

Title	Parasite genetic distance and local adaptation in co-evolving bacteria-bacteriophage populations
Authors	Scanlan, Pauline D.;Hall, Alex R.;Buckling, Angus
Publication date	2016-10-24
Original Citation	Scanlan, P. D., Hall, A. R. and Buckling, A. (2017) 'Parasite genetic distance and local adaptation in co-evolving bacteria-bacteriophage populations', <i>Molecular Ecology</i> , 26(7), pp. 1747-1755. doi: 10.1111/mec.13897
Type of publication	Article (peer-reviewed)
Link to publisher's version	<a href="https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13897">https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13897</a> - 10.1111/mec.13897
Rights	© 2016 John Wiley & Sons Ltd. This is the peer reviewed version of the following article: Scanlan, P. D., Hall, A. R. and Buckling, A. (2017), Parasite genetic distance and local adaptation in co# evolving bacteria-bacteriophage populations. <i>Mol Ecol</i> , 26: 1747-1755, which has been published in final form at <a href="https://doi.org/10.1111/mec.13897">https://doi.org/10.1111/mec.13897</a> . This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.
Download date	2024-05-04 15:43:29
Item downloaded from	<a href="https://hdl.handle.net/10468/6449">https://hdl.handle.net/10468/6449</a>

Received Date : 29-Jun-2016

Revised Date : 16-Sep-2016

Accepted Date : 12-Oct-2016

Article type : Special Issue

**Parasite genetic distance and local adaptation in coevolving bacteria-bacteriophage populations**

**Authors:** Pauline D. Scanlan<sup>1\*</sup> Alex R. Hall<sup>2\*</sup>, Angus Buckling<sup>3</sup>

**Affiliations:**

<sup>1</sup> APC Microbiome Institute, Bioscience Building, University College Cork, Ireland.

<sup>2</sup>Institute of Integrative Biology, ETH Zürich, Switzerland.

<sup>3</sup>ESI & CEC, Biosciences, University of Exeter, UK

\*These authors contributed equally to the manuscript.

**Correspondence:** p.scanlan@ucc.ie

**Running head:** Parasite genetic distance and local adaptation

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.13897

This article is protected by copyright. All rights reserved.

## Summary

Antagonistic coevolution between hosts and parasites can lead to local adaptation (LA), such that parasite fitness is greatest in sympatric hosts (or *vice versa*). The magnitude of LA typically increases with geographic distance, which is assumed to be because genetic (and hence phenotypic) distance increases with geographic distance. Here we explicitly test the relationships between parasite genetic and phenotypic distance and LA using isolates of coevolved viral parasites (lytic bacteriophage  $\phi 2$ ) and the host bacterium *Pseudomonas fluorescens* SBW25. We find positive relationships between parasite genotype and infectivity phenotype, but the strength of the relationship was greater when infectivity was defined by the identity of hosts that could be infected rather than the actual number of hosts infected (host range), and when measurements were compared within rather than among populations. Crucially, we find a monotonic relationship between LA and genetic distance across phage isolates from different populations, although in contrast to many geographic studies, parasite LA decreased with genetic distance. These results can be explained by the fact that bacteria can rapidly adapt to phage infectivity mutations, but that evolved resistance has a degree of specificity to the local phage population. Our results show that antagonistic coevolution alone can result in predictable links between genetic distance and host-parasite local adaptation.

## Introduction

Host-parasite antagonistic coevolution, the reciprocal evolution of defence and counter-defence, can result in local adaptation (LA) of either parasites or hosts, such that fitness is greater in the presence of sympatric versus allopatric host or parasite populations, respectively (Kawecki & Ebert 2004). Such LA has important implications for both the maintenance of diversity and epidemiology. The magnitude and sign of LA can be driven by a complex interplay of genetic and ecological variables (Gandon *et al.* 1996; Gandon & Nuismer 2009; Gomez *et al.* 2015; Greischar & Koskella 2007; Laine 2008; Morand *et al.* 1996) but a common predictor of LA in natural populations is the geographical distance between populations: LA typically, although not always (McCoy *et al.* 2002), increases with distance (*e.g.* (Ebert 1994; Imhoof & Schmid-Hempel 1998; Koskella *et al.* 2011). The general explanation for this pattern is that genetic distance, and hence dissimilarity of infectivity/ resistance phenotypes, increases with geographic distance (Kaltz & Shykoff 1998).

While it is commonly observed that parasite infectivity decreases as a function of genetic distance from the primary host (*e.g.* (Antonovics *et al.* 2013; Longdon *et al.* 2011; Perlman & Jaenike 2003), evidence for this relationship being driven by coevolution is limited (Antonovics *et al.* 2013). Specifically, the above studies investigate genetic distance at interspecific scales, and the inability of parasites to infect genetically distinct hosts probably represents an ancestral state rather than derived state of hosts and parasites (Antonovics *et al.* 2013) - so called “non-host resistance” (Heath 1981). In one recent study where infectivity and resistance traits are likely to be driven by coevolution (Lange *et al.* 2015), the infectivity of a microsporidian parasite population to populations of *Daphnia magna* was minimised when hosts were at intermediate genetic distance from the actual host populations (Lange *et al.* 2015), in contrast to the expected linear relationship. A possible explanation for this may be that the importance of genetic distance was masked by strong effects of other ecological selection pressures that correlate with susceptibility, or that genetic distance was determined from a mitochondrial gene not linked to resistance. Direct tests of the hypothesis that coevolution can drive monotonic relationships between genetic distances and LA are therefore lacking, and we carry out such a test here using coevolving populations of bacteria and viruses (bacteriophages; phages).

Critical to the importance of genetic distance increasing the magnitude of LA is that genetic distance correlates with phenotypic distance. If, for example, the same infectivity or resistance phenotype can be encoded by multiple alternative sequences, the impact of genetic distance on LA will be much weaker than if there is a perfect correlation between genotype and phenotype. While there is unsurprisingly clear evidence for a tight link between genotype and phenotype in some well characterised host-parasite systems (*e.g.* (Bull & Molineux 2008; Perry *et al.* 2015), this is not always the case (Scanlan *et al.* 2015). Moreover, the correlation between genotype and phenotype may differ across spatial scales: alternative genotypes encoding the same phenotype may be less likely to co-exist within than between populations as a result of diversity being lost through drift or selection.

We investigated the correlations between parasite genetic and phenotypic distance and LA in replicate populations of a lytic bacteriophage (SBW25  $\phi$ 2) that had been coevolved with the soil bacterium *Pseudomonas fluorescens* SBW25 for approximately 400 generations as part of a previous study (Hall *et al.* 2011). These phage and bacterial populations were evolved under identical physical and chemical conditions, to minimise any effect of adaptation to other selection pressures in determining genetic and phenotypic relationships. *P. fluorescens* SBW25

and SBW25  $\phi$ 2 undergo extensive coevolution (reciprocal evolution of bacterial resistance and phage infectivity) both in nutrient media and in soil microcosms (Buckling & Rainey 2002; Gomez & Buckling 2011), and show evidence of no LA, host LA and parasite LA depending on the ecological context and time spent (co)evolving (Gomez *et al.* 2015; Gomez & Buckling 2011; Gorter *et al.* 2016; Lopez Pascua *et al.* 2012; Morgan *et al.* 2005). For example, in soil communities phages are significantly locally adapted to bacterial hosts with this trend increasing when measured through time (24 versus 48 days) (Gomez & Buckling 2011). Conversely, bacteria are more resistant to sympatric rather than allopatric phages when coevolved in nutrient rich laboratory microcosms, again with LA increasing through time (Morgan *et al.* 2005). Moreover, phage LA is increased when measured between populations coevolved in different resource environments (Lopez Pascua *et al.* 2012), while a match between the temperature in which populations coevolved and are subsequently assayed increases LA in both bacteria and phages (Gorter *et al.* 2016).

Here, we focus on parasite, rather than host genetic and phenotypic distances because previous studies suggest a strong link between mutations in the phage gene that encodes the tail fibres and changes in the infectivity phenotype (Paterson *et al.* 2010; Scanlan *et al.* 2011). In contrast, resistance phenotypes have a more complex genetic basis (Scanlan *et al.* 2015). Our current study combines a re-analysis of a previous data set which investigated within- population (sympatric) interactions (Hall *et al.* 2011), but with the additional inclusion of an unpublished data set reporting among population (allopatric) interactions of bacteria and phage isolates collected at the same time-points.

## **Materials and Methods**

### ***Coevolution experiment***

Bacteria and phages used in this study were isolated from a previously published long-term coevolution experiment (Hall *et al.* 2011). In brief bacteria and phages were coevolved in static microcosms; 25 ml glass vial containing 6 ml of M9 salt solution supplemented with 10 g/L glycerol and 20 g/L of proteose peptone no. 3 (Kassen *et al.* 2000). Six replicate coevolving populations were established by inoculating each microcosm with  $10^8$  bacterial cells (from an overnight culture) and  $10^5$  clonal phage particles, obtained from a single plaque of SBW25  $\phi$ 2

(Buckling & Rainey 2002). Bacteria and phages were coevolved in a static incubator at 28 °C for 48 hours before 1% (60 µl) of the total population was transferred to a fresh microcosm. Every tenth transfer, a sample (600µl) of the total population was frozen at -80 °C in 20% (v/v) glycerol solution and a sample of phage was isolated by adding 100 µl of chloroform to 900 µl of culture, followed by centrifugation at 13,000 rpm for 3 minutes and storage at 4°C. This process was repeated for a total of 60 transfers (approx. 400 bacterial generations). Individual bacterial colonies and phage plaques were isolated as previously described (Hall *et al.* 2011).

### ***Phenotype distance matrix***

We assayed ten phage plaque isolates from each of the six different populations (A to F) at three different time-points (T10, T30 and T60,  $10 \times 6 \times 3 = 180$ ) for infectivity against 73 different bacterial hosts - two hosts were randomly selected from each of the six populations at every tenth transfer from T10 to T60, and the wild type (WT). Phages were applied to soft agar plate lawns of single host bacteria using a pin replicator as previously described (Hall *et al.* 2011) and plates were scored after 24 hours incubation in a binary fashion for phage infectivity (1) or host resistance (0). For a given pair of phage isolates, we calculated the phenotypic distance between them as the Manhattan distance between their infectivity profiles, which equals the number of hosts against which one phage is infectious and the other is not. These Manhattan distances can be considered as a multivariate measure of phage phenotypes, in that they account for both how many and which hosts each phage can infect. For some analyses we compared this to a univariate measure of phenotypic distance: the difference in the number of hosts each phage can infect (host range). A pair of phages that infect different but equally large sets of hosts will have a positive Manhattan distance, but no difference in host range. Thus, across the two measures we can compare differences among phages in terms of specificity (who can infect whom) and their degree of generalism, as emphasized by previous work (Betts *et al.* 2014; Poullain *et al.* 2008).

### ***Sequencing of tail fibre gene and genotype distance matrix***

Based on *a priori* sequence analysis (Paterson *et al.* 2010; Scanlan *et al.* 2011) we used the tail fibre gene as a biomarker for genotypic distances among phage isolates and from the ancestral sequence. The tail fibre gene of the ancestral phage is 1818 nucleotides in length and was fully sequenced using multiple PCR reactions to cover the entire length of the gene. A modified PCR

protocol was optimised to amplify the phage tail fibre gene directly from phage stocks and PCRs were conducted on all isolates ( $n = 180$ ). Products were assayed by gel electrophoresis prior to commercial sequencing. Sequence data could only be generated for 130 of the 180 phage isolates and as such only successfully sequenced genotypes/phenotypes were included in any further analyses. We then translated each tail fibre sequence, aligned it against the ancestral sequence using MEGA (Tamura *et al.* 2007), and scored each phage isolate for the presence of nonsynonymous mutations at each codon. To calculate the genetic distance between each pair of phages, we took the Euclidean distance between their states (mutated or not) at every amino acid residue on the sequence. Sequence data from a subset of unique genotypes was then used to construct a simple Neighbour-Joining phylogenetic tree based on the number of differences method in MEGA (Tamura *et al.* 2007) to illustrate the degree of genetic relatedness within and between populations through time together with their phenotypic distance (calculated as the average Manhattan distance to all other allopatric phages from the same time-point) and LA score.

### ***Comparing genetic and phenotypic distances***

We quantified the strength of the genotype-phenotype association across phages from different populations and time points by testing whether pairwise genetic distances (the lower triangle of the  $130 \times 130$  genetic distance matrix) were correlated with corresponding phenotypic distances using mantel tests. This is a non-parametric test for correlation between two distance matrices that tests for significance by randomly permuting one matrix and testing for better-than-observed correlations arising by chance.

To determine whether the genotype-phenotype association was different within single populations compared to across multiple populations, we tested it both (i) separately for each population (A-F), including all three time points, and (ii) across all populations and time points. Our rationale here is that if different populations reach similar phenotypes through different genetic changes, or *vice versa*, this will result in a stronger genotype-phenotype association in (i) compared to (ii). Because the distance matrix in (ii) is larger than in (i), this could generate a relatively strong association because of larger sample size. To account for this, we randomly re-sampled the data 999 times, each time taking a sample of 8, 7 and 7 isolates from time points 10, 30 and 60 respectively (the same as the average sample sizes in (i)), and testing the strength of the association for each sample. Finally, we (iii) tested the phenotype-genotype association by

the same method but separately for each time point (10, 30, 60 transfers), to determine whether genotype-phenotype mapping was stronger within time points than across the whole dataset. We visualized the phenotypic and genetic distances among pairs of phages using heatmaps and multidimensional scaling as described below.

### ***Local adaptation***

We calculated local adaptation based on the home versus away definition of LA as outlined in (Kawecki & Ebert 2004). For every phage isolate we took its average infectivity against sympatric hosts from the same time point (home) minus its average infectivity against allopatric hosts from the same time point (away) (Kawecki & Ebert 2004). We then tested whether phage isolates with high LA values also tended to have relatively large genetic or phenotypic distances from allopatric isolates at the same time point. We did this using linear mixed effects models, with phage population as a random effect, time point as a fixed effect, the average phenotypic (Manhattan) or genetic (Euclidean) distance to allopatric isolates at the same time point for each phage as a covariate, and local adaptation as the response variable. We also tested whether the slope of LA against distance varied among populations and time points by their fixed interaction terms. We fitted models by maximum likelihood, and tested effects by the change in log-likelihood between full and reduced models.

There is an alternative approach to calculating LA, local versus foreign, which is based on the fitness (in this case phage infectivity) of local parasites at home (against sympatric hosts) compared to that of parasites from other (foreign) populations tested against the same hosts. We also calculated this measure of LA for our phage isolates although, as we described in the supplementary information in more detail, variation of local adaptation among phages was less pronounced by this measure due to many of the hosts used in the analysis being resistant to all or none of the relevant phages, meaning that in many cases local adaptation is zero by this measure.

## **Results**

### ***Genetic variation of sequenced isolates***

We detected considerable genetic variation in our sequenced genotypes both within and between populations and through time. We detected 80 different mutations that altered the protein sequence of the tail fibre gene. For all unique genotypes, the mean $\pm$ s.d number of



mutations per genotype was  $12.6 \pm 5.4$ . The mean  $\pm$  s.d and range of mutations for unique genotypes isolated from T10, T30 and T60 were  $8.7 \pm 3.1$ , 5-18;  $16.2 \pm 5.5$ , 10-28 and  $14.5 \pm 4.3$ ; 8-23, respectively.

Only two insertions events were detected and were very rare (confined to a single genotype in one instance and restricted to a single population for a number of sequenced genotypes at T30 and T60). Three small-scale deletion events (removal of 1-4 Amino Acid (AA) residues) were detected, two of which were common and were observed in three of the six populations (and in 54 and 64% of all unique genotypes sequenced, respectively). The remaining mutations were nucleotide changes resulting in amino acid substitutions. Seventy-five different AA changes were detected, however, a limited number (eight) of these were different AA changes at the same site for different genotypes. The AA-changing mutations observed varied from very rare to common. For example, 16 out of 75 were only observed in a single genotype whereas one was present in all genotypes sequenced. Some mutations were confined to single populations, and consequently a phylogenetic tree made with this data shows clear clustering by population for the vast majority of genotypes (Fig. S1).

### ***Genetic distance predicts phenotypic distance, particularly within populations***

Genetic distances between pairs of phage isolates were correlated with their phenotypic distances measured as the number of hosts against which they have different infectivities (Table 1; Fig. 1). This association was stronger within populations than across the whole dataset ( $P=0.002$  in one-sample  $t$ -test of within-population correlations against the average value from 999 samples of equal size from the full dataset). In other words, sympatric row-column combinations in Fig. 1 (e.g. Population A vs. Population A) are more likely to have similar values in both panels than allopatric combinations (e.g. Population A vs. Population B). This indicates that allopatric pairs of phages are more likely to be genetically similar but have different phenotypes, or *vice versa*.

### ***Allopatric phages can be genetically different but phenotypically similar***

To determine whether the results of analyses above arise from allopatric phages being genetically similar but phenotypically different, or *vice versa*, we used multidimensional scaling to plot the distances among all phages in two dimensions (Fig. 2). In cases where allopatric

phages are genetically or phenotypically similar, this would lead to points (colours) from different populations being mixed together in these plots. We found that phages clustered genetically by population very closely (Fig. 2A), but there were several instances of phages from different populations being relatively close in phenotypic space (Fig. 2B). This is equivalent to the relatively large number of red squares in allopatric combinations in Fig. 1B compared to Fig. 1A, and is consistent with phages in different populations attaining similar phenotypes through different genetic pathways. We note that *within* some populations there were some cases of phages being genetically similar but phenotypically different, suggesting that mutations at other un-sequenced loci may also play a role in infectivity evolution.

### ***Genetic differences determine which hosts, not just how many, each phage infects***

If we consider only host range (total number of hosts infected), and not the multivariate infectivity phenotype, the genotype-phenotype association is no stronger within than between populations ( $P=0.91$ ). That is, the stronger link between genetic and phenotypic distance among sympatric pairs of phages is only detectable if we look at which hosts each phage can infect, not just how many. The discrepancy between the two measures of phenotypic distance was also evident within populations: for three populations (A, C, F), genetic distance was not correlated with differences in host range (Table 1), despite being a good predictor of Manhattan distances (Table 1). These populations showed relatively large differences between pairwise Manhattan distances and the corresponding differences in host range (Fig. S2; note that this difference increases with the extent to which a pair of phages infect different but similarly sized sets of hosts). This resulted in a negative correlation between the strength of the genetic distance-host range association and the difference between Manhattan distance and host range (correlation across population averages:  $r^2=0.76$ ,  $F_{1,5}=12.42$ ,  $P=0.02$ ). In other words, in populations where it was common for phages to infect different but similarly sized sets of hosts, genetic distance was a poor predictor of host range.

### ***Parasite local adaptation decreases with genetic distance***

The extent of local adaptation (measured as average infectivity against sympatric versus allopatric bacteria from the same time point) was negatively correlated with both genetic and phenotypic distances from allopatric phages at the same time point (effects of genetic and phenotypic distances as covariates in separate mixed models:  $\text{Chisq}=18.05$ ,  $P<0.0001$ ;  $\text{Chisq}=5.93$ ,  $P=0.015$ ; Fig. 3). Note that these genetic and phenotypic distances are correlated

with each other ( $P=0.01$  for average genetic allopatric distance as a predictor of average phenotypic allopatric distance in a mixed model), consistent with our analysis above. Moreover, phages with large phenotypic distances from allopatric phages at the same time point tended to have larger host ranges (host range as a predictor of allopatric phenotypic distance:  $P<0.0001$ ). Thus, phages with a relatively large number of phenotypic differences from contemporary phages in other populations tended to have relatively broad host ranges, be relatively genetically diverged and be less locally adapted.

Despite the overall correlation of LA with phenotypic and genetic distance, the slope of this association varied among populations in both cases (population  $\times$  distance fixed interaction:  $\text{Chisq}=129.07$ ,  $P<0.0001$  and  $\text{Chisq}=59.20$ ,  $P<0.0001$  respectively). In fact *within* most populations the association was weak, with a significant negative relationship only for phenotypic distance in population A ( $P<0.001$  in a  $t$ -test for the slope of LA against distance in ANCOVA including time as a factor and distance as a covariate). By contrast, the slope of LA against distance did not vary among time points for either genetic ( $\text{Chisq}=2.18$ ,  $P=0.33$ ) or phenotypic ( $\text{Chisq}=1.60$ ,  $P=0.45$ ) distance. That the LA-distance relationships are relatively weak within populations is perhaps unsurprising given that both phenotypes and genotypes vary among populations more than within populations (Fig. 1; Fig. S1). Consistent with this, the negative association of LA with genetic distance remains significant if we take the average for each population at each time point ( $\text{Chisq}=8.42$ ,  $P=0.004$ ). However, phenotypic distance is a weaker predictor of LA when averaged across populations/time points ( $\text{Chisq}=1.69$ ,  $P=0.19$ ), suggesting that this association arises from a combination of within- and among-population variation.

When LA is calculated as 'local versus foreign' parasites, as opposed to 'home versus away' above, it is not negatively associated with either genetic or phenotypic distance (Fig. S3). However, this is likely because the variation among phage isolates was much smaller for this measure of LA (variance of LA among phage isolates: 0.18 for home versus away, 0.07 for local versus foreign), with many phages having LA (local versus foreign) equal to or close to zero (Fig. S3). The high incidence of such cases is due to many of the hosts used in this analysis (21 out of 36) being resistant to all or none of the phages included in the analysis. Consequently, in this dataset there were multiple (52 out of 130) cases where phage isolates were equally infective compared to foreign parasites (local versus foreign = 0), but performed better or worse on sympatric compared to allopatric hosts (home versus away  $\neq$  0). More generally, the standard deviation of phage infectivity across different hosts was relatively high (average across all

phage isolates=0.40, s.d.=0.07) compared to the standard deviation of infectivity across different phages for individual bacterial hosts (average across bacteria=0.27, s.d.=0.19).

## Discussion

The primary aim of this study was to determine if the magnitude of host-parasite local adaptation was correlated with genetic distance among parasite populations using coevolved populations of bacteria and bacteriophages. We found a monotonic relationship between LA and genetic distance, but in apparent contrast to previous geographical studies of LA (Ebert 1994; Imhoof & Schmid-Hempel 1998; Koskella *et al.* 2011), we found that parasite LA decreased, rather than increased, with genetic distance to allopatric parasites. We also established that the genetic distance between a pair of parasite isolates was a good predictor of their phenotypic differences, but that this relationship was strongest when we looked at the identity of hosts they could infect, rather than the total number (host range), and when measured within compared to among populations.

The key result from our analysis showing parasite LA decreases with genetic distance can be readily explained. It is typically assumed that parasites have an evolutionary advantage when coevolving with their hosts, due to their larger population sizes, shorter generation times and stronger selection pressures, and hence on average parasites are locally adapted (Kaltz & Shykoff 1998). However, this is frequently not the case (Greischar & Koskella 2007), and typically not so in this and other bacteria-phage coevolutionary systems (reviewed in (Bohannon & Lenski 2000; Buckling & Brockhurst 2012; Dennehy 2012). Here hosts often have an evolutionary advantage, presumably due to comparable generation times and population sizes with viruses, but also a greater potential to adapt due to genomic complexity (Gilman *et al.* 2012; Zaman *et al.* 2014), and hence show greater local adaptation than viruses in this particular experimental setting. In this context we would expect that the more different parasites are to the local parasite population experienced by a host, the less likely it is that the host will have correlated adaptations required to resist them. Consequently, parasites are more likely to be locally maladapted when they are genetically distant from parasites in other populations, because such parasites are likely to be more infectious against foreign hosts that are themselves adapted to their local parasites, than their local hosts. This is precisely what we see: parasites becoming increasingly maladapted by the home versus away measure we used above with genetic distance, consistent with past observations that hosts becoming increasingly locally adapted (Morgan *et al.* 2005).

While our data are consistent with LA (of host or parasite) increasing with genetic distance, there are a number of reasons why this pattern may breakdown. First, ecological factors such as migration and environmental heterogeneity can shift LA in favour of host or parasites and these may vary in time and space (Forde *et al.* 2004; Gorter *et al.* 2016). Second, our finding that genotype-phenotype correlations are stronger when measured within compared to between populations (which can simply be explained by competition and drift resulting in the loss of phenotypically equivalent but genetically diverse parasites in mixed populations), suggests that this genotype-phenotype relationship is likely to decrease with spatial scale. If this is the case, then LA for a given genetic distance is in turn likely to become weaker with spatial scale. Third, the LA-genetic distance association could be modified by variation among populations in their degree of generalism (Thrall & Burdon 2003), such that that some populations are universally more infective/resistant than others. Such generalist populations would have LA close to zero because they are similarly infective/resistant against all hosts/parasites, even though generalist phenotypes may require multiple genetic changes and therefore relatively large genetic distance (Scanlan *et al.* 2011).

One possible alternative explanation for our finding that genetically different parasites could have similar phenotypes is the finite number of phenotypic traits (hosts) measured. By taking a random sample of host bacteria from the same coevolving communities as the viruses we analyzed, this should on average include relatively common host phenotypes, and therefore capture traits that were under direct selection. Additional hosts that were not sampled here are therefore less likely to reveal phenotypic differences that were important during coevolution. Consistent with this, our sample of host phenotypes was large and diverse enough to capture some of the specificity of the host-parasite interaction: some phages infected equally sized but different sets of hosts (Fig. S2). Consequently, the multivariate phenotype estimated here mapped to genetic distances more strongly within populations than a univariate equivalent (Table 1). Although we do not exclude the possibility that this could be further strengthened by sampling additional hosts, it was sufficient for us to detect the difference in genotype-phenotype mapping within- compared to among- populations.

In summary, we show that the genetic distance between parasites is a good predictor of the extent to which they are locally adapted, albeit in this case the greater the distances the lower the likelihood of parasite LA. This confirms a key assumption of host-parasite coevolution. Finally, our data also emphasises the importance of measuring host phenotypes appropriately

and accounting for the identity of hosts that can be infected compared as opposed to the number of host genotypes that can be infected when investigating links between genotype and phenotype. Indeed, this finding may account for the poor explanatory power of genotype-phenotype maps in a range of organisms (Lehner 2013).

**Acknowledgements.** This work was funded by NERC, ERC and AXA Research fund. We are grateful to three anonymous reviewers for helpful comments on the manuscript.

**Data Accessibility:** Data files used for genetic, phenotypic and local adaptation analysis together with sequence data used to construct the phylogenetic tree are available from the Dryad Digital Repository doi:10.5061/dryad.cs128

#### References:

- Antonovics J, Boots M, Ebert D, *et al.* (2013) The Origin of Specificity by Means of Natural Selection: Evolved and Nonhost Resistance in Host-Pathogen Interactions. *Evolution* **67**, 1-9.
- Betts A, Kaltz O, Hochberg ME (2014) Contrasted coevolutionary dynamics between a bacterial pathogen and its bacteriophages. *Proc Natl Acad Sci U S A* **111**, 11109-11114.
- Bohannan BJM, Lenski RE (2000) Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecology Letters* **3**, 362-377.
- Buckling A, Brockhurst M (2012) Bacteria-Virus Coevolution. *Evolutionary Systems Biology* **751**, 347-370.
- Buckling A, Rainey PB (2002) Antagonistic coevolution between a bacterium and a bacteriophage. *Proc Biol Sci* **269**, 931-936.
- Bull JJ, Molineux IJ (2008) Predicting evolution from genomics: experimental evolution of bacteriophage T7. *Heredity (Edinb)* **100**, 453-463.
- Dennehy JJ (2012) What Can Phages Tell Us about Host-Pathogen Coevolution? *Int J Evol Biol* **2012**, 396165.
- Ebert D (1994) Virulence and local adaptation of a horizontally transmitted parasite. *Science* **265**, 1084-1086.
- Forde SE, Thompson JN, Bohannan BJ (2004) Adaptation varies through space and time in a coevolving host-parasitoid interaction. *Nature* **431**, 841-844.

- Gandon S, Capowiez Y, Dubois Y, Michalakakis Y, Olivieri I (1996) Local adaptation and gene-for-gene coevolution in a metapopulation model. *Proceedings of the Royal Society B-Biological Sciences* **263**, 1003-1009.
- Gandon S, Nuismer SL (2009) Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *Am Nat* **173**, 212-224.
- Gilman RT, Nuismer SL, Jhvueng DC (2012) Coevolution in multidimensional trait space favours escape from parasites and pathogens. *Nature* **483**, 328-330.
- Gomez P, Bennie J, Gaston KJ, Buckling A (2015) The impact of resource availability on bacterial resistance to phages in soil. *PLoS One* **10**, e0123752.
- Gomez P, Buckling A (2011) Bacteria-phage antagonistic coevolution in soil. *Science* **332**, 106-109.
- Gorter FA, Scanlan PD, Buckling A (2016) Adaptation to abiotic conditions drives local adaptation in bacteria and viruses coevolving in heterogeneous environments. *Biol Lett* **12**, 20150879.
- Greischar MA, Koskella B (2007) A synthesis of experimental work on parasite local adaptation. *Ecology Letters* **10**, 418-434.
- Hall AR, Scanlan PD, Morgan AD, Buckling A (2011) Host-parasite coevolutionary arms races give way to fluctuating selection. *Ecology Letters* **14**, 635-642.
- Heath MC (1981) A Generalized Concept of Host-Parasite Specificity. *Phytopathology* **71**, 1121-1123.
- Imhoof B, Schmid-Hempel P (1998) Single-clone and mixed-clone infections versus host environment in *Crithidia bombi* infecting bumblebees. *Parasitology* **117** ( Pt 4), 331-336.
- Kaltz O, Shykoff JA (1998) Local adaptation in host-parasite systems. *Heredity* **81**, 361-370.
- Kassen R, Buckling A, Bell G, Rainey PB (2000) Diversity peaks at intermediate productivity in a laboratory microcosm. *Nature* **406**, 508-512.
- Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters* **7**, 1225-1241.
- Koskella B, Thompson JN, Preston GM, Buckling A (2011) Local biotic environment shapes the spatial scale of bacteriophage adaptation to bacteria. *Am Nat* **177**, 440-451.
- Laine AL (2008) Temperature-mediated patterns of local adaptation in a natural plant-pathogen metapopulation. *Ecology Letters* **11**, 327-337.
- Lange B, Kaufmann AP, Ebert D (2015) Genetic, ecological and geographic covariables explaining host range and specificity of a microsporidian parasite. *J Anim Ecol* **84**, 1711-1719.
- Lehner B (2013) Genotype to phenotype: lessons from model organisms for human genetics. *Nat Rev Genet* **14**, 168-178.



- Longdon B, Hadfield JD, Webster CL, Obbard DJ, Jiggins FM (2011) Host Phylogeny Determines Viral Persistence and Replication in Novel Hosts. *Plos Pathogens* **7**.
- Lopez Pascua L, Gandon S, Buckling A (2012) Abiotic heterogeneity drives parasite local adaptation in coevolving bacteria and phages. *J Evol Biol* **25**, 187-195.
- McCoy KD, Boulinier T, Schjorring S, Michalakakis Y (2002) Local adaptation of the ectoparasite *Ixodes uriae* to its seabird host. *Evolutionary Ecology Research* **4**, 441-456.
- Morand S, Manning SD, Woolhouse MEJ (1996) Parasite-host coevolution and geographic patterns of parasite infectivity and host susceptibility. *Proceedings of the Royal Society B-Biological Sciences* **263**, 119-128.
- Morgan AD, Gandon S, Buckling A (2005) The effect of migration on local adaptation in a coevolving host-parasite system. *Nature* **437**, 253-256.
- Paterson S, Vogwill T, Buckling A, *et al.* (2010) Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275-278.
- Perlman SJ, Jaenike J (2003) Infection success in novel hosts: An experimental and phylogenetic study of *Drosophila*-parasitic nematodes. *Evolution* **57**, 544-557.
- Perry EB, Barrick JE, Bohannan BJ (2015) The Molecular and Genetic Basis of Repeatable Coevolution between *Escherichia coli* and Bacteriophage T3 in a Laboratory Microcosm. *PLoS One* **10**, e0130639.
- Poullain V, Gandon S, Brockhurst MA, Buckling A, Hochberg ME (2008) The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. *Evolution* **62**, 1-11.
- Scanlan PD, Hall AR, Blackshields G, *et al.* (2015) Coevolution with bacteriophages drives genome-wide host evolution and constrains the acquisition of abiotic-beneficial mutations. *Mol Biol Evol* **32**, 1425-1435.
- Scanlan PD, Hall AR, Lopez-Pascua LD, Buckling A (2011) Genetic basis of infectivity evolution in a bacteriophage. *Mol Ecol* **20**, 981-989.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596-1599.
- Thrall PH, Burdon JJ (2003) Evolution of virulence in a plant host-pathogen metapopulation. *Science* **299**, 1735-1737.
- Zaman L, Meyer JR, Devangam S, *et al.* (2014) Coevolution Drives the Emergence of Complex Traits and Promotes Evolvability. *Plos Biology* **12**.



**Table 1.** Correlation between genetic and phenotypic distance at different scales and with different measures of phenotypic distance (Manhattan distance or host range difference).

Population	Time point	Correlation coefficient (Manhattan distance)	Correlation coefficient (Host range difference)
All	All	0.33 (s.d.=0.09)	0.23 (s.d.=0.09)
A	All	0.45 (P=0.002)	0.11 (P=0.11)
B	All	0.61 (P=0.002)	0.52 (P=0.002)
C	All	0.67 (P=0.002)	-0.1 (P=0.29)
D	All	0.67 (P=0.002)	0.52 (P=0.002)
E	All	0.78 (P=0.002)	0.27 (P=0.007)
F	All	0.52 (P=0.002)	-0.02 (P=0.49)
All	10	0.37 (P=0.002)	0.41 (P=0.002)
All	30	0.51 (P=0.002)	0.38 (P=0.002)
All	60	0.54 (P=0.002)	0.43 (P=0.002)

**Figure Legends:**

**Figure 1.** Genetic distances and phenotypic distances across all pairs of phages, shaded red to light yellow from high to low similarity. On each axis letters denote allopatrically separated populations and numbers give time points.

**Figure. 2.** Genetic (A) and phenotypic distances (B) among phage isolates (points) plotted using multidimensional scaling. Each population has a different colour; labels give the population and time-point from which each phage was isolated. Coordinates were obtained by classical multidimensional scaling of the genetic and phenotypic distance matrices using the *cmdscale* function of the *stats* package in R v3.2.3.

**Figure. 3.** Local adaptation for phage isolates from different populations (shown by colour) and time points (given as transfer number for each point). Each point is a single phage isolate, with local adaptation calculated as average infectivity against sympatric hosts from the same time point minus average infectivity against allopatric hosts from the same time point (home versus away), and genetic and phenotypic as the average distance for each phage to contemporary allopatric phages.



