	78	B	1	2869820	d4
WJL-d4-Plate2-25183-79	25183	WJL	Plate2	79	C	2	4934533	d4
GF-d4-Plate1-25210-83	25210	GF	Plate1	83	C	5	4113175	d4
WJL-d4-Plate1-25210-84	25210	WJL	Plate1	84	B	6	4684552	d4
GF-d4-Plate2-25208-86	25208	GF	Plate2	86	B	9	3324070	d4
GF-d4-Plate1-25183-87	25183	GF	Plate1	87	C	1	3728767	d4
WJL-d4-Plate1-25183-88	25183	WJL	Plate1	88	B	2	4564509	d4
WJL-d4-Plate1-25210-90	25210	WJL	Plate1	90	C	6	3714293	d4
GF-d4-Plate1-25210-91	25210	GF	Plate1	91	B	5	4179985	d4
GF-d4-Plate2-25208-93	25208	GF	Plate2	93	C	9	3569201	d4
WJL-d4-Plate1-25183-94	25183	WJL	Plate1	94	C	2	4200621	d4
GF-d4-Plate1-25183-95	25183	GF	Plate1	95	B	1	4373035	d4
GF-d4-Plate1-25208-98	25208	GF	Plate1	98	B	9	3652231	d4
WJL-d4-Plate2-25210-101	25210	WJL	Plate2	101	B	6	4457721	d4
GF-d4-Plate2-25210-103	25210	GF	Plate2	103	C	5	3903565	d4
WJL-d4-Plate2-25183-104	25183	WJL	Plate2	104	B	2	982388	d4
GF-d4-Plate2-25183-105	25183	GF	Plate2	105	C	1	3094592	d4
GF-d4-Plate2-25208-110	25208	GF	Plate2	110	A	9	1967561	d4
WJL-d4-Plate1-25210-112	25210	WJL	Plate1	112	A	6	3472086	d4
WJL-d4-Plate1-25183-116	25183	WJL	Plate1	116	A	2	4865847	d4
GF-d4-Plate2-25210-119	25210	GF	Plate2	119	A	5	3773438	d4
WJL-d4-Plate2-25208-120	25208	WJL	Plate2	120	F	10	2018688	d4
WJL-d4-Plate2-25208-121	25208	WJL	Plate2	121	D	10	2595705	d4
WJL-d4-Plate2-25208-123	25208	WJL	Plate2	123	E	10	1841390	d4
WJL-d4-Plate2-25208-124	25208	WJL	Plate2	124	A	10	3326544	d4
GF-d4-Plate2-25183-125	25183	GF	Plate2	125	A	1	1822797	d4
WJL-d4-Plate2-25208-126	25208	WJL	Plate2	126	B	10	3831425	d4
WJL-d4-Plate2-25208-127	25208	WJL	Plate2	127	C	10	3109485	d4
WJL-d4-Plate2-25210-129	25210	WJL	Plate2	129	A	6	1737064	d4
GF-d4-Plate1-25208-132	25208	GF	Plate1	132	A	9	3284211	d4
WJL-d4-Plate2-25183-135	25183	WJL	Plate2	135	A	2	4603643	d4
GF-d4-Plate1-25210-139	25210	GF	Plate1	139	A	5	2749602	d4
GF-d4-Plate1-25183-140	25183	GF	Plate1	140	A	1	2722703	d4

Transparent Methods

•Fly stocks and genetic crosses

Drosophila were kept at 25°C in a Panasonic Mir425 incubator with 12/12 hrs dark/light cycles. Routine stocks were kept on standard laboratory diet (see below "media preparation and NAC treatment") The 53 DGRP lines were obtained from Bloomington Drosophila Stock Center.

Field-collected flies were trapped with rotten tomatoes in a garden in Solaize (France) and reared on a medium without chemical preservatives to minimize the modification to their gut microbiota(Tefit et al., 2017). One liter of media contains 15g inactivated yeast, 25g sucrose (Sigma Aldrich, ref. #84100), 80g cornmeal and 10g agar.

To generate DGRP F_2s, four DGRP lines were selected for setting up seven different crosses: 25210 (RAL-859), 25183(RAL-335) are the lines with "large" larvae as germ-free, and 25208(RAL-820) and 28147(RAL-158) are the line with "small" larvae as germ-free (see figure legend Figure S3a).

All RNAi lines were crossed to the driver line *y,w;; tubulin*-GAL80^{ts} ,*daugtherless*-GAL4. To minimize lethality, we dampend the GAL4 strength by leaving the genetic crosses at 25°C. The following fly strains were used: *y,w*, UAS-*dpr-6*-IR(P{KK112634}VIE-260B), UAS-*CG13492*-IR, (w^{1118};P{GD14825}v29390), UAS-*daw*-IR(NIG #16987R-1), UAS-*sfl*-IR (w^{1118}; P{GD2336}v5070), UAS-*arr*-IR (w^{1118}; P{GD2617}v4818), UAS-*rg*-IR(w^{1118}; P{GD8235}v17407), UAS-*bol*-IR(w^{1118}; {GD10525}v21536), UAS-*glut1*-IR(y^1 v^1; P{TRiP.JF03060}attP2, Bloomington 28645), UAS-*CG32683*-IR (P{KK112515}VIE-260B), UAS-*CG42669*-IR(w^{1118};P{GD7292}v18081), UAS-*Eip75B*-IR (w^{1118}; P{GD1434}v44851), UAS-*mCherry*-IR (y^1 v^1; P{CaryP}attP2), VDRC GD control (VDRC ID60000).

•GWAS and data computing of heritability indice

To calculate heritability, we estimated variance components using a random effects model using the lme4 R package(Bates, 2015). To infer the differences in heritability between GF and Lp^{WJL}-monoassociated conditions, we chose to use a bootstrap approach as in (https://github.com/famuvie/breedR/wiki/Heritability). Strains and experiment dates were treated as random effects, and the heritability was calculated as VA/(VA+VD+VR), where VA is the additive genetic variance, and is equal to twice the Strain variance, VD is the experiment date variance, and VR is the residual variance. For the estimation of the empirical distribution of heritability indices, a bootstrap method within the R breedR package was used for 1000 simulations per condition. We used the online tool specifically designed for the DGRPs (http://dgrp2.gnets.ncsu.edu/)(Huang et al., 2014; Mackay et al., 2012) for GWAS. The Manhattan and QQ-plots were generated using R. Raw GWAS data can be accessed at https://data.mendeley.com/datasets/5m9ghb7vbs/4

•Single larva transcriptome analysis

RNA extraction from single larvae: Larvae were handpicked under the microscope using forceps and transferred to Eppendorf tubes filled with 100uL of beads and 350 uL of Trizol. The samples were then homogenized using a Precellys 24 Tissue Homogenizer at 6000 rpm for 30 seconds. After homogenization, the samples were transferred to liquid nitrogen for flash freezing and stored at –80°C. For RNA extraction, samples were thawed on ice, 350 uL of 100% Ethanol was then added to each sample before homogenizing again with the same parameters. Direct-zol™ RNA Miniprep R2056 Kit was used to extract RNA with these modifications: DNAse I treatment was skipped; after the RNA Wash step, an extra 2 min centrifugation step was added to remove residual wash buffer. Lastly, the sample was eluted in 10 uL of water, incubated at room

temperature for 2 min and then spun for 2 min to collect RNA. RNA was transferred to a low-binding 96 well plate and stored at -70°C.

RNA-sequencing: We prepared the libraries using the BRB-seq protocol and sequenced them using an Illumina NextSeq 500 (Alpern et al., 2018). Reads from the BRB-seq protocol generates two fastq files: R1 containing barcodes and UMIs and R2 containing the read sequences. R2 fastq file was first trimmed for removing BRB-seq-specific adapter and polyA sequences using the BRB-seqTools v1.0 suite (available at http://github.com/DeplanckeLab/BRB-seqTools). We then aligned the trimmed reads to the Ensembl r78 gene annotation of the dm3 genome mixed with *the Lactobacillus Plantarum WJL* genome using STAR (Version 2.5.3a)(Dobin et al., 2013), with default parameters (and extra "--outFilterMultimapNmax 1" parameter for completely removing multiple mapped reads). Then, using the BRB-seqTools v1.0 suite (available at http://github.com/DeplanckeLab/BRB-seqTools), we performed simultaneously the sample demultiplexing, and the count of reads per gene from the R1 FASTQ and the aligned R2 BAM files. This generated the count matrix that was used for further analyses. Genes were retained in the analysis if they had more than 10 reads in more than 50 samples. The data was subsequently transformed using the voom method. Differential expression was performed using the R Limma package(Law et al., 2014; Ritchie et al., 2015). Genes with a $\log_2$ fold change greater than 2 and a Benjamini-Hochberg adjusted P-value less than 0.05 were considered differentially expressed. Since the library preparation was performed in two plates, hence introducing a batch effect, we used the duplicateCorrelation function and included the batch as a blocking variable. Prior to PCA analysis and standard deviation calculations, we removed the batch effect using the removeBatchEffects function and then used the princomp function. We used the cluster profiler package to perform GSEA analyses. The gmt file containing the gene ontology annotations was obtained from GO2MSIG data. Specifically, we used the highquality GO annotations for *Drosophila melanogaster*. For each GSEA analysis, we used 100,000 permutations to obtain adjusted p-values and only included gene set sizes to between 6 and 1000 genes. The raw expression data has been deposited in ArrayExpress (accession number: E-MTAB-6518)

RNA-sequencing: We prepared the libraries using the BRB-seq protocol and sequenced them using an Illumina NextSeq 500(Alpern et al., 2018). Reads from the BRB-seq protocol generates two fastq files: R1 containing barcodes and UMIs and R2 containing the read sequences. R2 fastq file was first trimmed for removing BRB-seq-specific adapter and polyA sequences using the BRB-seqTools v1.0 suite (available at http://github.com/DeplanckeLab/BRB-seqTools). We then aligned the trimmed reads to the Ensembl r78 gene annotation of the dm3 genome mixed with *the Lactobacillus Plantarum WJL* genome using STAR (Version 2.5.3a)(Dobin et al., 2013), with default parameters (and extra "--outFilterMultimapNmax 1" parameter for completely removing multiple mapped reads). Then, using the BRB-seqTools v1.0 suite (available at http://github.com/DeplanckeLab/BRB-seqTools), we performed simultaneously the sample demultiplexing, and the count of reads per gene from the R1 FASTQ and the aligned R2 BAM files. This generated the count matrix that was used for further analyses. The data was subsequently transformed using the voom method and analyzed using the R Limma package(Law et al., 2014; Ritchie et al., 2015).

The raw expression data of BRB-Seq has been deposited in ArrayExpress (accession number: E-MTAB-6518)

•The making and maintenance of germ-free flies

Axenic flies were generated by dechorionating embryos with 50% household bleach for five minutes; eggs were then washed in successive 70% ethanol and sterile distilled water for three minutes each. After washing, eggs were transferred to tubes containing standard diet and a

cocktail of antibiotics containing 50µg/mL ampicillin, 50µg/mL kanamycin, 15µg/mL erythromycin, 50µg/mL tetracyclin for stock maintenance. Axeny was routinely verified by plating larvae and adult lysates on LB and MRS plates. For experiments food without antibiotics was used.

•Media preparation and NAC treatment

Standard laboratory fly food consists of 50g/L inactivated yeast (Springaline™), 80g/L cornmeal, 7.14g/L agar, 5.12g/L Moldex (Sigma M-50109) and 0.4% propionic acid. Where applicable, experiments comparing variations in larval size, developmental timing, adult emergence were performed on diet with 6g or 8g inactivated yeast per liter of media while keeping the same concentrations for the other ingredients. Where appropriate, 1.7g/L of N-Acetylcystein (SigmaA7250-25g) was added to the low-protein diet.

•Larval Length Measurement

All live *Drosophila* larvae were collected from each nutritive cap containing low yeast diet by temporary immersion in sterile PBS, transferred on a microscopy slide, killed with a short pulse of heat (5 sec at 90°C), mounted with 80% glycerol/PBS. The images were taken with the Leica stereomicroscope M205FA and the lengths of individual larvae were measured using ImageJ software(Schneider et al., 2012). For each DGRP strain and each cross and/or condition, at least three biological replicates were generated.

•Developmental timing and Adult emergence

Developmental timing and adult emergence of the flies were quantified by counting the number of individuals appearing every 24 hours until the last pupa/adult emerges. Each animal is assigned to the number that corresponds to the day it appeared, and the population mean and variance were calculated based on the cumulative numbers.

•Adult trait measurements

2-3 days old adult flies were anesthetized with CO_2 and immersed in 70% ethanol, and individual body and its corresponding organ (wing and eye) were imaged under a Leica M205 stereomicroscope. Specifically, the adult body length was measured from the top of the head to the tip of the abdomen. The eye area was measured by manually tracing the circumference of both eyes. The wings were gently nipped at the base of the hinge and imaged, and the area was measured by tracing the edge of the wing. All images were taken measured using ImageJ software

•Bacteria culture and mono-association

For each mono-association experiment, Lp^{WJL} (Ryu et al., 2008) was grown in Man, Rogosa and Sharpe (MRS) medium (Difco, ref. #288110) over-night at 37°C, and diluted to O.D.=0.5 the next morning to inoculate 40 freshly laid eggs on a 55mm petri dish or standard 28mm tubes containing fly food of low yeast content. The inoculum corresponds to about $5x10^7$ CFUs. Equal volume of sterile PBS was spread on control axenic eggs.

To contaminate the garden-collected flies with their own microbiota, eggs were dechorionated and directly seeded onto appropriate food caps. Sterile PBS was used to wash the side of the bottles where the adult wild flies were raised to recover more fecal content, and 300 ul of the wash was inoculated to the dechorionated eggs. For GF control, 300 ul of sterile PBS was used to inoculate the dechorionated eggs. The microbial composition of this community can be founded here(Tefit et al., 2017).

•Bacteria niche load

Five to six 24 hour old germ-free larvae were collected from the low-protein diet food cap and transferred to a microtube containing 400ul of low-protein diet, and inoculated with 50ul of Lp^{WJL} of 0.5 O.D.. On the day of harvest, ~0.75-1mm glass micro-beads and 900µl PBS were added to each microtube and the entire content of the tube was homogenized with the Precellys-24 tissue homogenizer (Bertin Technologies). Lysate dilutions (in PBS) are plated on MRS agar with Easyspiral automatic plater (Intersciences). The MRS agar plates were incubated for 24h at 37°C. The CFU/ml count was calculated based on the readings by the automatic colony counter Scan1200 (Intersciences)

•Statistical Analysis and data representation

GraphPad Prism software version 6.0f for Macintosh (GraphPad Software, La Jolla California USA, www.graphpad.com) was used to compare GF and Lp^{WJL}-associated conditions for larval length, developmental timing, adult emergence, allometry and linear regression analysis for the buffering effect. For small samples with less than 10 data points, nonparametric analysis was conducted. For all each sample set, we first conducted D'agostino-Pearson normality test. If the samples assume normal distribution, the F test of equality of variances were conducted to compare variability among the datasets. For samples assuming non-normal distribution, Levene's test is conducted based on the deviation from the median of each dataset.

Supplemental Reference

Alpern, D., Gardeux, V., Russeil, J., and Deplancke, B. (2018). Time- and cost-efficient high-throughput transcriptomics enabled by Bulk RNA Barcoding and sequencing. bioRxiv.

Bates, D., Mâchler M, Bolker B, Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software *67*, 1-48.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Hu, Y., Comjean, A., Perkins, L.A., Perrimon, N., and Mohr, S.E. (2015). GLAD: an Online Database of Gene List Annotation for Drosophila. J Genomics *3*, 75-81.

Huang, W., Massouras, A., Inoue, Y., Peiffer, J., Ramia, M., Tarone, A.M., Turlapati, L., Zichner, T., Zhu, D., Lyman, R.F., *et al.* (2014). Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. Genome Res *24*, 1193-1208.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol *15*, R29.

Mackay, T.F., Richards, S., Stone, E.A., Barbadilla, A., Ayroles, J.F., Zhu, D., Casillas, S., Han, Y., Magwire, M.M., Cridland, J.M., *et al.* (2012). The Drosophila melanogaster Genetic Reference Panel. Nature *482*, 173-178.

McFall-Ngai, M., Hadfield, M.G., Bosch, T.C., Carey, H.V., Domazet-Loso, T., Douglas, A.E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F., *et al.* (2013). Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci U S A *110*, 3229-3236.

Powell, J.A. (2014). GO2MSIG, an automated GO based multi-species gene set generator for gene set enrichment analysis. BMC Bioinformatics *15*, 146.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47.

Ryu, J.H., Kim, S.H., Lee, H.Y., Bai, J.Y., Nam, Y.D., Bae, J.W., Lee, D.G., Shin, S.C., Ha, E.M., and Lee, W.J. (2008). Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in Drosophila. Science *319*, 777-782.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675.

Tefit, M.A., Gillet, B., Joncour, P., Hughes, S., and Leulier, F. (2017). Stable association of a Drosophila-derived microbiota with its animal partner and the nutritional environment throughout a fly population's life cycle. J Insect Physiol.