MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 760–774 doi: 10.1111/j.1365-294X.2009.04508.x

Identification of a candidate adaptive polymorphism for *Drosophila* life history by parallel independent clines on two continents

ANNALISE B. PAABY,* MARK J. BLACKET,† ARY A. HOFFMANN† and PAUL S. SCHMIDT*

Department of Biology, University of Pennsylvania, Philadelphia, PA, USA, †Centre for Environmental Stress and Adaptation Research, Department of Genetics, University of Melbourne, Melbourne, Vic., Australia

Abstract

Life history traits are critical components of fitness and frequently reflect adaptive responses to environmental pressures. However, few genes that contribute to natural life history variation have been identified. Insulin signalling mediates the determination of life history traits in many organisms, and single gene manipulation in *Drosophila melanogaster* suggests that individual genes in the pathway have the potential to produce major effects on these quantitative traits. We evaluated allelic variation at two insulin signalling genes, the *Insulin-like Receptor* (*InR*) and its substrate, *chico*, in natural populations of *D. melanogaster*. We found different patterns of variation: *InR* shows evidence of positive selection and clines in allele frequency across latitude; *chico* exhibits neutral patterns of evolution. The clinal patterns at *InR* are replicated between North America and Australia, showing striking similarity in the distribution of specific alleles and the rate at which allele frequencies change across latitude. Moreover, we identified a polymorphism at *InR* that appears to be functionally significant and consistent with hypothetical patterns of selection across geography. This polymorphism provides new characterization of genic regions of functionality within *InR*, and is likely a component in a suite of genes and traits that respond adaptively to climatic variation.

Keywords: *chico*, cline, *Insulin-like Receptor*, latitude, life history

Received 3 June 2009; revision received 2 October 2009; accepted 14 October 2009

Introduction

Life history traits include aspects of an organism's biology that directly affect reproduction and survival (Roff 1992; Stearns 1992). These traits are closely connected with fitness, and life history strategies have provided valuable examples of adaptive responses to natural selection (e.g. Endler 1986; Reznick *et al.* 1990; Martin 1995). Many phenotypic and genetic components contribute to life history, and complex and correlated suites of traits have evolved to maximize the fitness of life history regimes. But despite the vast diversity of life history strategies across taxa, empirical studies have demonstrated widely shared correlations among life history traits. Two major correlations underlying life history determination include the negative correlation between reproduction and survival, and the positive correlation between longevity and stress tolerance (Reznick 1985; Stearns 1991; Partridge *et al.* 2005; Vermeulen & Loeschcke 2006; Harshman & Zera 2007; Toivonen & Partridge 2009). These relationships between traits are mediated by pleiotropic genic elements, or by genic elements that affect single traits but co-occur through linkage disequilibrium and act as pleiotropic alleles. Such pleiotropic effects of individual genes are routinely observed in mutation genetics experiments, and indicate that genetic determinants that affect one component of life history are also likely to affect others (Partridge *et al.* 2005; Paaby & Schmidt 2009).

Patterns of phenotypic variation across environmental heterogeneity can indicate adaptive responses to selection, and evaluating these patterns can yield insight into

Correspondence: Annalise B. Paaby, Department of Biology, New York University, New York, NY, USA, Fax: +1 212 995 4015; E-mail: apaaby@nyu.edu

the evolutionary dynamics of natural populations. For example, variation in coat pigmentation in the oldfield mouse *Peromyscus polionotus* suggests adaptation to changes in substrate reflectance (Mullen & Hoekstra 2008), and flower colour in the desert plant *Linanthus parryae* likely represents adaptation to different soil environments (Schemske & Bierzychudek 2007). Examination of life history variation across climatic gradients may provide similar understanding into how this complex suite of traits evolves. Natural populations of *Drosophila melanogaster* exhibit variation in many traits across latitudinal clines, including life history phenotypes: high latitude populations exhibit longer lifespan, lower fecundity, higher incidence of reproductive diapause, larger body size, and higher cold stress tolerance than low latitude populations (Capy *et al.* 1993; Mitrovski & Hoffmann 2001; De Jong & Bochdanovits 2003; Schmidt *et al.* 2005a; Trotta *et al.* 2006). Genetic variance for and genetic correlations among these traits indicate that selection in the local environment may act on certain phenotypes but drive expression of others through tradeoffs (David 1975; Anderson *et al.* 2003; De Jong & Bochdanovits 2003; Schmidt *et al.* 2005b; Rako *et al.* 2007; Schmidt & Paaby 2008). Variation in stress tolerance has been especially well described, and may be a key component in life history evolution (Hoffmann & Harshman 1999; Hoffmann *et al.* 2001, 2005; 2007). This framework suggests a hypothetical selection regime: high latitude, seasonally cold climates impose seasonal stress and favour genotypes that confer stress tolerance; correlated traits, which may evolve as co-adapted responses to the same selection regime or by indirect selection via pleiotropy, are characterized by better overwintering ability, larger body size, longer lifespan, slower development and lower fecundity relative to low latitude populations (Paaby & Schmidt 2009).

Despite the well-characterized life history variation in natural populations of *D. melanogaster*, few genes have been shown to contribute to these observed phenotypic patterns (De Luca *et al.* 2003; Carbone *et al.* 2006; Paaby & Schmidt 2008; Schmidt *et al.* 2008). However, likely candidate genes include those characterized by extended longevity mutant or overexpression phenotypes in model systems (e.g. Kenyon *et al.* 1993; Lin *et al.* 1998; Rogina *et al.* 2000; Clancy *et al.* 2001; Tatar *et al.* 2001; Hwangbo *et al.* 2004), which routinely show correlated responses in other life history traits (Partridge *et al.* 2005; Paaby & Schmidt 2009). In particular, the insulin signalling pathway and its pleiotropic determination of a suite of correlated traits has been well characterized. Insulin signalling is conserved across taxa and appears universally relevant to the determination of life history phenotypes in metazoans: mutations that reduce insulin signalling in *D. melanogaster*, mutations in the dauer pathway in *Caenorhabditis elegans* and reduction of IGF-I in mice produce correlated phenotypic responses that include increases in lifespan and stress tolerance and decreases in reproductive success (Partridge & Gems 2002; Tatar *et al.* 2003; Giannakou & Partridge 2007). A reduction in insulin signalling by manipulation at almost any point in the pathway can induce these pleiotropic effects in either *D. melanogaster* or *C. elegans*, as demonstrated by genetic manipulations of multiple elements in both systems (Giannakou & Partridge 2007). Whether loci within this pathway harbour significant natural genetic variation may be a function of position: across 12 *Drosophila* species, downstream elements show evidence of stronger purifying selection than upstream elements (Alvarez-Ponce *et al.* 2009). Two upstream pathway members include the *Insulin-like Receptor* (*InR*) and the receptor substrate, *chico*. Disruption at either gene reduces insulin signalling and produces similar mutant phenotypes, including longevity extension, increased stress tolerance, decreased growth and development and compromised reproduction (Clancy *et al.* 2001; Tatar *et al.* 2001).

The characterization of these genes in mediating organism life history offers an opportunity to identify the loci that are involved in life history evolution. However, aging genes characterized by mutation genetics do not necessarily contribute to genetic variance for traits in the wild (Curtsinger 2003; Flatt 2004). For example, genes under strong selective constraints may vary little in natural populations, or genes may harbour only neutral polymorphisms. But populations do harbour substantial allelic variation for lifespan: artificial selection experiments (e.g. Rose & Charlesworth 1981; Promislow *et al.* 1996), genetic correlation analyses (e.g. Tatar *et al.* 1996; Schmidt *et al.* 2005a), and identification of aging genes or chromosomal regions through QTL analysis (Nuzhdin *et al.* 1997; Leips & Mackay 2000; Mackay 2002; De Luca *et al.* 2003; Geiger-Thornsberry & Mackay 2004; Pasyukova *et al.* 2004; Wilson *et al.* 2006) have demonstrated significant genetic variance for this trait in natural populations of *D. melanogaster*. Consequently, the evaluation of natural genetic variation at candidate genes has the potential to identify functionally significant polymorphisms that affect lifespan and other life history phenotypes, and provide insight into genic function that is complementary to investigations using mutation genetics.

In *D. melanogaster*, characterization of nucleotide targets of selection is facilitated by the distribution of natural populations across latitudinal gradients. *D. melanogaster* originated in tropical Africa, and has colonized temperate regions in the North American and Australian continents within the last few hundred years (David & Capy 1988). Surveys at multiple loci have

revealed patterns of variation along latitudinal clines in which the frequencies of the derived alleles increase with latitude, suggesting adaptation to novel, temperate habitats (Sezgin *et al.* 2004; Hoffmann & Weeks 2007). While patterns across geography can result from stochastic processes, independent clinal patterns replicated on multiple continents imply selection. For example, clinal variation at the alcohol dehydrogenase locus has been documented on several continents, implicating thermal regimes in the maintenance of allelic variation (Oakeshott *et al.* 1982). Furthermore, analysis of genome-wide tiling arrays generated from temperate and subtropical populations has revealed multiple regions of differentiation and parallel responses between North America and Australia (Turner *et al.* 2008). Consistency between continents likely reflects responses to independent but similar selection pressures, as North American and Australian populations were founded at different times and from different source populations (Bock & Parsons 1981; David & Capy 1988). Clines have also been observed in the distribution of multiple chromosomal inversions, which are likely maintained by selection on genes within or near the inversions (Hoffmann *et al.* 2004).

The genetic correlations between life history traits in natural populations (Schmidt *et al.* 2005b) and the pleiotropic expression of these same traits by mutation analyses in insulin signalling genes (Clancy *et al.* 2001; Tatar *et al.* 2001) suggest that *InR* and *chico* might contribute to the observed genetic variance for these phenotypes in the wild. Here, we evaluated allelic variation at these loci, derived from populations spanning latitudinal gradients on two continents, for evidence of historical and contemporaneous selection. We observed substantial differences in the molecular evolution of *InR* and *chico*, suggesting that different members of the pathway have responded differently to selection. Furthermore, we identified a polymorphism in the first exon of *InR* that shows a striking nonrandom distribution across both continents and functional effects on phenotype. These results suggest that this polymorphism may play a role in the determination of adaptive life history phenotypes, and contributes new characterization of genic regions of functionality within the *Drosophila* insulin receptor.

Materials and methods

Population samples

To sequence *InR* and *chico*, *D. melanogaster* isofemale lines were established from North American populations in Bowdoinham, ME (44.01 °N latitude, 69.90 °W longitude), New Hope, PA (40.36 °N, 74.95 °W), Lawrenceville, NJ (40.29 °N, 74.73 °W), Orlando, FL (28.54 °N, 81.38 °W) and Homestead, FL (25.47 °N, 80.48 °W). Second and third chromosomes were extracted using the CyO (stock 5439 from the Bloomington Stock Center) and TM6B (stock 279 from the Bloomington Stock Center) balancers, to permit sequencing at *chico* (second chromosome) and *InR* (third chromosome). Putatively clinal *InR* polymorphisms identified by sequencing were then screened at larger sample sizes, using genomic preparations from isofemale lines or wild-caught males from additional collections from North American populations (encompassing a total of nine populations across 18.6 ° latitude) and from wild-caught females (or single F1 individuals from wild-caught females) from populations in eastern Australia (encompassing a total of 17 populations across 27.3 ° latitude). The North American samples were collected from Bowdoinham, ME, Lawrenceville, NJ, Eutawville, SC (33.39 °N, 80.34 °W), Morven, GA (30.94 °N, 83.50 °W), Jasper, FL (30.54 °N, 82.95 °W), Fort Pierce, FL (27.45 °N, 80.33 °W) and Homestead, FL. North American lines were also provided by W. Eanes, which included chromosome-extracted lines from Mount Sinai, NY (40.95 °N, 72.84 °W) and by T. Morgan, which included inbred lines from Raleigh, NC (35.77 °N, 78.64 °W). A total of 537 North American lines were used in the polymorphism screen. The Australian samples were collected from Cooktown (15.47 °S, 145.25 °E), Cape Tribulation (16.02 °S, 145.48 °E), Innisfail (17.52 °S, 146.03 °E), Cardwell (18.25 °S, 146.03 °E), Mackay (21.13 °S, 149.18 °E), Gladstone (23.83 °S, 151.25 °E), Maryborough (25.53 °S, 152.93 °E), Brisbane (27.47 °S, 153.02 °E), Kingscliff (28.25 °S, 153.57 °E), Coffs Harbour (30.30 °S, 153.13 °E), Port Maquarie (31.42 °S, 152.92 °E), Sydney (33.87 °S, 151.22 °E), Nowra (34.87 °S, 150.60 °E), Bega (36.67 °S, 149.83 °E), Dromana (38.33 °S, 144.97 °E), Spreyton (41.22 °S, 146.25 °E), and Sorrell (42.78 °S, 147.42 °E). A total of 384 Australian lines were used in the polymorphism screen. Both alleles were counted in all samples except those from Mount Sinai and Raleigh, which had isogenic chromosomes. Lines were screened for the *In(3R)Payne* inversion following Matzkin *et al.* (2005) for the North American samples and Anderson *et al.* (2005) for the Australian samples. The association between the *InR* indel polymorphism and *In(3R)Payne* was evaluated by nominal logistic regression, modelling the log odds (standard/inverted), using JMPv5 (SAS Institute).

Sequencing

To sequence North American samples, genomic DNA was extracted from whole flies using the Wizard SV Genomic DNA Purification System (Promega), the *InR*

and *chico* genes were amplified in sections of approximately 500 bp by polymerase chain reaction, and the products were sequenced in both directions by the University of Pennsylvania DNA Sequencing Facility using an ABI 3100 sequencer (Applied Biosystems). Overlapping sequence fragments were assembled into gene sequences for each line using Sequencher 4.8 (Gene Codes Corporation). A total of 27 *chico* and 41 *InR* sequences were determined. Sequences have been deposited in GenBank under accession nos. GQ927177–GQ927244. To characterize the glutamine-histidine indel length polymorphism in the Australian populations, an approximately 240 bp fragment in the first exon at *InR* was sequenced from 28 Australian samples. *InR* and *chico* sequences from *D. simulans* and *D. yakuba* were acquired electronically through the UCSC Genome Bioinformatics website, http://genome.ucsc.edu/. Comparison of the *InR* indel sequence region across 12 *Drosophila* species was performed with the Vista Genome Browser, http://pipeline.lbl.gov/.

Sequence analysis

Sequences were aligned using ClustalW. Analyses of polymorphism and linkage disequilibrium, estimates of nucleotide diversity, and tests for neutrality and divergence were performed with DnaSPv3.14 (Rozas & Rozas 1999). Linkage disequilibrium between the indel polymorphism in the first *InR* exon and the 177 observed SNPs at that locus was evaluated by computing Fisher Exact Test *P*-values for three by two contingency tables in the R programming language. For expediency, only sequences with the three most common indel alleles (representing 87% of the total alleles) were used in the analysis (seven sequences were excluded). The selection coefficient was estimated using unfolded and folded configurations of silent and replacement changes by applying the Poisson Random Field method to the *InR* and *chico* sequence datasets, following Hartl *et al.* (1994). The frequency distributions at sites were polarized using the *D. simulans* sequence as an outgroup.

Survey for polymorphism

Three SNPs were screened at larger sample sizes, described above, using differential restriction enzyme digestion after PCR amplification of specific fragments. *Btg*I cut the SNP at position 1468, *Mbo*I cut the SNP at position 3052 and *Bsg*I cut the SNP at position 4531 (positions relative to GenBank accession no. GQ927244). Primer pairs were as follows: for SNP 1468, forward AACCCAACTGGTGGTGCTG, reverse GCAGAGTTTGCTGTTCCAG; for SNP 3052, forward GGTCATGGTATTAAGCAATTTG, reverse AGGCTTATCTAGGTAGCTCC; for SNP 4531, forward TCAATAGCGGGATACGGC, reverse GAGCCAACTGAATGATGTTC. Fragments were amplified in 30 cycles using 1.5 mM Mg^{2+}; annealing temperatures were 55 °C (SNP 1468) and 52 °C (SNPs 3052 and 4531). The *InR* indel polymorphism was screened at larger sample sizes after fragment amplification with a fluorescent-tagged primer, using either an Applied Biosystems 3100 capillary sequencer or using 6% acrylamide gels on a Licor system (IR^2, BioSciences). Primer pairs for amplification of this fragment were forward CAATATCTTTAGCAACTGTCAC and reverse TTTAGGGCTTAAACTCAGTC. This fragment was amplified in 30 cycles using 2.5 mM Mg^{2+} with an annealing temperature of 51.1 °C. Digestion conditions followed protocols provided by New England Biolabs. Purification by drop dialysis (Millipore 0.025 μm VSPW filter) of the PCR product containing SNP 4531 was usually required for clean digestion by *Bsg*I. The indel identities named in this paper refer to fragment lengths generated by these primers. The relationships between allele frequency and latitude were analysed by linear regression, using JMPv5 (SAS Institute).

Phenotype assays

The effects of the 248 and 254 *InR* alleles on phenotype were tested in three assays, using lines developed to minimize confounding background effects. All lines were maintained on standard cornmeal-molasses media at room temperature and subject to ambient light cycles. First, a third chromosome carrying the 248 allele and a third chromosome carrying the 254 allele, derived from the Mount Sinai population, were extracted using balancers and put in a background with isogenic X (w* from stock 2475 from the Bloomington Stock Center) and second (from stock 6326) chromosomes. A stock with fused second and third chromosomes (stock 2475) was used to facilitate the selection of +/CyO;+/TM3 progeny in the F1 generation. The two extracted lines carrying the 248 and 254 chromosomes were crossed, and the offspring permitted to recombine for four generations. Individual third chromosomes in the F4 generation were again extracted with balancers, and each line was genotyped for the *InR* indel polymorphism and for two of the clinal SNPs, 3052 and 4531 (SNP 1468 was identical between the parentals). Four 248 lines and four 254 lines, for which genetic variation was randomly distributed on the third chromosome (outside of the interval containing *InR*) across all lines, and for which the X and second chromosomes were isogenic across all lines, were used in each assay. Bottle cultures were reared at low density

to minimize confounding environmental effects. Flies were collected, freshly eclosed, over 24 h, and sorted into vials of five males and five females each. To measure tolerance to oxidative stress, flies in three replicate vials of each of the eight lines were aged at 25 °C for 5 days, then transferred into media-free vials with cotton saturated with 1 mL of 30 mM methyl viologen (paraquat) in 5% sucrose solution. After 48 h of continuous exposure, patterns of mortality were determined for females in all replicates. The oxidative stress data were analysed with nominal logistic regression, modelling the log odds (mortality/survivorship). To measure recovery from chill coma, flies in five replicate vials of each of the eight lines were aged at 25 °C for 4 days, then females were removed (with minimal CO_2 exposure) and aged another 24 h in vials with fresh media. To induce cold stress, vials were completely covered in ice and placed at 4 °C for 3 h, then restored to room temperature. Time to recovery (transition to the upright position) was recorded using a video camera and analysed by ANOVA. To measure fecundity, flies were transferred to fresh vials (without topical yeast) and eggs were counted every day for 2 weeks. Three replicates of each of the eight lines were used in this assay. Cumulative fecundity was analysed by ANOVA. Line nested within allele was treated as a random effect in all statistical analyses, which were performed using JMPv5 (SAS Institute).

Results

Polymorphism

Sequencing revealed a trend in allele frequency across latitude for several *InR* polymorphisms, including an amino acid insertion–deletion (indel) polymorphism in the first exon. This indel polymorphism disrupts repeats of glutamine and histidine, and the same six sequence variants were observed in both the North American and Australian sequence datasets (Fig. 1). The remainder of the *InR* locus exhibited substantial polymorphism in the North American sequence data: 88 synonymous single nucleotide polymorphisms (SNPs), 15 nonsynonymous SNPs, and 14 other indels (all but one of which are intronic) were observed. Of the observed polymorphisms at *InR*, 13 showed trends in frequency across latitude, including the indel polymorphism in the first exon and 12 SNPs well-distributed across the locus (Fig. 2). Considerable polymorphism was also found at *chico*: 25 synonymous SNPs, 12 nonsynonymous SNPs, and one indel were observed. However, no trend in allele frequency across latitude was observed at *chico*.

Our sequencing results suggested that there might be four regions of functional interest at *InR*. Linkage disequilibrium analysis showed that the 12 clinal SNPs at *InR* cluster tightly into three regions of disequilibrium (Fig. 2). Within each cluster, the SNPs are in perfect or

```
D. melanogaster     RRRRQHQQQHHH-YQH---------HHQQHLQRQQANVSYT-KK   236
                    RRRRQHQQQHHH-YQH------HQHHHQQHLQRQQANVSYT-KF   245
                    RRRRQHQQQHHHHYQHH------QHHHQQHHQRQQANVSYT-KF   248
                    RRRRQHQQQHHHHYQHHH-----QQHHQQHHQRQQANVSYT-KF   251
                    RRRRQHQQQHHH-YQHHQH---HQHHHQQHLQRQQANVSYT-KF   254
                    RRRRQHQQQHHH-YQHHQHHQHHQHHHQQHLQRQQANVSYT-KF   263

D. simulans         RRRRQHQQQQHHHHQHHP-----QHHHQQ-----QAHVSYT-KF
D. sechellia        RRRRQHQQQHHHNHQHHP-----QHHHQQHHQRQQAHVSYT-KF
D. yakuba           RRRRQHQHKQQ--------------------QQQQVHVNYT-KF
D. erecta           RRRRQQQQQQQQ------------------YHQQQAHVNYT-KF
D. ananassae        RRRR--EQRQQ--------------------QQQ--DFNFT-KF
D. pseudoobscura    RRRRQQQQQRQE------------------PHQKQKFVNYT-KF
D. persimilis       RRRRQQQQQRQE------------------PHQKQKFVNYT-KF
D. willistoni       SHHHQHQQQQQH---------------QHRHHHRRNQICIN-DI
D. virilis          QQQRQHPPHQQT---------------VQHCRLQQQHAAPA-TC
D. mojavensis       QQQQQQQQYKY--------------------QQEQRASST-TF
D. grimshawi        QQQTPKQQQRES--------------LRQFMKDKNPRRAYTYTF
```

Fig. 1 Polymorphism and divergence of *InR* amino acid sequences containing the indel polymorphism. The same six variants were observed in both the North American and Australian sequence datasets, shown here; four additional, rare fragment length variants were observed in the extended survey (233, 239, 242, 275). The most common allele, 248, which is at high frequency at high latitudes, and the second most common allele, 254, which is at high frequency at low latitudes, are in black; the four remaining alleles are in grey.

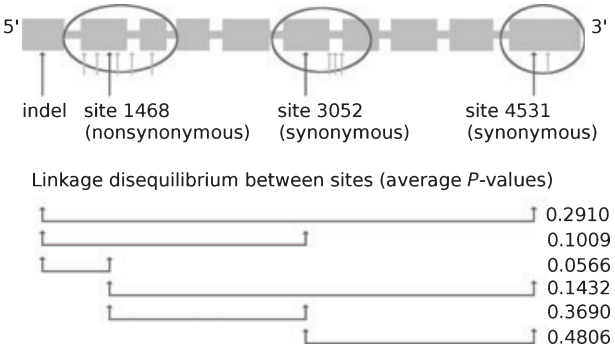

Fig. 2 Diagram of the *InR* gene showing clinally varying polymorphisms. Grey boxes are exons, interior grey lines are introns. The indel polymorphism and the 12 SNPs exhibiting clines in frequency in the sequence data are indicated by arrows. Black arrows indicate the four polymorphisms that were screened at larger sample sizes. From left to right, the positions of the 13 polymorphisms, beginning with the indel and with reference to GenBank accession no. GQ927244, are: 430, 1277, 1466, 1468, 1493, 1566, 1753, 3052, 3093, 3122, 3124, 4531, 4534. Circles indicate which SNPs cluster together by linkage disequilibrium, and the horizontal lines show the average disequilibrium between clusters. Linkage disequilibrium was estimated by Fisher exact test; the numbers report the average *P*-values for associations across clusters. All disequilibrium *P*-values for sites within clusters were < 0.0001. The diagram is not accurate to scale.

near-perfect disequilibrium; among clusters, the SNPs are independent. These three SNP clusters are also independent of the indel polymorphism in the first exon. Consequently, we hypothesized that the indel polymorphism and any one (or several) of the SNPs within each cluster might be functionally important, affecting phenotype and experiencing differential selection across latitude. But because only three latitudinal regions from five populations were represented in our sequence dataset, and because clinal patterns can occur randomly, we expanded our polymorphism survey with larger sample sizes, replicated on two continents.

Geographical patterns

Screened at larger sample sizes, the indel polymorphism in the first exon at *InR* exhibits a significant cline in frequency across latitude in both North America and Australia (Fig. 3, Table 1). The two most common alleles, which comprise 65.1% and 66.1% of the total alleles in North America and Australia, respectively, vary significantly with latitude. The 248 allele (named for its PCR fragment length) is rare at low latitudes and increases in frequency with latitude ($P = 0.0018$ in North America; $P < 0.0001$ in Australia); the 254 allele is common at low latitudes and decreases in frequency with latitude ($P = 0.0134$ in North America; $P = 0.0006$ in Australia). The reciprocal frequencies of the 248 and 254 alleles and the redundant clines on both continents strongly suggest nonrandom distribution by selection. Moreover, the rates at which the allele frequencies change across latitude are very similar between continents (Fig. 3). Comparisons of the regression coefficients for both alleles show no significant differences between the continents (for 248, $F_s = 0.4930$; for 254, $F_s = 3.6908$; df=1,22; critical $F = 4.301$), and on both continents the two common alleles are present at equal frequencies at approximately 25 ° latitude (Fig. 3). In addition to the six fragment length variants observed in the sequence data, this survey revealed another four (rare) alleles.

None of the other *InR* polymorphisms screened at large sample size demonstrates a significant trend across latitude on both continents. Representative SNPs from each of the three clusters showing clinal patterns in the sequence dataset were screened, but only the Australian populations showed a significant pattern of allele frequency across latitude (Fig. 4, Table 1). However, the regression coefficients for each SNP are not significantly different between continents (for SNP 1468, $F_s = 0.1384$; for SNP 3052, $F_s = 0.2274$; for SNP 4531, $F_s = 0.6512$; df = 1,22; critical $F = 4.301$). This may reflect similar responses across the continents but insufficient power for detection in North America, where fewer

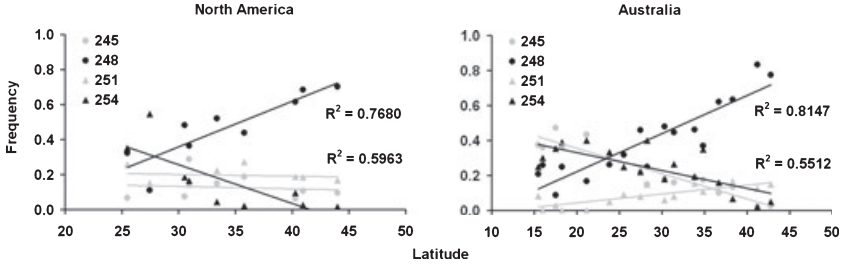

Fig. 3 Allele frequencies across latitude for the four most common variants of the *InR* indel polymorphism. The two most common variants, 248 and 254, show significant clines in frequency on both continents (see Table 1) and are shown in black. The slopes of the regression lines also show similarity between continents (North America: slope 248 = 0.0260, SE = 0.0053, slope 254 = −0.0221, SE = 0.0067; Australia: slope 248 = 0.0218, SE = 0.0027, slope 254 = −0.0102, SE = 0.0024; see text for statistical comparisons).

Table 1 Results from the linear regression analyses evaluating *InR* allele frequencies across latitude

North America					Australia				
Indel polymorphism (All samples)									
Allele	*F*	d.f.	R^2	*P*	Allele	*F*	d.f.	R^2	*P*
245	0.1183	1,7	0.0166	0.7410	245	104.0094	1,15	0.8740	< 0.0001***
248	23.6822	1,7	0.7680	0.0018**	248	65.9628	1,15	0.8147	< 0.0001***
251	0.1612	1,7	0.0225	0.7001	251	26.5220	1,15	0.6387	0.0001***
254	10.7729	1,7	0.5963	0.0134*	254	18.4251	1,15	0.5512	0.0006***
SNPs representative of the three linked clusters									
Site	*F*	d.f.	R^2	*P*	Site	*F*	d.f.	R^2	*P*
1508	1.1085	1,7	0.1303	0.3274	1508	13.0197	1,15	0.4647	0.0026**
3039	3.9521	1,7	0.3553	0.0871	3039	32.7394	1,15	0.6858	< 0.0001***
4574	0.3462	1,7	0.0431	0.5747	4574	11.7721	1,15	0.4397	0.0037**
Indel polymorphism (Standard chromosomes only)[a]									
Allele	*F*	d.f.	R^2	*P*	Allele	*F*	d.f.	R^2	*P*
245	1.3779	1,7	0.1645	0.2788	245	1.2729	1,5	0.2030	0.3104
248	8.9143	1,7	0.5601	0.0204*	248	2.3958	1,5	0.3242	0.1823
251	4.4678	1,7	0.3896	0.0724	251	0.7031	1,5	0.1231	0.4400
254	1.1042	1,7	0.1363	0.3283	254	5.4709	1,5	0.5226	0.0665
					Allele	*F*	d.f.	R^2	*P*
Indel polymorphism (Inverted chromosomes only)[a,b]									
					245	6.0345	1,9	0.4014	0.0364*
					248	11.1870	1,9	0.5542	0.0086**
					251	2.9945	1,9	0.2497	0.1176
					254	1.1762	1,9	0.1156	0.3063

Statistical significance at: $P < 0.05$ (*), $P < 0.01$ (**) and $P < 0.001$ (***) levels. [a]Only populations with seven or more of the indicated chromosomal states were used in the analysis. [b]The difference between heterozygous and homozygous inverted chromosomal arrangements could not be determined in the North American samples, so results from inverted chromosomes from these populations are not reported.

populations were screened over a narrower range of latitude. Of the 12 SNPs comprising these clusters, one affects the amino acid sequence (Fig. 2); the other 11 are synonymous or intronic. While any of these SNPs could be under direct selection, the robust clinal patterns of the indel polymorphism make it the strongest candidate functional polymorphism.

The clinal pattern of the *InR* indel polymorphism does not appear to be caused by linkage with another site on the third chromosome. The *In(3R)Payne* inversion exhibits a cline in frequency across latitude in both North America and Australia (Knibb 1982; Anderson *et al.* 2005), and worldwide there is a significant association between the indel polymorphism and the chromosomal arrangement: the 248 allele is 3.29 times more likely to be associated with the standard chromosome than is the 254 allele (odds ratio = 25.74; $P < 0.0001$). Such an association would be expected between any two clinal elements. Furthermore, two pieces of evidence suggest that the observed changes in the indel allele frequency are not driven by association with this inversion. First, in North America, the magnitudes of the indel clines are stronger than those reported for the inversion (Knibb 1982). Second, indel clines generated using only standard chromosomes in North America and only inverted chromosomes in Australia still show significance (Fig. 5, Table 1); low sample size of specific alleles at the ends of the clines (i.e. 248 at low latitudes, 254 at high latitudes) probably constrained detection of significance in additional comparisons. It is also possible that the indel's geographical patterns are driven by disequilibrium with another nucleotide polymorphism. However, of the 177 SNPs we observed within the *InR* locus, only 11 are in significant disequilibrium with the indel polymorphism (Fig. 6). Nine of these are close neighbours; the remaining two encode a

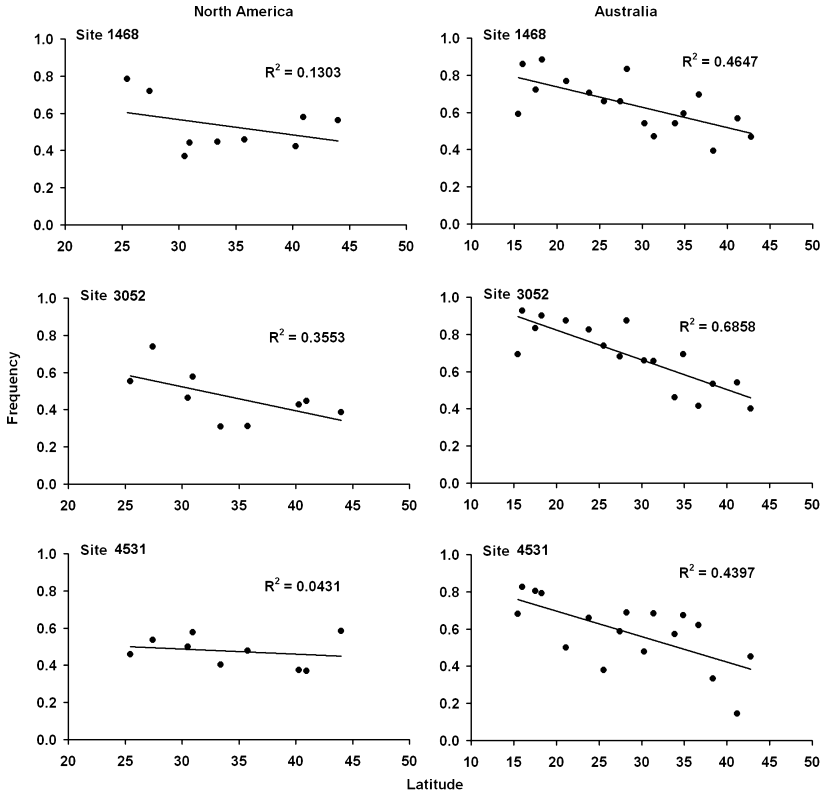

Fig. 4 Allele frequencies across latitude for three *InR* SNPs which showed clinal trends in the sequence data. Each SNP represents a cluster of SNPs exhibiting high linkage disequilibrium. Only the Australian populations showed statistically significant changes in allele frequency across latitude (see Table 1), but the slopes of the regression lines are similar between continents (North America: slope 1468 = −0.0082, SE = 0.0078, slope 3052 = −0.0129, SE = 0.0065; slope 4531 = −0.0028, SE = 0.0047; Australia: slope 1468 = −0.0110, SE = 0.0030, slope 3052 = −0.0161, SE = 0.0028; slope 4531 = −0.0136, SE = 0.0040; see text for statistical comparisons).

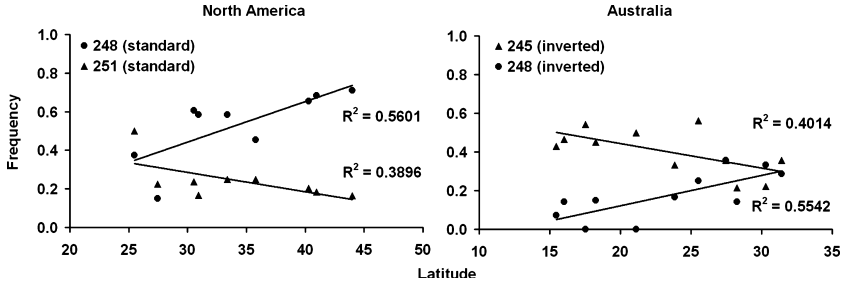

Fig. 5 Allele frequencies across latitude for *InR* indel alleles on either standard or *In3RPayne* inverted chromosomes. The alleles shown here exhibit statistically significant or near-significant changes across latitude (see Table 1).

silent third-position site and an intronic site. Hence, it is unlikely that the indel polymorphism is linked to a site within *InR* that is responsible for the cline. Furthermore, despite occurring in the 5′ region of the sequence, we observe significant decay of linkage disequilibrium both 5′ and 3′ of the indel polymorphism (Fig. 6). While we cannot rule out the possibility that this polymorphism is linked to a site outside the locus we sequenced, all our evidence suggests that the indel polymorphism is an independent site under direct selection.

Phenotype assays

In assays testing the effect of the *InR* polymorphism on stress tolerance and fecundity, lines carrying the 248 and 254 alleles showed significantly different contributions to phenotype, and in the direction predicted by the allele frequencies at low and high latitudes (Fig. 7, Table 2). Flies homozygous for the 248 allele, which is at high frequency at high latitudes, tolerated oxidative stress better than flies homozygous for the 254 allele (P = 0.0007), and recovered faster from chill coma (P = 0.0305). Alternately, flies with the 254 allele laid more eggs than those with the 248 allele (P = 0.0030). These effects are accordant with phenotypic effects of laboratory-induced mutations at *InR* and *chico*, which show increased stress tolerance and reduced fecundity when insulin signalling is reduced (Clancy *et al.* 2001; Tatar *et al.* 2001). Consequently, the differences in performance between lines carrying the naturally derived *InR* alleles imply that the 248 allele reduces insulin signalling relative to the 254 allele. The results are also

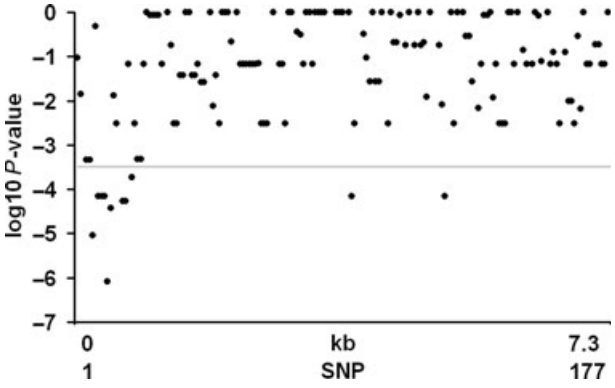

Fig. 6 Linkage disequilibrium estimates between the *InR* indel polymorphism and the 177 SNPs observed across the *InR* locus. Disequilibrium was estimated by Fisher exact test; *P*-values below the horizontal line indicate associations that are statistically significant at the 0.05 level, after correction for multiple tests. Most of the SNPs in significant linkage disequilibrium reside with a few hundred bases of the indel polymorphism, and decay of disequilibrium is demonstrated in both the 5′ and 3′ directions.

consistent with the hypothesis that the nonrandom allele distributions across latitude are driven by alternate selection regimes imposed by environments with varying degrees of seasonal stress. The effect of line (nested within allele) was significant or marginally significant for the two stress assays, but there was no line effect at all in the fecundity assay (Table 2). These results do not provide direct evidence that the 248 and 254 alleles confer differences in stress tolerance or fecundity, as the indel polymorphism was not perfectly isolated from other allelic variation in the recombinant strains used in the tests. In these strains, the two clinal SNPs at sites 3093 and 4574 retained the linked allele associations present in the original parental lines (the clinal SNP at site 1508 was identical between the parentals). Consequently, although these results demonstrate a significant association between the indel alleles and predicted phenotypes, they cannot distinguish between contributions from the indel and any other linked polymorphisms within the interval containing *InR*.

Tests for neutrality and selection

Classical tests of molecular evolution revealed evidence of selection at *InR*, but not at *chico*. *InR* showed evidence of adaptive protein evolution over the approximately 2 Myr since *D. melanogaster* shared a common ancestor with *D. simulans*: the McDonald–Kreitman test (McDonald & Kreitman 1991) for divergence demonstrated an excess of fixed replacement changes between the species (Table 3). These differences between species

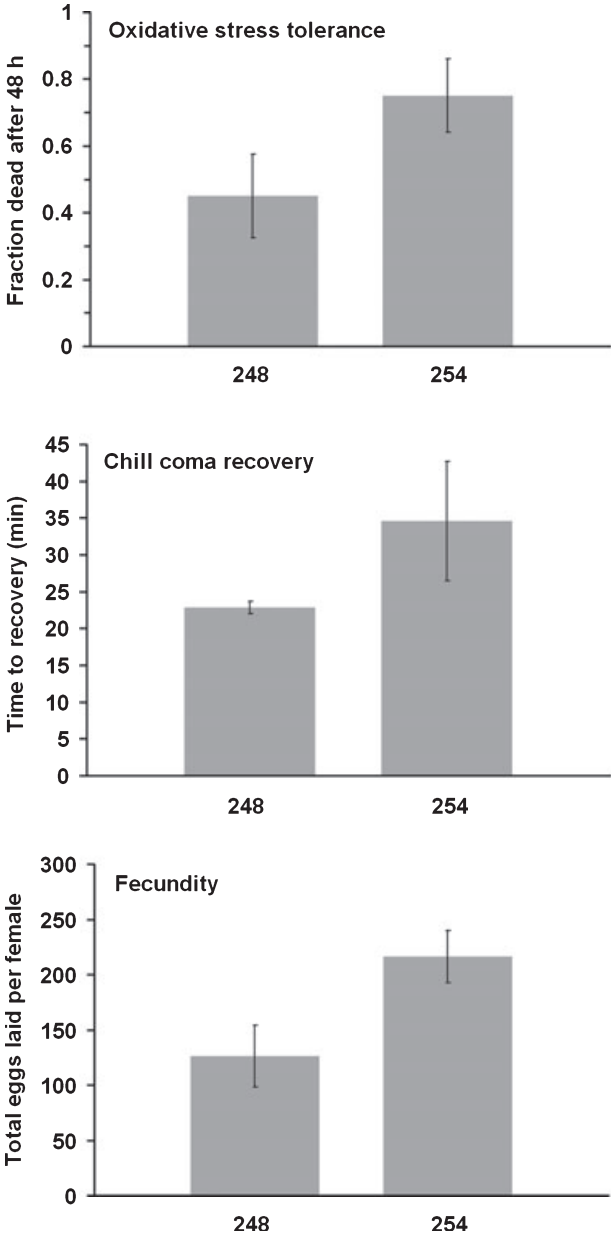

Fig. 7 Effects on phenotype of the 248 and 254 *InR* alleles. In all three assays, the alleles show contributions to phenotype that are consistent with the predicted adaptive life history response. Flies carrying the 248 allele showed better tolerance to oxidative stress, faster recovery from chill coma, and lower fecundity relative to flies carrying the 254 allele, which accords with a selection regime favouring stress tolerant alleles at high latitudes. Error bars show 95% confidence.

were evenly distributed across both lineages, according to polarization by an outgroup *D. yakuba* sequence (data not shown). Our results are qualitatively identical using sequences derived from just a single population (from Bowdoinham, ME, $N = 20$) and gene trees showed no evidence of sequence similarity by geography, indicat-

Table 2 Results of the assays testing for functional differences between *InR* alleles

Ox stress assay: Nominal logistic model effect Wald tests			
Source	d.f.	χ^2	*P*
Allele	1	11.4927	0.0007***
line[allele]	6	11.5941	0.0717

Chill coma assay: ANOVA				
Source	d.f.	MS	*F*	*P*
Allele	1	2.477×10^7	7.9284	0.0305*
line[allele]	6	3124151	7.7423	< 0.0001***
Error	16	403518		

Fecundity assay: ANOVA				
Source	d.f.	MS	*F*	*P*
Allele	1	1213650	23.0509	0.0030**
line[allele]	6	52651	0.9345	0.4973
Error	16	56341		

Statistical significance at: $P < 0.05$ (*), $P < 0.01$ (**) and $P < 0.001$ (***) levels.

Table 3 Results of the McDonald–Kreitman test for divergence at *InR* and *chico*

	Fixed differences		Polymorphisms		
	S	NS	S	NS	P^a
InR	101	37	88	15	0.03
standard	109	38	59	14	0.31
In(3R)P	106	37	62	5	< 0.01
chico	56	27	25	12	1.0

[a]*P*-values were computed by Fisher's exact test using a two-tailed distribution.

ing that the test is unlikely to be biased by population structure within our sample set (data not shown). The polymorphisms we screened at larger sample size showed that each allele state was present in each chromosomal arrangement, indicating that there is significant genetic exchange even at sites relatively close to the *In(3R)P* inversion break points. Consequently, we included sequences derived from both chromosomal arrangements in these analyses. However, the McDonald–Kreitman test at *InR* does show different results when only standard and only inverted chromosomes are evaluated: samples with standard chromosomes lose statistical significance ($P = 0.31$), largely due to a decrease in the number of synonymous polymorphisms, but samples with inverted chromosomes gain significance, following decreases in both synonymous and replacement polymorphisms ($P < 0.01$) (Table 3). The McDonald–Kreitman test showed no significant results at *chico*, and the null hypothesis of neutral protein evolution could not be rejected (Table 3). The McDonald–Kreitman test evaluates evolution over a relatively long timescale, while the clinal polymorphism screen assumes contemporaneous evolution. Since the analyses show adaptive patterns at *InR* and neutral patterns at *chico* over both macro- and microevolutionary time, these results may reflect differences in functional constraint between the genes that is both ancient and modern. *InR* also shows evidence of selection according to estimates of the selection coefficient. These estimates were determined by applying the Poisson random field (PRF) method to the frequency distributions of both silent and replacement SNPs (Hartl *et al.* 1994). Although the unfolded frequency distribution at replacement sites yielded a statistically neutral estimate of the selection coefficient ($\gamma_{est} = 0.42$, CI = −2.18 to 2.63), both the unfolded and folded distributions of silent SNPs at *InR* showed significantly positive estimates ($\gamma_{es\ t} = 4.09$, CI = 1.85 to 11.82; $\gamma_{est} = 1.98$, CI = 0.24 to 18.71). We interpret these results as evidence of selection on replacement sites linked to silent sites. Although the PRF method assumes independence between sites, this is not biologically realistic: an analysis of all polymorphisms across *InR* showed patterns of linkage disequilibrium that are typical for *D. melanogaster* (Miyashita *et al.* 1993; Long *et al.* 1998; Langley *et al.* 2000), in which disequilibrium extends for 1–2 kb (data not shown). Estimates for *chico* by both silent and replacement unfolded sites were not significantly different from zero ($\gamma_{est} = -0.23$, CI = −1.65 to 2.20; $\gamma_{est} = 0.97$, CI = −1.30 to 17.80). In our study, the estimation of the selection coefficient may not be an especially valuable measure because of the assumption of site independence in the face of known linkage disequilibrium. However, these estimates are consistent with our other results, which show evidence of selection at *InR* but neutrality at *chico*. For both genes, nucleotide diversity was estimated at levels typical for *D. melanogaster* (*InR* $\pi_{est} = 0.0049$; *chico* $\pi_{est} = 0.0033$). Tests for neutrality, including the Tajima (1989) test and the Fu & Li (1993) test, did not reveal any significant departures from the null hypothesis (for *InR*, Tajima's $D = -0.16$, $P > 0.10$, Fu and Li's $D = -1.63$, $P > 0.10$; for *chico*, Tajima's $D = -0.21$, $P > 0.10$, Fu and Li's $D = 0.26$, $P > 0.10$). While the presence of the indel polymorphism at reciprocal frequencies across latitude is evidence of balancing selection, the absence of significantly positive *D* values for the neutrality tests at *InR* may be a function of an overrepresentation of high latitude sequences in our dataset, coupled with relatively few

polymorphisms contributing to a signal for overdominance by linkage with the indel (Fig. 6).

Discussion

By examining patterns of nucleotide variation across latitude, we have identified a putative functional polymorphism in the first exon at *InR*. The distribution of the 248 allele at high latitudes and the 254 allele at low latitudes, replicated across two continents, is strongly suggestive of selection. Moreover, the functional tests suggest that the alleles confer effects that are consistent with adaptation to the environment in which they are most common: lines carrying the 248 allele, which is common in cold climates, show enhanced stress tolerance; lines carrying the 254 allele, which is common in warm climates, show higher fecundity. We hypothesize that the 248 allele reduces insulin signalling relative to the 254 allele, and that seasonally imposed stress maintains high frequencies of the 248 allele at high latitudes, which would otherwise experience negative selection due to its pleiotropic cost to reproductive success. We are currently measuring insulin signalling using biochemical and gene expression assays, to determine if signalling levels differ between the alleles. This mechanism would be consistent with the role of insulin signalling in mediating stress tolerance, body size, lipid content, longevity and reproduction (reviewed in Giannakou & Partridge 2007), observed clines in these traits across latitude (e.g. David 1975; Karan *et al.* 1998; Hoffmann *et al.* 2005; Schmidt *et al.* 2005a; Trotta *et al.* 2006; Schmidt & Paaby 2008), and the observed genetic correlations between these traits among natural strains (Hoffmann *et al.* 2001). Consequently, the identification of this polymorphism might be an example of how a specific, pleiotropic genic element may contribute to the life history tradeoffs between reproduction, stress tolerance, and very likely, longevity.

However, the phenotype assays do not provide direct evidence that the *InR* indel alleles confer differences in stress tolerance or fecundity, as the polymorphism was not completely isolated from other genetic variation within the recombinant test lines. Rather, the assays provide compelling preliminary results in support of our adaptive hypothesis, by demonstrating a significant association between chromosomal fragments harbouring distinct alleles and the expected phenotypic outcome. We have not characterized the recombination breakpoints in the test lines, but it is likely that the intervals containing the distinct *InR* indel alleles span many genes. Current work is underway to test the functional significance of this amino acid polymorphism more robustly, using higher replication and additional alleles from other source populations, and to evaluate how the polymorphism affects a broad suite of traits, including lifespan and other phenotypes, levels of insulin signalling, and other dynamics within the pathway.

Natural genetic variation at other sites may contribute to the patterns we observe. However, clinal patterns are weaker at the other three polymorphisms we screened, showing significance only in Australia. We conclude that the indel polymorphism is the most likely functional candidate polymorphism at *InR*. However, associations between phenotypes and cosmopolitan inversions have been observed in natural populations of *D. melanogaster*, and are likely due to selection at loci within or near the inversions (Hoffmann *et al.* 2004). The association of multiple insulin signalling loci with inversions, along with observed clines in body size, development, and inversion frequencies, have been proposed to be a part of the same adaptive strategy (De Jong & Bochdanovits 2003). *InR* is associated with the inversion *In(3R)Payne*, which contains alleles under selection (Kennington *et al.* 2006), is associated with clinal variation in body size (Weeks *et al.* 2002; Rako *et al.* 2006), and contributes to additive genetic variance for other phenotypes that vary clinally (Kennington *et al.* 2007). Consequently, it may be difficult to distinguish between selection at *InR* and selection on *In(3R)Payne*. For example, if *InR* experiences direct selection and contributes to differences in phenotypes mediated by insulin signalling, its proximity to *In(3R)Payne* may also contribute to the observed inversion clines; alternately, selection at other loci associated with the inversion may be driving the observed distribution of *InR* alleles. However, the persistence of clinal variation among *InR* indel alleles in standard and inverted lines, the relatively stronger cline at *InR* compared to the inversion in North America, and the decay of linkage disequilibrium both 5′ and 3′ of the indel polymorphism suggest a good measure of independence between *InR* and *In(3R)Payne*.

Our data demonstrate that for two genes in the insulin signalling pathway, only *InR* exhibits significant patterns of polymorphism and divergence that are suggestive of adaptive evolution. The clinal pattern of the indel polymorphism and the evidence for protein evolution and positive selection at *InR* suggest that this locus has undergone adaptive evolution on both short and long timescales; the evidence for protein evolution by the McDonald-Kreitman method is significant considering the potential for this test to underestimate adaptive responses (Charlesworth & Eyre-Walker 2008). We do not assume that the hypothetical selection pressures that may be influencing life history evolution in contemporaneous populations are necessarily relevant to or predictive of pressures that drove the amino acid evolution of *InR* over phylogenetic time. However, in the face of sim-

ilar functional characterization between *InR* and *chico* by mutational genetics (Clancy *et al.* 2001; Tatar *et al.* 2001), our consistent results for adaptive evolution at *InR* and neutral evolution at *chico* suggest that these genes may be independently constrained. These results are consistent with the hypothesis that although disruption of multiple genes within the insulin signalling pathway can lead to a reduction in signalling, the receptor molecule may be one of only a subset responsive to selection. This situation may be analogous to metabolic control theory, in which only some points can control flux through the pathway (Kacser & Burns 1973).

Characterization of other insulin signalling genes in other systems suggests that this pathway may be a fruitful research target in the identification of additional loci that contribute to additive genetic variance for life history traits. The transcription factor *dFOXO* is a promising candidate: overexpression of *dFOXO* extends lifespan (Hwangbo *et al.* 2004), upregulation of *dFOXO* is associated with lifespan extension by other genes in the pathway, and the *C. elegans* ortholog *daf-16* is essential for lifespan extension by insulin signalling (reviewed in Giannakou & Partridge 2007). Natural variation at the human ortholog of *dFOXO*, *FOXO3A*, contributes significantly to differences in longevity in human populations (Flachsbart *et al.* 2008; Willcox *et al.* 2008), but variation in other natural populations remains to be investigated. Polymorphism in the human *insulin-like growth factor I receptor* (*IGFIR*) is also associated with variation in human longevity (Suh *et al.* 2008). In teleost fishes, the *insulin-like growth factor II* (*IGF2*) shows evidence of selection responses that are coincident with the evolution of placentation (O'Neill *et al.* 2007). Transgenic experiments have shown that *Dp110*, which encodes the insulin-regulated phosphatidylinositol 3-kinase, affects the expression of reproductive diapause, a trait important in overwintering (Williams *et al.* 2006). Diapause is variable in natural populations, shows genetic correlations to multiple life history traits (Schmidt *et al.* 2005a), and shows strong selective responses over both spatial (Schmidt *et al.* 2005b) and temporal (Schmidt & Conde 2006) scales. However, there is an absence of pronounced polymorphism between two natural *Dp110* alleles that differentially affect diapause: of 20 polymorphisms detected, none affect the amino acid sequence, and no differences in RNA levels have been detected (Williams *et al.* 2006). Future characterization of natural populations may yield further insight into whether this gene, or others in the insulin signalling pathway, vary significantly in the wild. Such investigations of natural genetic variation at target genes can identify genic elements that contribute to phenotype evolution, as well as elucidate important dynamics within pathways and characterize genic functionality on a fine scale. Genetic manipulations in the laboratory typically demonstrate functions of whole genes; the function of smaller regions and even specific nucleotides can be resolved by evaluating natural alleles if there is sufficient variation in the wild, where natural selection can impose subtle pressure over many generations. In our study, the putatively functional *InR* allele identifies a mutation that likely contributes to genetic variance for lifespan along with associated life history phenotypes, and provides fresh insight into the functional genetics of insulin signalling.

Acknowledgements

We thank J. Plotkin for generous assistance in analysing linkage disequilibrium and estimating the selection coefficient. We also thank W. Eanes and T. Morgan for providing *D. melanogaster* samples. This work was supported by an American Federation for Aging Research/Glenn Foundation Scholarship to ABP, by National Science Foundation-DEB grant 0542859 to PSS, by a Federation Fellowship to AAH, and by the Australian Research Council via their Special Research Centre Scheme.

References

Alvarez-Ponce D, Aguadé M, Rozas J (2009) Network-level molecular evolutionary analysis of the insulin/TOR signal transduction pathway across 12 *Drosophila* genomes. *Genome Research*, **19**, 234–242.

Anderson AR, Collinge JE, Hoffmann AA, Kellett M, McKechnie SW (2003) Thermal tolerance tradeoffs associated with the right arm of chromosome 3 and marked by the *Hsr-omega* gene in *Drosophila melanogaster*. *Heredity*, **90**, 195–202.

Anderson AR, Hoffmann AA, McKechnie SW, Umina PA, Weeks AR (2005) The latitudinal cline in the *In(3R)Payne* inversion polymorphism has shifted in the last 20 years in Australian *Drosophila melanogaster* populations. *Molecular Ecology*, **14**, 851–858.

Bock IR, Parsons PA (1981) Species of Australia and New Zealand. In: *Genetics and Biology of Drosophila*, Vol. 3a (eds Asherburner M, Carson HL, Thompson JN), pp. 291–308. Academic Press, London, UK.

Capy P, Pla E, David JR (1993) Phenotypic and genetic variability of morphometrical traits in natural populations of *Drosophila melanogaster* and *Drosophila simulans*. *Evolution*, **25**, 517–536.

Carbone MA, Jordan KW, Lyman RF *et al.* (2006) Phenotypic variation and natural selection at *Catsup*, a pleiotropic quantitative trait gene in *Drosophila*. *Current Biology*, **16**, 912–919.

Charlesworth J, Eyre-Walker A (2008) The McDonald–Kreitman test and slightly deleterious mutations. *Molecular Biology and Evolution*, **25**, 1007–1015.

Clancy DJ, Gems D, Harshman LG *et al.* (2001) Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. *Science*, **292**, 104–106.

Curtsinger JW (2003) Peeking under QTL peaks. *Nature Genetics*, **34**, 358–359.

David JR (1975) Evolution of a cosmopolitan species: genetic latitudinal clines in *Drosophila melanogaster* wild populations. *Experientia*, **31**, 164–166.

David JR, Capy P (1988) Genetic variation of *Drosophila melanogaster* natural populations. *Trends in Genetics*, **4**, 106–111.

De Jong G, Bochdanovits Z (2003) Latitudinal clines in *Drosophila melanogaster*: body size, allozyme frequencies, inversion frequencies, and the insulin signalling pathway. *Journal of Genetics*, **82**, 207–223.

De Luca M, Roshina NV, Geiger-Thornsberry GL, Lyman RF, Pasyukova EG, Mackay TFC (2003) Dopa decarboxylase (*Ddc*) affects variation in *Drosophila* longevity. *Nature Genetics*, **34**, 429–433.

Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton.

Flachsbart F, Caliebe A, Kleindorp R *et al.* (2008) Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proceedings of the National Academy of Sciences, USA*, **106**, 2700–2705.

Flatt T (2004) Assessing natural genetic variation in genes affecting *Drosophila* lifespan. *Mechanisms of Ageing and Development*, **125**, 155–159.

Fu Y-Y, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Geiger-Thornsberry G, Mackay TFC (2004) Quantitative trait loci affecting natural variation in *Drosophila* longevity. *Mechanisms of Ageing and Development*, **125**, 179–189.

Giannakou ME, Partridge L (2007) Role of insulin-like signalling in *Drosophila* lifespan. *Trends in Biochemical Science*, **32**, 180–188.

Harshman LG, Zera AJ (2007) The cost of reproduction: the devil in the details. *Trends in Ecology and Evolution*, **22**, 80–86.

Hartl DL, Moriyama EN, Sawyer SA (1994) Selection intensity for codon bias. *Genetics*, **138**, 227–234.

Hoffmann AA, Harshman LG (1999) Desiccation and starvation resistance in *Drosophila*: patterns of variation at the species, population and intrapopulation levels. *Heredity*, **83**, 637–643.

Hoffmann AA, Weeks AR (2007) Climatic selection on genes and traits after a 100 year-old invasion: a critical look at the temperate-tropical clines in *Drosophila melanogaster* from eastern Australia. *Genetica*, **129**, 133–147.

Hoffmann AA, Hallas R, Sinclair C, Mitrovski P (2001) Levels of variation in stress resistance in *Drosophila* among strains, local populations, and geographic regions: patterns for desiccation, starvation, cold resistance and associated traits. *Evolution*, **55**, 1621–1630.

Hoffmann AA, Sgro CM, Weeks AR (2004) Chromosomal inversion polymorphisms and adaptation. *Trends in Ecology and Evolution*, **19**, 482–488.

Hoffmann AA, Shirriffs J, Scott M (2005) Relative importance of plastic *vs* genetic factors in adaptive differentiation: geographical variation for stress resistance in *Drosophila melanogaster* from eastern Australia. *Functional Ecology*, **19**, 222–227.

Hwangbo DS, Gershman B, Tu M-P, Palmer M, Tatar M (2004) *Drosophila* dFOXO controls lifespan and regulates insulin signalling in brain and fat body. *Nature*, **429**, 562–566.

Kacser H, Burns JA (1973) The control of flux. *Symposium for the Society of Experimental Biology*, **27**, 65–104.

Karan D, Dahiya N, Munjal AK *et al.* (1998) Desiccation and starvation tolerance of adult *Drosophila*: opposite latitudinal clines in natural populations of three different species. *Evolution*, **52**, 825–831.

Kennington WJ, Partridge L, Hoffmann AA (2006) Patterns of diversity and linkage disequilibrium within the cosmopolitan inversion *In(3R)Payne* in *Drosophila melanogaster* are indicative of coadaptation. *Genetics*, **172**, 1655–1663.

Kennington WJ, Hoffmann AA, Partridge L (2007) Mapping regions within cosmopolitan inversion *In(3R)Payne* associated with natural variation in body size in *Drosophila melanogaster*. *Genetics*, **177**, 549–556.

Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R (1993) A *C. elegans* mutant that lives twice as long as wild type. *Nature*, **366**, 461–464.

Knibb WR (1982) Chromosome inversion polymorphism in *Drosophila melanogaster*. II. Geographic clines and climatic associations in Australasia, North America and Asia.. *Genetica*, **58**, 213–222.

Langley CH, Lazzaro BP, Phillips W, Heikkinen E, Braverman JM (2000) Linkage disequilibria and the site frequency spectra in the *su(s)* and *su(w(a))* regions of the *Drosophila melanogaster X* chromosome. *Genetics*, **156**, 1837–1852.

Leips J, Mackay TFC (2000) Quantitative trait loci for life span in *Drosophila melanogaster*: interactions with genetic background and larval density. *Genetics*, **155**, 1773–1788.

Lin Y, Seroude S, Benzer S (1998) Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*. *Science*, **282**, 943–946.

Long AD, Lyman RF, Langley CH, Mackay TFC (1998) Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics*, **149**, 999–1017.

Mackay TFC (2002) The nature of quantitative genetic variation for *Drosophila* longevity. *Mechanisms of Ageing and Development*, **123**, 95–104.

Martin TE (1995) Avian life history evolution in relation to nest sites, nest predation, and food. *Ecological Monographs*, **65**, 101–127.

Matzkin LM, Merritt TJS, Zhu C-T, Eanes WF (2005) The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion in *In(3R)Payne* in *Drosophila melanogaster*. *Genetics*, **170**, 1143–1152.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature*, **351**, 652–654.

Mitrovski P, Hoffmann AA (2001) Postponed reproduction as an adaptation to winter conditions in *Drosophila melanogaster*: evidence for clinal variation under semi-natural conditions. *Proceedings of the Royal Society of London, Series B*, **268**, 2163–2168.

Miyashita NT, Aguade M, Langley CH (1993) Linkage disequilibrium in the *white* locus region of *Drosophila melanogaster*. *Genetical Research*, **62**, 101–109.

Mullen LM, Hoekstra HE (2008) Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution*, **62**, 1555–1570.

Nuzhdin SV, Pasyukova EG, Dilda CA, Zeng Z-B, Mackay TFC (1997) Sex-specific quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA*, **94**, 9734–9739.

Oakeshott JG, Gibson JB, Anderson PR, Knibb WR, Anderson DG, Chambers GK (1982) Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. *Evolution*, **36**, 86–96.

O'Neill MJ, Lawton BR, Mateos MD *et al.* (2007) Ancient and continuing Darwinian selection on *insulin-like growth factor II* in placental fishes. *Proceedings of the National Academy of Sciences, USA*, **104**, 12404–12409.

Paaby AB, Schmidt PS (2008) Functional significance of allelic variation at *methuselah*, an aging gene in *Drosophila*. *PLoS ONE*, **3**, e1987.

Paaby AB, Schmidt PS (2009) Dissecting the genetics of longevity in *Drosophila melanogaster*. *Fly*, **3**, 29–38.

Partridge L, Gems D (2002) The evolution of longevity. *Current Biology*, **12**, R544–R546.

Partridge L, Gems D, Wither DJ (2005) Sex and death: what is the connection? *Cell*, **120**, 461–462.

Pasyukova EG, Roshina NV, Mackay TFC (2004) *Shuttle craft*: a candidate quantitative trait gene for *Drosophila lifespan*. *Aging Cell*, **3**, 297–307.

Promislow DEL, Tatar M, Khazaeli AA, Curtsinger JW (1996) Age-specific patterns of genetic variance in *Drosophila melanogaster*. I. Mortality. *Genetics*, **143**, 839–848.

Rako L, Anderson AR, Sgro CM, Stocker AJ, Hoffmann AA (2006) The association between inversion *In(3R)Payne* and clinally varying traits in *Drosophila melanogaster*. *Genetica*, **128**, 373–384.

Rako L, Blacket MJ, McKechnie SW, Hoffmann AA (2007) Candidate genes and thermal phenotypes: identifying ecologically important genetic variation for thermotolerance in the Australian *Drosophila melanogaster* cline. *Molecular Ecology*, **16**, 2948–2957.

Reznick D (1985) Costs of reproduction: an evaluation of the empirical evidence. *Oikos*, **44**, 257–267.

Reznick D, Bryga H, Endler JA (1990) Experimentally induced life-history evolution in a natural population. *Nature*, **346**, 357–359.

Roff DA (1992) *The Evolution of Life Histories: Theory and Analysis*. Chapman and Hall, New York.

Rogina B, Reenan RA, Nilsen SP, Helfand SL (2000) Extended life-span conferred by cotransporter gene mutations in *Drosophila*. *Science*, **290**, 2137–2140.

Rose MR, Charlesworth B (1981) Genetics of life history in *Drosophila melanogaster*. II. Exploratory selection experiments. *Genetics*, **97**, 187–196.

Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.

Schemske DW, Bierzychudek P (2007) Spatial differentiation for flower color in the desert annual Linanthus parryae: was Wright right? *Evolution*, **61**, 2528–2543.

Schmidt PS, Conde DR (2006) Environmental heterogeneity and the maintenance of genetic variation for reproductive diapause in *Drosophila melanogaster*. *Evolution*, **60**, 1602–1611.

Schmidt PS, Paaby AB (2008) Reproductive diapause and life history clines in North American populations of *Drosophila melanogaster*. *Evolution*, **62**, 1204–1215.

Schmidt PS, Matzkin LM, Ippolito M, Eanes WF (2005a) Geographic variation in diapause incidence, life history traits, and climatic adaptation in *Drosophila melanogaster*. *Evolution*, **59**, 1721–1732.

Schmidt PS, Paaby AB, Heschel MS (2005b) Genetic variance for diapause expression and associated life histories in *Drosophila melanogaster*. *Evolution*, **59**, 2616–2625.

Schmidt PS, Zhu CT, Das J, Batavia M, Yang L, Eanes W (2008) An amino acid polymorphism in the couch potato gene forms the basis for climatic adaptation in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA*, **105**, 16207–16211.

Sezgin E, Duvernell DD, Matzkin LM *et al.* (2004) Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster*. *Genetics*, **168**, 923–931.

Stearns SC (1991) Trade-offs in life-history evolution. *Functional Ecology*, **3**, 259–268.

Stearns SC (1992) *The Evolution of Life Histories*. Oxford University Press, New York.

Suh Y, Atzmon G, Cho M-O *et al.* (2008) Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proceedings of the National Academy of Sciences, USA*, **105**, 3438–3442.

Tajima F (1989) Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tatar M, Promislow DEL, Khazaeli AA, Curtsinger JW (1996) Age-specific patterns of genetic variance in *Drosophila melanogaster* II. Fecundity and its genetic covariance with age-specific mortality. *Genetics*, **143**, 849–858.

Tatar M, Kopelman A, Epstein D, Tu M-P, Yin C-M, Garofalo RS (2001) A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science*, **292**, 107–110.

Tatar M, Bartke A, Antebi A (2003) The endocrine regulation of aging by insulin-like signals. *Science*, **299**, 1346–1351.

Toivonen JM, Partridge L (2009) Endocrine regulation of aging and reproduction in *Drosophila*. *Molecular and Cellular Endocrinology*, **299**, 39–50.

Trotta V, Calboli FCV, Ziosi M *et al.* (2006) Thermal plasticity in *Drosophila melanogaster*: a comparison of geographic populations. *BMC Evolutionary Biology*, **6**, 67–79.

Turner TL, Levine MT, Eckert ML, Begun DJ (2008) Genomic analysis of adaptive differentiation in *Drosophila melanogaster*. *Genetics*, **179**, 455–473.

Vermeulen CJ, Loeschcke L (2006) Longevity and the stress response in *Drosophila*. *Experimental Gerontology*, **42**, 153–159.

Weeks AR, McKechnie SW, Hoffmann AA (2002) Dissecting adaptive clinal variation: markers, inversions and size/stress associations in *Drosophila melanogaster* from a central field population. *Ecology Letters*, **5**, 756–763.

Willcox BJ, Donion TA, He Q *et al.* (2008) FOXO3A genotype is strongly associated with human longevity. *Proceedings of the National Academy of Sciences, USA*, **105**, 13987–13992.

Williams KD, Busto M, Suster ML *et al.* (2006) Natural variation in *Drosophila melanogaster* diapause due to the insulin-regulated PI3-kinase. *Proceedings of the National Academy of Sciences, USA*, **103**, 15911–15915.

Wilson RH, Morgan TJ, Mackay TFC (2006) High-resolution mapping of quantitative trait loci affecting increased life span in *Drosophila melanogaster*. *Genetics*, **173**, 1455–1463.

Annalise Paaby investigated the evolutionary genetics of *Drosophila* life history as a graduate student in Paul Schmidt's lab, and currently studies cryptic genetic variation in *C. elegans* at New York University. This paper comprises a chapter of her dissertation. Mark Blacket and Ary Hoffmann are researchers at the Centre for Environmental Stress and Adaptation Research (CESAR), which investigates how organisms, and in particular, insects, adapt to environmental stress. The research is pitched at the genetic, ecological, physiological and biochemical levels. The prime focus of the studies has been the clinal variation in heat and cold responses up and down the temperate-tropical coastal transect in eastern Australia, identification of genes linked to climatic adaptation, and the identification of species and populations with limited evolutionary adaptation potential. Paul Schmidt is broadly interested in evolutionary dynamics in heterogeneous environments, using *Drosophila* as a model genetic system and the North Atlantic rocky intertidal as a model ecological system.