Studies reporting relationships between epistasis and background fitness

Table S1 Summary of the 30 reviewed studies reporting an empirical relationship between epistasis and background fitness. These were identified by means of a Google Scholar search for "*diminishing returns epistasis*". We chose for further review the first 30 papers, found either through this search or because they were cited in one of these papers, reporting a relationships between epistasis and expected fitness. Information is given on the type of organism and experimental setting that was used, whether mutations had beneficial or detrimental effects, the locations of the mutations, and under which circumstances selection on mutations was measured. Approximate F- and P-values are given for the epistatic variance components, on which the published correlations were based. The type of test reported in the original publication is given, and if the study corrected for regression-to-the-mean. In a few studies neither point estimates nor significance tests were provided, but instead conclusions in these studies were based on graphical exploration of the relationship between expected fitness and epistasis, and hence they were included here. For all studies for which error variances could be extracted, the published correlation is given along with an estimate of the corrected correlation using Eq. 5 in the main text.

Study	Organism	Setting	Effects	Location	Selection	Appr.F	Appr.P	Published r	Corrected r	Test	Correction
Caudle et al. 2014	ssDNA Bacteriophage ID11	9 isolated mutations tested in novel environment (41°C)	beneficial	different	general	$F_{17,72} = 49.4$	< 0.001	-0.72	-0.72	$r_{E_{ij},[w_i+w_j]}$	n.a
		9 isolated mutations tested in ancestral environment (37°C)	beneficial	different	general	$F_{17,72} = 8.4$	< 0.001	0.12	0.24	$r_{E_{ij},[w_i+w_j]}$	n.a
		9 isolated mutations tested in novel environment (33°C)	both	different	general	$F_{17,72} = 4.6$	< 0.001	-0.54	-0.53	$r_{E_{ij},[w_i+w_j]}$	n.a
Perfeito et al. 2014	*Escherichia coli*	Fitness increase following mutation accumulation in 23 genetic backgrounds	beneficial	different	general	$F_{22,46} = 6.8$	< 0.001	-0.76	-0.78	$r_{w_i,[w_{ij}-w_i]}$	no
		Fitness decline in 23 genotypes during mutation accumulation	deleterious	different	general	$F_{22,46} = 5.9$	< 0.001	-0.34	-0.31	$r_{w_i,[w_{ij}-w_i]}$	no
Elena & Lenski 1997	*Escherichia coli*	Single, double and triple mutants from 225 genotypes generated by mutagenesis	deleterious	different	general						n.a
Trindade et al. 2009	*Escherichia coli*	19 mutations conferring antibiotic resistance, tested in absence of drugs	deleterious	different	general	$F_{102,309} = 14.6$	< 0.001	-0.19	-0.16	$r_{E_{ij},[w_i+w_j]}$	no
Trindade et al. 2012	*Escherichia coli*	25 mutations conferring antibiotic resistance, tested in absence of drugs	deleterious	different	general						n.a
Bonhoeffer et al.	HIV-1	Natural variation in 9466	deleterious	different	general					selection	no

2004		isolated natural strains								
Burch & Chao 2004	RNA bacteriophage Φ6	Fitness decline during mutation accumulation	deleterious	different	general				$r_{w_i,[w_{ij}-w_i]}$	no
Kryazhimskiy et al. 2014	*Saccharomyces cerevisiae*	Fitness increase during experimental evolution, and fitness decline of gene knockouts, in 64 genetic backgrounds	beneficial	different	general				$r_{w_i,[w_{ij}-w_i]}$	yes
Kvitek & Sherlock 2011	*Saccharomyces cerevisiae*	15 mutations isolated during experimental evolution, tested in ancestral environment	beneficial	different	general					n.a
Szafraniec et al. 2003	*Saccharomyces cerevisiae*	74 genotypes generated by mutagenesis	deleterious	different	general				$b_{w_{ij},[w_i+w_j]} < 1$	yes
Jasnos et al 2008	*Saccharomyces cerevisiae*	Multiple single and double deletions generated through genetic engineering, tested in 5 environments	deleterious	different	general					n.a
Xu et al. 2012	*Saccharomyces cerevisiae*	Comparative meta-data	deleterious	different	general				selection	no
Pearson et al. 2012	ssDNA Bacteriophage ID11	Assayed isolated mutation F on 8 fitness backgrounds	beneficial	different	general	$F_{7,32} = 10.0$	< 0.001	-0.71	-0.73	r_{E_{ij},w_i}	no
		Assayed isolated mutation B on 8 fitness backgrounds	beneficial	different	general	$F_{7,32} = 25.1$	< 0.001	-0.53	-0.53	r_{E_{ij},w_i}	no
Rokyta et al. 2011	ssDNA Bacteriophage ID11	9 mutations isolated from directed mutagenesis	beneficial	different	general	$F_{17,72} = 8.7$	< 0.001	0.12	0.22		n.a
Lalic & Elena 2012	Tobacco potyvirus	20 single mutants combined to create 53 double mutants	deleterious	different	general	$F_{42,344} = 2.9$	< 0.001	-0.42	-0.23	$r_{E_{ij},[w_i+w_j]/2}$	no
Sanjuan et al. 2005	Vesicular stomatitis virus	Fitness increase in 12 genetic backgrounds following mutation accumulation	beneficial	different	general	$F_{11,48} = 8.7$	< 0.001	-0.49	-0.42	$r_{w_i,[w_{ij}/w_i]}$	no
Martin et al. 2007	Vesicular stomatitis virus	Multiple mutations generated by mutagenesis, tested in ancestral environment	both	different	general				selection	no
Sanjuan et al. 2004	Vesicular stomatitis virus	47 genotypes generated by mutagenesis	both	different	general				selection	no
Hall & MacLean 2011	*Pseudomonas aeruginosa*	14 mutations in antibiotic resistance gene, tested in absence of antibiotics	deleterious	same gene	general	$F_{13,98} = 3.0$	0.001	-0.51	-0.45	$r_{E_{ij},[w_iw_j]}$	no
		14 mutations in antibiotic resistance gene, tested in absence of antibiotics	deleterious	same gene	general			-0.55		$r_{E_{ij},[w_iw_j]}$	no
		14 mutations in antibiotic resistance gene, tested in	deleterious	same gene	general			-0.45		$r_{E_{ij},[w_iw_j]}$	no

		absence of antibiotics									
Hayden & Wagner 2012	Azoarcus group 1 ribozyme	4 isolated mutations tested in 3 environments	beneficial	different	specific						n.a
Flynn et al. 2013	*Escherichia coli*	5 mutations isolated during experimental evolution, tested in novel environment	beneficial	different	specific	$F_{31,128} = 4.6$	< 0.001	-0.75	-0.74	$r_{E_{ij},[w_i+w_j]}$	no
		5 mutations isolated during experimental evolution, tested in novel environment	beneficial	different	specific	$F_{17,72} = 17.2$	< 0.001	-0.28	-0.25	$r_{E_{ij},[w_i+w_j]}$	no
Khan et al. 2011	*Escherichia coli*	5 mutations isolated during experimental evolution, tested in ancestral environment	beneficial	different	specific	$F_{25,364} = 4.6$	< 0.001	-0.58	-0.56	$r_{E_{ij},[w_i+w_j]}$	no
		rbs-mutation	beneficial	different	specific	$F_{15,224} = 10.5$	< 0.001	-0.26	-0.24	$r_{w_i,[w_{ij}-w_i]}$	no
		topi-mutation	beneficial	different	specific	$F_{15,224} = 13.9$	< 0.001	-0.59	-0.59	$r_{w_i,[w_{ij}-w_i]}$	no
		spoT-mutation	beneficial	different	specific	$F_{15,224} = 16.6$	< 0.001	-0.50	-0.50	$r_{w_i,[w_{ij}-w_i]}$	no
		glmUS-mutation	beneficial	different	specific	$F_{15,224} = 13.0$	< 0.001	-0.50	-0.50	$r_{w_i,[w_{ij}-w_i]}$	no
		pykF-mutation	beneficial	different	specific	$F_{15,224} = 5.2$	< 0.001	0.65	0.78	$r_{w_i,[w_{ij}-w_i]}$	no
Wang et al. 2012	*Escherichia coli*	2 mutations isolated during experimental evolution, tested in ancestral environment on 8 different genetic backgrounds	beneficial	different	specific			-0.81		r_{E_{ij},w_k}	yes
Bull et al 2000	Bacteriophage ΦX174	Benefit of new mutation on 4 genetic backgrounds evolved under heat stress	beneficial	same pathway	specific					$r_{w_i,[w_{ij}-w_i]}$	no
Schenk et al. 2013	*Escherichia coli*	4 antibiotic resistance mutations	beneficial	same pathway	specific	$F_{10,33} > 1000$	< 0.001	-0.85	-0.85	$r_{E_{ij},[w_i+w_j]}$	no
		4 antibiotic resistance mutations	beneficial	same pathway	specific	$F_{10,44} > 1000$	< 0.001	-0.75	-0.75	$r_{E_{ij},[w_i+w_j]}$	no
da Silva et al. 2010	HIV-1	7 mutations generated by directed mutagenesis, tested in ancestral environment	both	same pathway	specific						n.a
Chou et al. 2011	*Methylobacterium extorquens*	4 mutations isolated during experimental evolution, tested in ancestral environment	beneficial	same pathway	specific	$F_{31,64} = 15.8$	< 0.001	-0.8	-0.82	$r_{w_i,[w_{ij}/w_i]}$	no
		fghA-mutation	beneficial	same pathway	specific	$F_{7,16} = 2.2$	0.09	-0.89	undefined	$r_{w_i,[w_{ij}/w_i]}$	no
		pntAB-mutation	beneficial	same pathway	specific	$F_{7,16} = 0.58$	0.76	-0.53	undefined	$r_{w_i,[w_{ij}/w_i]}$	no
		gshA-mutation	beneficial	same pathway	specific	$F_{7,16} = 5.0$	0.004	-0.94	-0.97	$r_{w_i,[w_{ij}/w_i]}$	no

		GB-mutation	beneficial	same pathway	specific	$F_{7,16} = 2.2$	0.09	-0.84	undefined	$r_{w_i,[w_{ij}/w_i]}$	no
Chou et al. 2014	*Methylobacterium extorquens*	4 mutations isolated during experimental evolution, tested in ancestral environment	beneficial	same pathway	specific	$F_{27,56} = 289.7$	< 0.001	-0.53	-0.53	$r_{E_{ij},[\,w_i w_j]}$	yes
MacLean 2009	*Pseudomonas aeruginosa*	Experimental activation of two antibiotic resistance pathways	both	same pathway	specific	$F_{7,21} = 1.9$	0.12	-0.73	undefined	$r_{E_{ij},[\,w_i w_j]}$	no
MacLean et al. 2010	*Pseudomonas aeruginosa*	Antibiotic resistance increase in three different genetic backgrounds	beneficial	same pathway	specific			-1		$r_{w_i,[w_{ij}/w_i]}$	no

Setting: See original publications for more detail.

Location: *same gene* refers to the authors reporting that most studied mutations were located physically in the same gene, *same pathway* refers to mutations affecting the same physiological pathway (as often was the case for beneficial mutations isolated during experimental evolution on a specific growth medium), and *different* refers to mutations that were random and often their effects and locations were not known a priori.

Selection: *direct* corresponds to when fitness of mutations was assessed under the same specific selective conditions as they first were identified in (as for fitness of mutations conferring antibiotic resistance tested on a growth medium containing the antibiotic), and *general* corresponds to when fitness of mutations were scored in less specific conditions (as for fitness of antibiotic resistance mutations on a growth medium not containing the antibiotic, or for random mutations acquired by mutagenesis or through mutation accumulation experiments).

Appr.F: an approximation of the F-ratio for the epistatic interaction variance [i.e. $\sigma^2(E_{ij}) / \sigma^2(e_i+e_j+e_{ij})$ for the case when absolute epistasis had been estimated]. The first term in the subscript of F gives the degrees of freedom for the effect of mutant genotype ($n_{genotypes}$-1), and the second term gives the degrees of freedom for the error term ($n_{tot} – n_{genotypes}$).

Appr.P: the accompanying approximation of the P-value for the epistatic interaction variance. Note that no P-values were calculated for the corrected correlations as it would require simulation and resampling using the original datasets (see "Correcting the correlations for measurement error" below for further details).

Test: $r_{E_{ij},[w_i+w_j]}$, $r_{E_{ij},[w_i+w_j]/2}$ and $r_{E_{ij},[\,w_i w_j]}$ refer to a correlation between epistasis and expected fitness, r_{E_{ij},w_i} refers to a correlation between epistasis and background fitness, $r_{w_i,[w_{ij}/w_i]}$ and $r_{w_i,[w_{ij}-w_i]}$ refer to a correlation between background fitness and fitness improvement or decline, $b_{w_{ij},[w_i+w_j]} < 1$ refers to a test of a major axis (MA) regression slope of observed fitness on expected fitness of double mutants being significantly

below 1, *selection* refers to subsampling of mutations based on their fitness effects and subsequent comparisons of the strength of epistasis between mutations with above and below average fitness effects.

Correction: whether the statistical method itself, or any additional measure was taken to reduce or correct for the effect of regression-to-the-mean. *n.a* refers to cases where relationships with epistasis were not directly tested by regressions using background or expected fitness. Many of these studies instead compared epistasis across high and low quality environments, or between high and low order gene interaction.

Correcting the correlations for measurement error

As outlined in the main text, the correlation between [$w_i + w_j$] and E_{ij} is a function of not only $\sigma^2(a_i + a_j)$, $\sigma^2(i_{ij})$ and $\sigma(i_{ij}, [a_i+a_j])$, but also of $\sigma^2(e_i + e_j)$ and $\sigma^2(e_{ij})$ (see eqn. 4). Having knowledge of $\sigma^2(e_i + e_j)$ and $\sigma^2(e_{ij})$, we are able to obtain the corrected correlation between i_{ij} and [$a_i + a_j$] that is not biased by measurement error variance, using eqn. 5 (see main text for more details). In the main text we derive eq. 5 for the case when the expected fitness of double mutants assuming purely additive effects ($w_i + w_j$) is correlated with the absolute amount of epistasis ($E = w_{ij} - [w_i + w_j]$). However, some studies used a relative measure of epistasis (i.e. $w_{ij} / [w_i * w_j]-1$), or they used the absolute (i.e. $w_{ij} - w_i$) or relative (i.e. w_{ij} / w_i-1) fitness improvement associated with introducing mutation *j* into a genetic background containing mutation *i* (see table S1 for further details). To accommodate this, eq. 5 was modified appropriately.

In cases where the true variance between mutants is low (small $\sigma^2(a_i)$, $\sigma^2(a_j)$ and $\sigma^2(a_{ij})$ relative to measurement error ($\sigma^2(e_i)$, $\sigma^2(e_j)$ and $\sigma^2(e_{ij})$), it becomes clear that correcting the variance components in the denominator inflates the corrected correlation to take on extreme values. Hence, correlations based on non-significant variance components will be erroneous, and confidence limits and significance of correlations needs to be estimated using data resampling techniques. Therefore we calculated the F-ratio and accompanying P-value for the epistatic variance component using $\sigma^2(E_{ij}) / \sigma^2(e_i+e_j+e_{ij})$ as an indicator of the reliability of the published estimate of the correlation. Indeed, in four cases the epistatic variance was approximated to be non-significant (Table S1), and correction lead to the correlation taking on values outside the theoretical boundary ($r < -1$) due to the corrected epistatic variance (i_{ij}), present in the denominator of Eq. 5, approaching zero.

Approximate measurement error variances were derived from the mean squared standard errors of mean fitness for both single and double mutants. In cases where we could not find separate estimates of error variance for both single and double mutants, equal error variances were assumed. Estimates of errors in expected fitness ($\sigma^2(e_i + e_j)$) and epistasis ($\sigma^2(e_i+e_j+e_{ij})$) were obtained through error propagation of single and double mutant estimation errors.

It is clear that these corrected correlations and F-values only serve as approximations. Indeed, in all our corrections we estimated measurement error variance from the mean standard error across all measured genotypes for one class of mutant (single or double). Thereby we assumed that measurement error was the same for all genotypes of one class of mutant. In addition, in Khan et al. (2011), Chou et al. (2011;2014), Flynn et al. (2013) and Schenk et al. (2013), complex higher-order epistatic interactions between 2 to 5 mutations were studied, requiring more elaborate statistical corrections using resampling methods to arrive at exact estimates of the correlations and their statistical significance. Nevertheless, our analysis shows that for the majority of studies measurement error variance is relatively small, and as a consequence, correction has little effect on the qualitative conclusions drawn from the combined body of literature reviewed here concerning diminishing returns epistasis (Table S1).

References

Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM, Petropoulos CJ. 2004. Evidence for positive epistasis in HIV-1. Science 306: 1547-1550.

Bull JJ, Badgett MR, Wichman HA. 2000. Big-benefit mutations in a bacteriophage inhibited with heat. Molecular Biology and Evolution 17: 942-950.

Burch CL, Chao L. 2004. Epistasis and its relationship to canalizations in the RNA virus Φ6. Genetics 167: 559-567.

Caudle SB, Miller CR, Rokyta DR. 2014. Environment determines epistatic patterns for a ssDNA virus. Genetics 196: 267-279.

Chou H, Chiu H, Delaney NF, Segre D, Marx CJ. 2011. Diminishing returns epistasis among beneficial mutations decelerates adaptation. Science 332: 1190-1192.

Chou H, Delaney NF, Draghi JA, Marx CJ. 2014. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and epistasis between adaptive mutations. PLoS Genetics 10: e1004149.

da Silva J, Coetzer M, Nedellec R, Pastore C, Mosier DE. 2010. Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. Genetics 185: 292-303.

Elena SF, Lenski RE. 1997. Test of synergistic interactions among deleterious mutations in bacteria. Nature 395-398.

Flynn KM, Cooper TF, Moore FGB, Cooper VS. 2013. The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. PLoS Genetics 9: e1003426.

Hall AR, MacLean RC. 2011. Epistasis buffers the fitness effects of Rifampicin-resitance mutations in Pseudomonas aeruginosa. Evolution 65: 2370-2379.

Jasnos L, Tomala K, Paczesniak D, Korona R. 2008. Interactions between stressful environment and gene deletions alleviate the expected average loss of fitness in yeast. Genetics 178: 2105-2111.

Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. Science 332: 1193-1196.

Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. 2014. Global epistasis makes adaptation predictable despite sequence-level stochasticity. bioRxiv, doi: 10.1101/001784.

Kvitek DJ, Sherlock G. 2011. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. PLoS Genetics 7: e1002056.

Lalic J, Elena SF. 2012. Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. Heredity 109: 71-77.

MacLean RC. 2009. Predicting epistasis: an experimental test of metabolic control theory with bacterial transciption and translation. Journal of Evolutionary Biology 23: 488-493.

MacLean RC, Perron GG, Gardner A. 2010. Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for Rifampicin resistance in Pseudomonas aeruginosa. Genetics 186: 1345-1354.

Martin G, Elena SF, Lenormand T. 2007. Distributions of epistasis in microbes fit predictions from a fitness landscape model. Nature Genetics 39: 555-560.

Martinez JP, Bocharov G, Ignatovich A, Reiter J, Dittmar MT, Wain-Hobson S, Meyerhans A. 2011. Fitness ranking of individual mutants drives patterns of epistatic interactions in HIV-1. PLoS ONE 6: e18375.

Pearson VM, Miller CR, Rokyta DR. 2012. The constancy of beneficial fitness effects of mutations across diverse genetic backgrounds. PLOS ONE 7: e43864.

Perfeito L, Sousa A, Bataillon T, Gordo I. 2013. Rates of fitness decline and rebound suggest pervasive epistasis. Evolution 68: 150-162.

Rokyta DR, Joyce P, Caudle SB, Miller C, Beisel CJ, Wichman HA. 2011. Epistasis between beneficial mutations and the phenotype-to-fitness map for a ssDNA virus. PLoS Genetics 7: e1002075.

Sanjuan R, Cuevas JM, Moya A, Elena SF. 2005. Epistasis and the adaptability of an RNA virus. Genetics 170: 1001-1008.

Sanjuan R, Moya A, Elena SF. 2004. The contribution of epistasis to the architecture of fitness in an RNA virus. Proceedings of the National Academy of Sciences USA 101: 15376-15379.

Schenk MF, Szendro IG, Salverda MLM, Krug J, de Visser JAGM. 2013. Patterns o fepistasis between beneficial mutations in an antibiotic resistance gene. Molecular Biology and Evolution 30: 1779-1787.

Szafraniec K, Wloch DM, Sliwa P, Borts RH, Korona R. 2003. Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast Saccharomyces cerevisiae. Genetical Reseach of Cambridge 82: 19-31.

Trindade S, Sousa A, Xavier KB, Dionisio F, Ferreira MG, Gordo I. 2009. Positive epistasis drives the acquisition of multidrug resistance. PLoS Genetics 5: e 1000578.

Trindade S, Sousa A, Gordo I. 2012. Antibiotic resitance and stress in the light of Fisher's model. Evolution 66: 3815-3824.

Wang Y, Arenas CD, Stoebel DM, Cooper TF. 2012. Genetic background affects epistatic interactions between two beneficial mutations. Biology Letters 9: 20120328.

Xu L, Barker B, Gu Z. 2012. Dynamic epistasis for different alleles of the same gene. Proceedings of the National Academy of Sciences USA 109: 10420-10425.

Table S2 Numerical example of (post-hoc) correction for measurement error variance

Available for download as an Excel file at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.169870/-/DC1