Phenotypic plasticity explains violation of Dollo's law

Bertanne Visser[1*], Hans T. Alborn[2], Suzon Rondeaux[1], Manon Haillot[1], Thierry Hance[3], Darren Rebar[4], Jana M. Riederer[5], Stefano Tiso[5], Timo J.B. van Eldijk[5], Franz J. Weissing[5] & Caroline M. Nieberding[6]

[1]Evolution and Ecophysiology group, Biodiversity Research Centre, Earth and Life Institute, UCLouvain, Croix du Sud 4-5, 1348 Louvain-la-Neuve, Belgium.
[2]Chemistry Research Unit, Center of Medical, Agricultural, and Veterinary Entomology, Agricultural Research Service, United States Department of Agriculture, 1600 SW 23rd Drive, Gainesville, FL 32608, USA.
[3]Ecology of Interactions and Biological Control group, Biodiversity Research Centre, Earth and Life Institute, UCLouvain, Croix du Sud 4-5, 1348 Louvain-la-Neuve, Belgium.
[4]Department of Biological Sciences, Emporia State University, 1 Kellogg Circle, Campus Box 4050, Emporia, KS 66801, USA.
[5]Groningen Institute of Evolutionary Life Sciences, University of Groningen, Nijenborgh 7, 9747 AG Groningen, the Netherlands.
[6]Evolutionary Ecology and Genetics group, Biodiversity Research Centre, Earth and Life Institute, UCLouvain, Croix du Sud 4-5, 1348 Louvain-la-Neuve, Belgium.

*Correspondence to: bertanne.visser@uclouvain.be

Abstract
Dollo's law of irreversibility states that once a complex adaptation has been lost in evolution, it will not be regained. Recently, various violations of this principle have been described. Here, we argue that the logic underlying Dollo's law only applies to traits that are constitutively expressed, while it fails in case of 'plastic' traits that are up- or downregulated according to needs. We tested this hypothesis for an archetypal violation of Dollo's law, the loss and regain of fat synthesis in parasitic wasps. Wasps from lineages that supposedly had lost lipogenic ability more than 200 million years ago were grown under various conditions. In line with our hypothesis, it turned out that fat synthesis had not been lost but was only switched on in low-fat environments. Such plasticity cannot only explain supposed violations of Dollo's law, but also the maintenance of adaptations to rarely occurring extreme events.

Introduction
Universal laws are rare or absent in ecology and evolution, but Dollo's principle of the irreversibility of evolution seems to come close to such a law[1,2]. In the numerous cases where a complex adaptation has been lost in the evolutionary history of a lineage[3], there are only a handful examples where the adaptation has later been regained[4,5]. Snakes did not regain legs, birds did not regain teeth, and ratites did not regain the ability to fly. Complex adaptations seem to be highly vulnerable: if they are of no use for an extended period of time, they can be expected to be selected against and/or to decay by genetic drift and the accumulation of deleterious mutations[6]; if they are later needed again, they cannot easily re-evolve[7], since the build-up of a complex trait generally requires the co-evolution of whole sets of genes. If this logic is indeed correct, Dollo's principle may have important implications. In times of global climate change, previously rare and extreme events (like periods of drought or flooding) are expected to occur much more frequently. How can organisms cope with such events, if their adaptations to such events were irreversibly lost in the intervening period between such events?

The line of argumentation behind Dollo's law sounds plausible for constitutively expressed organs like legs or teeth. It is not surprising that cave fish lose their eyes when living in complete darkness over many generations, and that new eyes do not easily evolve once they change to a habitat where vision is an advantage. However, for three reasons the situation is different for adaptations that are phenotypically plastic (like most physiological and behavioural traits) and expressed according to local environmental conditions. First, such adaptations are less costly to maintain than constitutively expressed traits, and the underlying genetic architecture may be down- or upregulated as needed. Second, novel phenotypes may be expressed, and thus selected upon, when individuals experience non-standard conditions. Third, regulatory pathways underlying these adaptations tend to combine the properties of robustness and evolvability, meaning that they do not easily decay when subjected to the accumulation of mutations and that only few mutational steps are required for restoring their function when conditions change[8]. Therefore, we hypothesize that *(i)* phenotypically plastic traits are much less affected by evolutionary decay and irreversible loss than constitutively expressed traits, and that *(ii)* many supposed violations of Dollo's law are unrecognized instances of phenotypic plasticity.

In the light of these considerations, we scrutinize an archetypal violation of Dollo's law, the apparent loss and regain of an essential metabolic trait: the synthesis of fat[9,10]. Fat is synthesized when a surplus of sugars and other carbohydrates is available in the diet[11], providing a reserve for

future use. How much fat is synthesized at any time largely depends on the fat content of the diet, e.g., parasitic worms will slow down the rate of *de novo* fatty acid synthesis when fat can be scavenged from a host[12]. Fat is critical for survival and reproduction in nearly all living organisms. The importance of fat and the ability to synthesize fat *de novo* explains why underlying metabolic and genetic pathways for fat synthesis are typically highly conserved, from bacteria to humans[13–16]. Therefore, it came as a surprise that the ability to synthesize fat for storage appeared to have been repeatedly lost and regained in parasitic insects[10]. Parasitic wasps lost the ability to synthesize fat more than 200 million years ago[17] and the trait re-appeared in several lineages, including in the genus *Leptopilina*[10,18] (about 80 million years ago). Replicated experiments with different populations of two *Leptopilina* species then revealed that some populations synthesized fat, while others did not[10,18–21]. Recently, similar results were also found for the wasp *Nasonia vitripennis*[22].

We hypothesized that these findings do not reflect the constitutive loss and regain of fat synthesis due to mutational changes in the metabolic pathway, but rather extreme plastic expression (on or completely off) of fat synthesis in response to the local environment. Wasp development occurs in or on a host insect[23]; hence stored fat of the host can be carried over directly by the wasp. If the host contains plenty of fat, there is no need for *de novo* fatty acid and triglyceride synthesis and the pathway should be completely shut off. If, in contrast, a fat-poor host is encountered, the wasp has to synthesize additional fat itself by activating the pathway. It is, therefore, conceivable that fat synthesis became plastic when the wasps started to parasitize fat-rich hosts (more than 200 MA) and that it was switched off, except when the wasps encountered fat-poor hosts.

Results and Discussion

The question arises whether a switching device that is not used for extensive periods of time (more than 100 million years) should not be lost during the course of evolution. To investigate this, we ran individual-based simulations that monitored the sustained functionality of a switching device (a gene regulatory network that could decay by mutation) that is only sporadically used in evolutionary time (Material and methods Section 1). Figure 1 shows that the switching device rapidly disintegrates (red simulations) if it is never used. However, even very infrequent use (pink: every 100 generations; purple: every 1000 generations) suffices to keep the switching device largely intact. Interestingly, the switching device does not erode gradually, but instead slowly evolves an improved performance over evolutionary time (i.e., the percentage of correct decisions increases with the increasing number of generations). An inspection of the evolving gene regulatory networks reveals that they become more and more robust (i.e., less and less affected by mutational decay), in line with earlier findings on network evolution[8].

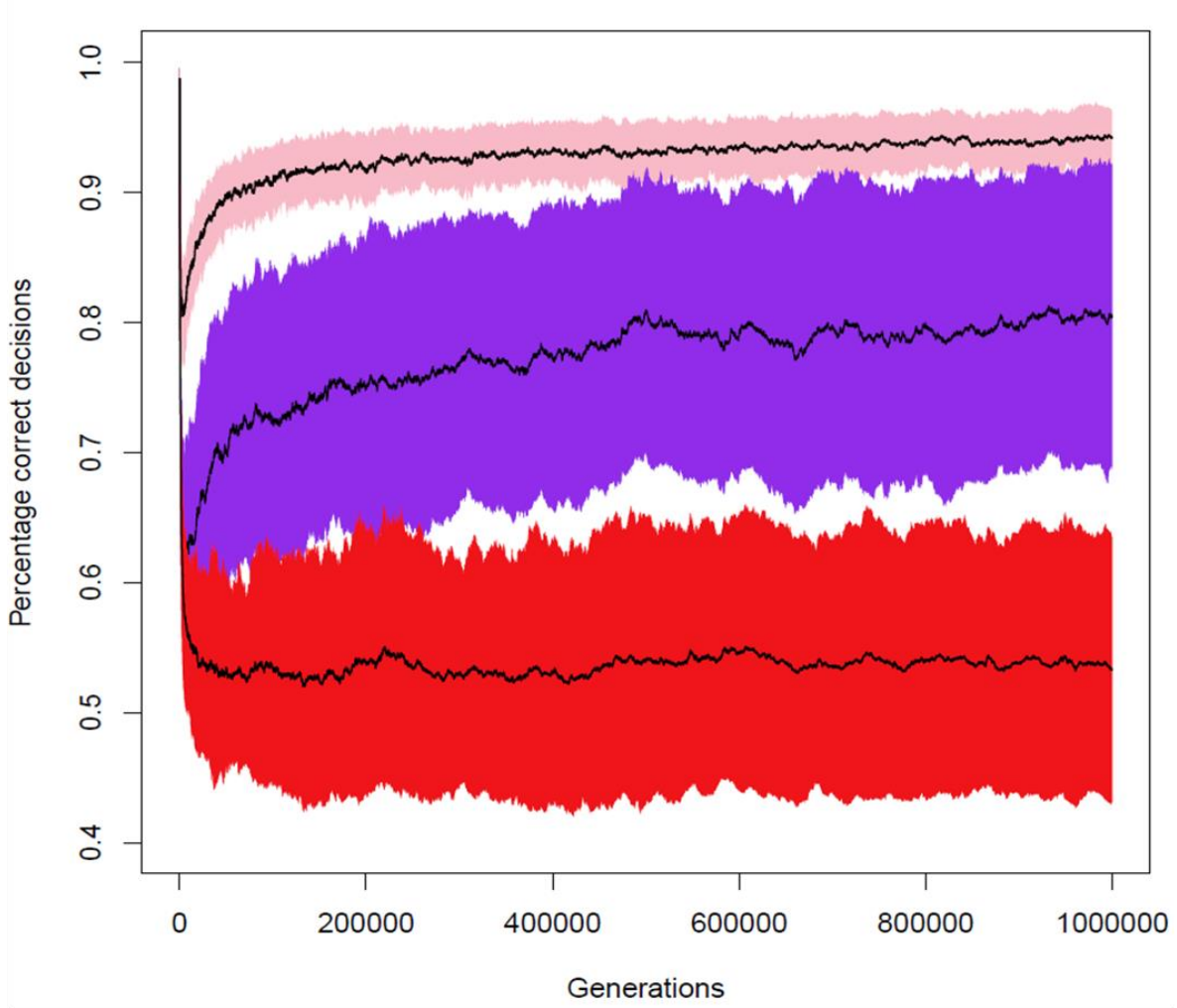


Fig. 1: Sporadic activation is sufficient for the maintenance of adaptive plasticity. Long-term individual-based simulations showing how the performance of a gene-regulatory network (GRN) underlying adaptive plasticity changes in time when plasticity is only sporadically activated. We first evolved replicate GRNs in a variable environment where it is adaptive to switch on a metabolic pathway (fat synthesis) under low-fat conditions and to switch it off under high-fat conditions. In generation 0, a monomorphic population was established, where all N=10,000 individuals were endowed with the same well-performing GRN (different across replicates). Subsequently, the population evolved subject to selection, mutation (μ=0.001 per gene locus) and genetic drift in a fat-rich environment, where it is adaptive to constitutively switch off the metabolic pathway. Every 100 generations, we monitored the performance of a representative sample of GRNs (percentage correct decisions) in the original (fat-variable) environment: 1.0 means that the GRN is still making 100% adaptive decisions; 0.5 means that the GRN only makes 50% adaptive decision, as would be expected by a random GRN or a GRN that switches the pathway constitutively on or off. The coloured graphs show the average performance (± standard deviation) of the GRNs for three scenarios (100 replicates per simulation). Red: the population never again encounters the fat-variable environment, leading to the loss of adaptive plasticity (convergence to 0.5). Pink: the individuals encounter a fat-variable environment on average every 100 generations; after an initial rapid drop in performance, a sustained high performance (>90% correct decisions) of the GRNs is regained after about 100,000 generations. Purple: the individuals encounter a fat-variable environment on average every 1000 generations; after an initial rapid drop in performance, an intermediate performance (>75% correct decisions) is regained gradually.

From the simulations, we conclude that phenotypic plasticity (i.e., switching on or off of a metabolic pathway) can be maintained over long evolutionary time periods, even if plasticity is only needed sporadically. We, therefore, tested whether wasps that seemingly had lost the ability for fat synthesis were still able to synthesize fat in a low-fat environment. To this end, we let females from four field-caught populations of *L. heterotoma* develop on two naturally co-occurring host species: low-fat ("lean") *Drosophila simulans* and high-fat ("fat") *D. melanogaster* (containing 63 ± 3 µg and 91 ± 4 µg, mean ± 1SE storage fat, respectively; $F_{1,17} = 35.95$; $p < 0.0001$; Material and methods Section 2). When developing on the low-fat *D. simulans* host, the fat content of the wasps was in three of the four populations significantly higher after feeding than at emergence, indicating that fat synthesis had indeed occurred (Table 1). In contrast, the wasps showed only a marginal (and non-significant) increase when developing on the fatter *D. melanogaster*. These data suggest that fat synthesis does indeed depend on the host environment, at least in some wasp populations.

Table 1: Wasps supposedly having lost lipogenic ability synthesize fat in a fat-poor environment. Mean absolute fat amount ± 1se (in µg) was quantified in adult wasps from field-caught *L. heterotoma* populations raised on two hosts (fat-poor *D. simulans*, left part of the table; fat-rich *D. melanogaster*, right part of the table) and at two developmental stages (Emerged: just after emergence; Fed: having fed for 7 days after emergence). P-values reveal whether 7 days of feeding led to a significant increase in fat content, indicating the occurrence of fat synthesis. Three of the four populations tested on *D. simulans* exhibited fat synthesis on the lean host but no fat synthesis on the fat host, that is, plasticity in fat synthesis. (*) T-tests were performed when data was normally distributed and variances equal with (^) or without log transformation. The non-parametric Mann-Whitney U test was used for non-normal data or data with unequal variances (*).

	Development on *D. simulans*				Development on *D. melanogaster*			
Population	Sample size	Emerged	Fed	p-value	Sample size	Emerged	Fed	p-value
Belgium 1**	10	17.50 ± 6.84	43.50 ± 11.38	**0.043**	38	36.00 ± 2.54	40.00 ± 3.22	0.336
Belgium 2	38	15.60 ± 1.02	36.83 ± 4.86	**<0.001 (*)**	32	38.50 ± 4.24	43.91 ± 3.45	0.331
UK 1	21	24.20 ± 1.49	30.91 ± 3.22	0.142 (^)	29	40.00 ± 4.41	44.30 ± 2.09	0.375
UK 2	-	-	-	-	17	33.60 ± 3.82	39.50 ± 1.50	0.522
Japan	20	12.20 ± 1.55	24.80 ± 4.86	**0.011 (^)**	13	29.17 ± 6.27	28.67 ± 8.19	0.964

The population-level comparison of wasp fat content at two points in time is only a crude measure that not always detects the occurrence of fat synthesis reliably. Even in case of active fat synthesis, fat content can stay constant or even decrease if, for example, fats are burned at a faster rate than at which they are produced[24]. To unequivocally demonstrate that fat synthesis can be induced plastically, we turned to stable isotope tracing followed by GC-MS (Gas Chromatography-Mass Spectrometry) analyses[25,26] (Material and methods Section 2). Incorporation of stable isotopes after feeding depends on fat synthesis; hence a significant increase in stable isotope levels compared to controls (without access to stable isotopes) demonstrates active fat synthesis, even when lipids are burned. We used a split-brood family design where daughters of a single mother were allowed to develop on either lean *D. simulans* or fat *D. melanogaster*. Seventeen families, belonging to five field-caught populations, showed a (much) higher fat metabolism in the fat-poor environment (*D.*

simulans) than in the fat-rich environment (*D. melanogaster*) (Figure 2; Supplementary Table 1). These results confirm that fat synthesis is indeed a plastic trait that is induced in response to low host fat content. Notice that the 17 families strongly differ in their environmental response, both in their baseline level of fat synthesis (on fat *D. melanogaster*) and in the slopes of their reaction norms. We conclude that fat synthesis is plastic and that there is (across-family) genetic variation in the degree of plasticity.

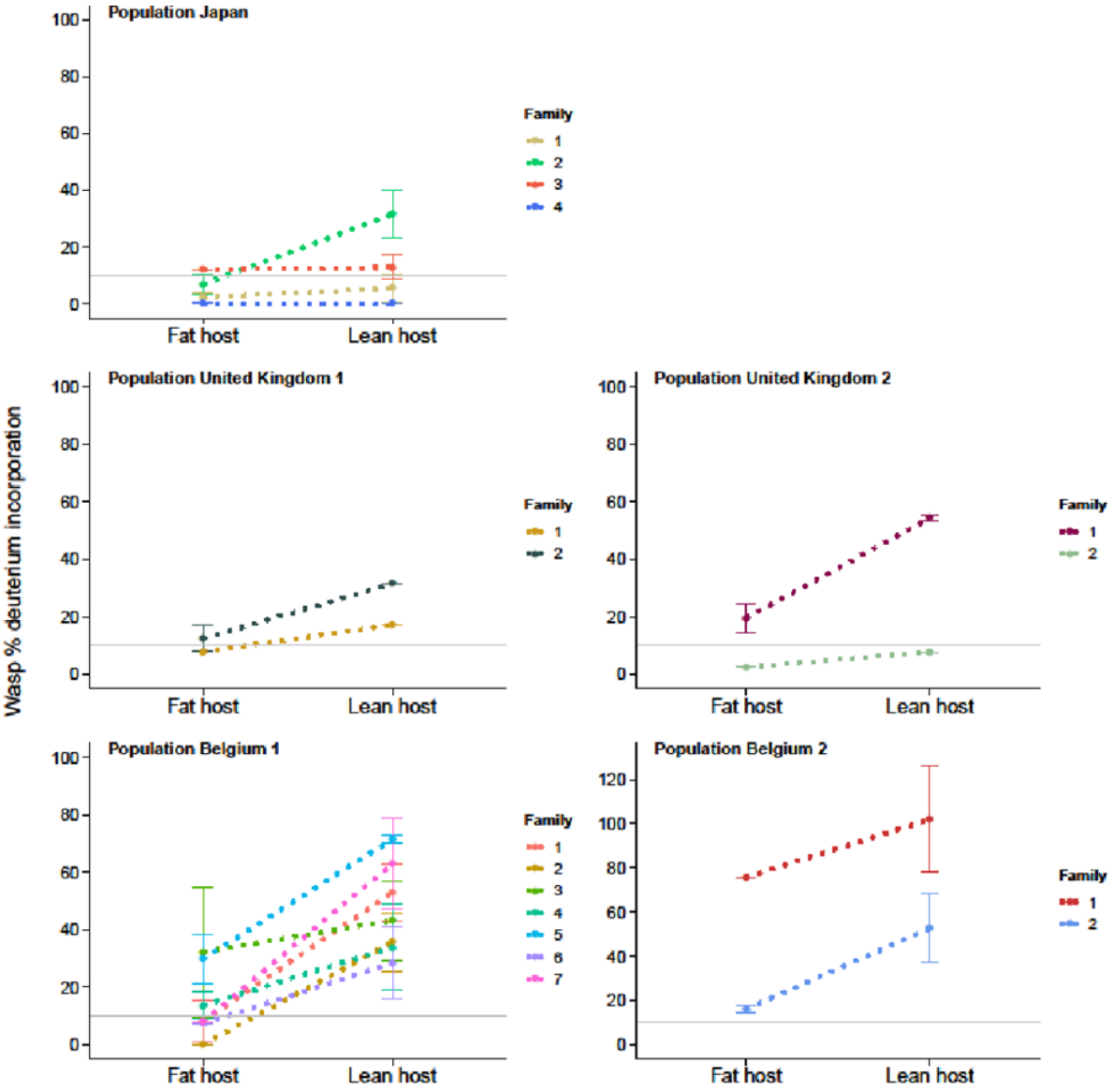
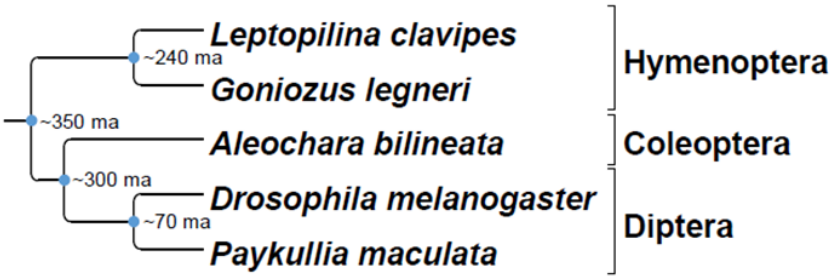


Fig. 2: Phenotypic plasticity in five field-caught wasp populations. Incorporation of stable isotopes into the fatty acid fraction of offspring from 17 families developing in a fat-rich environment (fat host *Drosophila melanogaster*, left in each graph) and in a fat-poor environment (lean host *D. simulans*, right in each graph; n = 138). The horizontal gray line indicates that a stable isotope incorporation below 10% is considered insufficient evidence for the occurrence of fat synthesis.

To rule out the possibility that the above results reflect a host-species effect, rather than an effect of the fat content of the environment, we repeated the experiment reported in Table 1, but now replacing lean *D. simulans* by lean *D. melanogaster* hosts. By reducing the sugar content in the diet of *D. melanogaster*, we were able to generate leaner flies (i.e. pupae containing 52 ± 3 µg storage lipids, mean ± 1SE, compared to 91 ± 4 µg storage lipids, mean ± 1SE; $F_{1,22} = 71.18$, $p < 0.0001$). In line with our earlier results, three out of four wasp populations indeed showed fat synthesis on these leaner *D. melanogaster* hosts (Supplementary Table 2). We conclude that plastic fat synthesis is induced by host fat content, rather than other traits differing between *D. melanogaster* and *D. simulans.*

The ability to synthesize fat when being placed in a non-standard (low-fat) environment indicates that key genes for fat synthesis have not lost their functionality in the *Leptopilina* genus. Making use of the fact that the genetic molecular pathway underlying fatty acid synthesis is highly conserved across animal taxa[13–16], we conducted a comparative analysis of coding sequences of acetyl coenzyme A carboxylase (ACC)[27] and fatty acid synthase (FAS)[15], two enzymes that are critical for the production of fatty acids, the raw materials for storage fat. We used the *acc* and *fas* gene coding sequences of *D. melanogaster* as a starting point, because this fly readily synthesizes fat. Similar gene sequences were indeed found in the genome of *L. clavipes,* a sister species of *L. heterotoma,* and all functional domains of ACC and FAS enzymes were recovered, suggesting fully functional coding sequences in the *L. clavipes* genome (Figure 3). We then expanded our search for *acc* and *fas* functional coding sequences and protein domains to more distantly related parasitoids presumed to have lost fat synthesis independently[10]: the hymenopteran *Goniozus legneri* (family Bethylidae), the dipteran *Paykullia maculata* (family Rhinophoridae), and the coleopteran *Aleochara bilineata* (family Staphilinidae)(Figure 3). ACC and FAS amino acid sequences of all these species aligned (Supplementary Texts 1 and 2), suggesting that these two critical genes for fat synthesis have been conserved throughout the repeated evolution of parasitism in insects.

a



b Acetyl CoA carboxylase

		Leptopilina clavipes scf7180005162486		*Goniozus legneri* scaffold2431		*Aleochara bilineata* scaffold525		*Drosophila melanogaster* NM_136498.3		*Paykullia maculata* scaffold367327	
Domain	Accession	Interval	E-value	Interval	E-value	Interval	E-value	Interval	E-value	Interval	E-value
BC	COG0439	958-2451	5.31E-144	631-2124	2.65E-141	337-1812	3.05E-139	967-2460	1.24E-138	103-1596	1.71E-135
BCCP	cd06850	2851-3084	5.68E-12	2524-2718	4.63E-16	2230-2424	1.28E-13	2860-3045	2.70E-16	1996-2181	1.32E-16
ACCC	pfam08326	3085-5244	0	2719-4874	0	2425-4629	0	3055-5268	0	2191-4401	0
CT	pfam01039	5578-7191	2.44E-159	5221-6852	9.34E-155	4954-6663	3.65E-164	5587-7221	1.53E-160	4720-6138	2.54E-156

c Fatty acid synthase

		Leptopilina clavipes scf7180005162060		*Goniozus legneri* scaffold5		*Aleochara bilineata* scaffold43971		*Drosophila melanogaster* FBtr0077659		*Paykullia maculata* scaffold4856	
Domain	Accession	Interval	E-value	Interval	E-value	Interval	E-value	Interval	E-value	Interval	E-value
KS	cd00833	381-1589	6.28E-164	85-1293	1.79E-166	52-1260	7.01E-170	742-1956	9.12E-163	412-1623	9.69E-165
AT	smart00827	1857-2708	1.48E-60	1561-2412	1.86E-64	1525-2376	5.65E-61	2227-3084	2.43E-61	1894- 2775	3.47E-62
DH	pfam14765	2925-3392	1.53E-15	2629-3225	1.23E-15	2626-3327	2.37E-09	3292-4026	4.16E-20	2959-3714	9.21E-15
ER	smart00829	4704-5576	2.79E-127	4408-5280	2.40E-130	4339-5211	2.12E-130	5101-5973	9.57E-126	4768-5640	3.58E-129
KR	cd08954	5637-6368	2.02E-79	5341-6051	9.53E-80	5266-5982	6.33E-77	6025-6747	1.24E-90	5692-6414	1.20E-87
ACP	smart00823	6348-6566	3.89E-07	6091-6249	1.55E-08	6019-6177	4.17E-07	6784-6942	4.39E-08	6451-6609	2.06E-07
TE	pfam00975	6762-7046	9.69E-15	6412-6699	5.41E-17	6346-6783	1.40E-19	7180-7497	4.82E-12	6814-7161	1.76E-17

Fig. 3: Conservation of two genes crucial for fatty acid synthesis in four parasitoid insects that supposedly had lost lipogenic activity. Long evolutionary divergence times (up to 350 MA) separate the insect *Drosophila melanogaster* (that synthesizes lipids constitutively) and 4 parasitoid insects that were assumed to have lost the ability to synthesize lipids(10) (A). Acetyl coenzyme A carboxylase (ACC) and fatty acid synthase (FAS) are two essential genes for the production of fatty acids: the presence of all domains of ACC (B) and FAS (C) genes from *D. melanogaster* in the four parasitoid genomes reveals that the functional activity of the two genes is conserved in these insects. A table containing the detailed length and position of the different functional domains forming the two genes, as well as conservation level of the nucleotide sequence of the domains (e-values; the lower the e-value, the higher the significance of the match) are shown for each species. Abbreviations: BC = Biotin carboxylase; BCCP = Biotin carboxyl carrier protein; ACCC = Acetyl-coA carboxylase central region; CT = Carboxyl transferase domain; KS = Ketoacyl synthase; AT = Acyl transferase; DH = Dehydratase; ER = Enoyl reductase; KR = Ketoacyl reductase; ACP = Acyl carrier protein; TE = Thioesterase. Accession numbers refer to the conserved domain identifier on NCBI's Conserved Domain Database. Parasitoid transcript identifiers are provided underneath each species name.

Summarizing, we have provided compelling evidence that parasitoid wasps have not lost their ability to synthesize storage fat in the distant past (>200 million years ago), but that fat synthesis is a plastic trait that can be switched off when the wasps are developing in a fat-rich environment. In other animals, fat synthesis is a constitutively expressed trait, but the rate of fat synthesis depends on the nutrient content of the diet. The ability of wasps to switch off fat synthesis completely, despite continued feeding on sugars, is unique and exceptional, and we are unaware of a similar finding in other animals. A crucial pathway like fat metabolism is thus not constitutively expressed in parasitoid wasps, but activated or deactivated in response to environmental conditions. This makes perfect sense, since they typically develop on fat-rich hosts that provide all the storage fat needed by the wasps. Yet, plasticity is required since there is considerable spatio-temporal variation in host availability and quality. *L. heterotoma* is a generalist wasp that can parasitize more than ten different *Drosophila* species that differ substantially in size and fat availability[28]. Moreover, there is considerable geographic and seasonal variation in host species diversity and community composition[29]. Hosts are further patchily distributed with overlapping generations, suggesting considerable spatial variation at a local scale[28]. *Drosophila* are further well known to show large variation in starvation resistance, which is typically correlated with fat content[30]. Plasticity in wasp fat synthesis is thus likely adaptive and evolved in response to highly variable environmental conditions in host fat content.

Previous documented cases of trait regain over long evolutionary time, in addition to the regaining of fat synthesis in parasitoids[10], include the regaining of wings in stick insects[31], the evolution of sexual reproduction from asexuality in mites[32], among other examples[33]. These cases were all based on comparative analyses, which was shown to be problematic, because phylogenies do not necessarily provide a reliable representation of trait evolution[34–37]. Our results provide the first experimental evidence that macro-evolutionary patterns of trait reversals may in fact reflect trait plasticity: the trait is not "lost" or "regained" but is rather switched off or on, depending on environmental conditions. Intriguingly, such a regulatory switch can remain largely intact, even if it is only sporadically activated (Figure 1). We consider it plausible that our findings are not restricted to fat metabolism in parasitoid wasps: the plastic regulation of trait expression could explain more cases of apparent trait loss and reappearance at macro-evolutionary time scales. Wing formation, for example, is often observed as an atavism (the sporadic occurrence of an ancestral phenotype) in otherwise wingless insects[38], and wing polymorphism, i.e. plasticity in wing development, is common in insects in general[39]. Similarly, many asexual populations sporadically produce sexually reproducing individuals and plasticity in reproductive mode has evolved in several insect systems[40–42]. Hence, plasticity may be a common principle explaining apparent violations of Dollo's law. As indicated by our simulation study, plasticity can also explain the puzzling fact that adaptations to rare and extreme events are not lost, even if they are only sporadically used.

Materials and Methods

Section 1: Modelling study

We consider the general situation where phenotypic plasticity is only sporadically adaptive and ask the question whether and under what circumstances plasticity can remain functional over long evolutionary time periods when the regulatory processes underlying plasticity are gradually broken down by mutations. To fix ideas, we consider a regulatory mechanism that switches on or off a pathway (like fat synthesis) in response to environmental conditions (e.g. host fat content).

Fitness considerations

We assume that the local environment of an individual is characterized by two factors: fat content *F* and nutrient content *N*, where nutrients represent sugars and other carbohydrates that can be used to synthesize fat. Nutrients are measured in units corresponding to the amount of fat that can be synthesized from them. We assume that fitness (viability and/or fecundity) is directly proportional to the amount of fat stored by the individual. When fat synthesis is switched off, this amount is equal to *F*, the amount of fat in the environment. When fat synthesis is switched on, the amount of fat stored is assumed to be $N - c + (1-k)F$. This expression reflects the following assumptions: *(i)* fat is synthesized from the available nutrients, but this comes at a fitness cost *c*; *(ii)* fat can still be absorbed from the environment, but at a reduced rate (1-*k*). It is adaptive to switch on fat synthesis if $N - c + (1-k)F$ is larger than *F*, or equivalently if $F < \frac{1}{k}(N-c)$.

The right-hand side of this inequality is a straight line, which is illustrated by the blue line in Fig. 4. The three boxes in Fig. 4 illustrate three types of environmental conditions.

- Red box: low-fat environments. Here, $F < \frac{1}{k}(N-c)$ is always satisfied, implying that fat synthesis should be switched on constitutively.
- Yellow box: high-fat environments. Here, $F > \frac{1}{k}(N-c)$, implying that fat synthesis should be switched off constitutively.
- Orange box: intermediate-fat environments. Here, fat synthesis should be plastic and switched on if for the given environment (*N*, *F*) the fat content is below the blue line and switched off otherwise.

The simulations reported here were all run for the parameters $k = \frac{1}{2}$ and $c = \frac{1}{4}$.

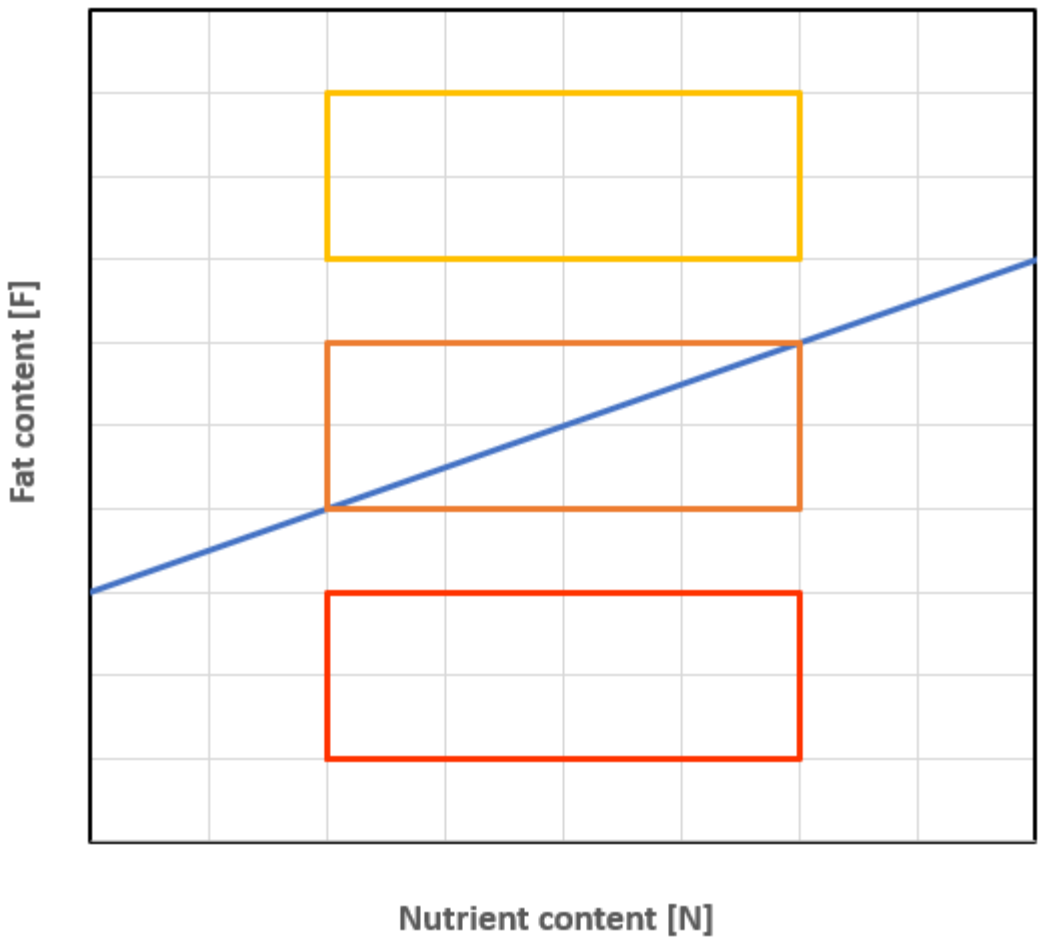


Fig. 4: Environmental conditions encountered by the model organisms. For a given combination of environmental nutrient content *N* and environmental fat content *F*, it is adaptive to switch on fat synthesis if (*N*, *F*) is below the blue line (corresponding to $F < \frac{1}{k}(N - c)$) and to switch it off otherwise. The three boxes illustrate three types of environment: a low-fat environment (red) where fat synthesis should be switched on constitutively; a high-fat environment (yellow) where fat synthesis should be switched off constitutively; and an intermediate-fat environment (orange) where a plastic switch is selectively favoured.

Gene regulatory networks (GRN)

In our model, the switching device was implemented by an evolving gene regulatory network (as in (43)). The simulations shown in Fig. 1 of the main text are based on the simplest possible network that consists of two receptor nodes (sensing the fat and the nutrient content in the local environment, respectively) and an effector node that switches on fat synthesis if the combined weighted input of the two receptor nodes exceeds a threshold value *T* and switches it off otherwise. Hence, fat synthesis is switched on if $w_F F + w_N N > T$ (and off otherwise), where the weighing factors w_F and w_N and the threshold *T* are genetically determined evolvable parameters. We considered many alternative network structures (all with two receptor nodes and one effector node) and obtained very similar results (see below).

For the simple GRN described above, the switching device is 100% adaptive when the switch is on (i.e., $w_F F + w_N N > T$) if $F < \frac{1}{k}(N - c)$ and off otherwise. A simple calculation yields that this is the case if: $w_N > 0,\ w_F = -kw_N$ and $T = cw_N$.

Evolution of the GRN

For simplicity, we consider an asexual haploid population with discrete, non-overlapping generations and fixed population size $N = 10{,}000$. Each individual has several gene loci, each locus encoding one parameter of the GRN. (In case of the simple network described above, there are three gene loci, which encode the parameters w_F, w_N and T). At the start of its life, each individual is placed in a randomly chosen environment (*N*, *T*). Based on its (genetically encoded) GRN, the individual decides on whether to switch on or off fat synthesis. If synthesis is switched on, the individual's fitness is given by $N - c + (1-k)F$; otherwise its fitness is given by *F*. Subsequently, the individuals produce offspring, where the number of offspring per individual is proportional to the individual's fitness. Each offspring inherits the genetic parameters of its parents, subject to mutation. With probability μ (per locus) a mutation occurs. In such a case the parental value is changed by a certain amount; the mutational step size is drawn from a normal distribution with mean zero and standard deviation σ. In the reported simulations, we chose $\mu = 0.001$ and $\sigma = 0.1$.

Preadaptation of the GRNs

Starting with randomly initialized population, we first let the population evolve in the intermediate-fat environment (orange box in Fig. 4) for 10,000 generations. In all replicate simulations, "perfectly adapted switch" (corresponding to $w_N > 0,\ w_F = -kw_N$ and $T = cw_N$) evolved, typically within 1,000 generations. These evolved networks were used to seed the populations in the subsequent "decay" simulations.

Evolutionary decay of the GRNs

For the decay experiments reported in Fig. 1 of the main text, we initiated a large number of monomorphic replicate populations with one of the perfectly adapted GRNs from the preadaptation phase. These populations were exposed for an extended period of time (1,000,000 generations) to a high-fat environment (yellow box in Fig. 4), where all GRNs switched off fat synthesis constitutively. However, in some scenarios, the environmental conditions changed back sporadically (with probability *q*) to the intermediate-fat environment, where it is adaptive to switch on fat metabolism in half of the environments. In Fig. 1, we report on the changing rates $q = 0.0$ (no changing back; red), $q = 0.001$ (changing back once every 1,000 generations; purple), and $q = 0.01$ (changing back once every 100 generations; pink). When such a change occurred, the population was exposed to the intermediate-fat environment for *t* generations (Fig. 1 is based on *t* = 3).

Throughout the simulation, the performance of the network was monitored every 100 generations as follows: 100 GRNs were chosen at random from the population, and each of these GRNs was

exposed to 100 randomly chosen environmental conditions from the intermediate-fat environment. From this, we could determine the average percentage of "correct" decisions (where the network should be switched on if and only if $F < \frac{1}{k}(N-c)$. 1.0 means that the GRN is still making 100% adaptive decisions; 0.5 means that the GRN only makes 50% adaptive decision, as would be expected by a random GRN or a GRN that switches the pathway constitutively on or off. This measure for performance in the "old" intermediate-fat environment was determined for 100 replicate simulations per scenario and plotted in Fig. 1 (mean ± standard deviation).

Evolving robustness of the GRNs

The simulations in Fig. 1 are representative for all networks and parameters considered. Whenever $q = 0.0$, the performance of the regulatory switch eroded in evolutionary time, but typically at a much lower rate in case of the more complex GRNs. Whenever $q = 0.01$, the performance of the switch went back to levels above 90% and even above 95% for the more complex GRNs. Even for $q = 0.001$, a sustained performance level above 75% was obtained in all cases.

Intriguingly, in the last two scenarios the performance level first drops rapidly (from 1.0 to a much lower level, although this drop is less pronounced in the more complex GRNs) and subsequently recovers to reach high levels again. Apparently, the GRNs have evolved a higher level of robustness, a property that seems to be typical for evolving networks[8]. For the simple GRN studied in Fig. 1, this outcome can be explained as follows. The initial network was characterized by the genetic parameters $w_N > 0,\ w_F = -kw_N$ and $T = cw_N$ (see above), where w_N was typically a small positive number. In the course of evolutionary time, the relation between the three evolving parameters remained approximately the same, but w_N (and with it the other parameters) evolved to much larger values. This automatically resulted in an increasingly robust network, since mutations with a given step size distribution affect the performance of a network much less when the corresponding parameter is large in absolute value.

Section 2: Experimental study

Insects

Hosts and parasitoids were maintained as previously described[21]. Five *Leptopilina heterotoma* (Hymenoptera: Figitidae) populations were used for experiments: a population from Japan (Sapporo), two populations from the United Kingdom (1: Whittlesworth; 2: Great Shelford) and two populations from Belgium (1: Wilsele; 2: Eupen). Information on collection sites, including GPS coordinates, can be found in[21].

Determination of host fat content

D. simulans and *D. melanogaster* hosts were allowed to lay eggs over 24 hours in glass flasks containing ~50mL standard medium[21]. After two days, developing larvae were sieved and ~200 were larvae placed in a *Drosophila* tube (ø x h 25x95; Dominique Dutscher) containing ~10mL medium. Seven days after egg laying, newly formed pupae were frozen at -18˚C, after which fat content was determined as described in[21], where dry weight before and after neutral fat extraction was used to calculate absolute fat amount (in µg) for each host. The host pupal stage was chosen for estimating fat content, because at this point the host ceases to feed, while the parasitoid starts

consuming the entire host[28]. All data were analysed using R Project version 3.4.3(46). Fat content of hosts was compared using a one-way ANOVA with host species as fixed factor.

Manipulating host fat content

To generate leaner *D. melanogaster* hosts, we adapted our standard food medium[21] to contain 100 times less (0.5g) sugar per litre water. Manipulating sugar content did not alter the structure of the food medium, thus maintaining similar rearing conditions, with the exception of sugar content. Fat content of leaner and fatter *D. melanogaster* hosts was determined and analysed as described above.

Fat synthesis quantification with wasp populations

Mated female *L. heterotoma* were allowed to lay eggs on host fly larvae collected as described above with *ad libitum* access to honey as a food source until death. Honey consists of sugars and other carbohydrates that readily induce fat synthesis. After three weeks, adult offspring emergence was monitored daily and females were haphazardly placed in experimental treatments: emergence or feeding for 7 days on honey. Wasps were frozen at -18˚C after completion of experiments. Fat content was determined as described above. The ability for fat synthesis was then determined by comparing fat levels of recently emerged and fed individuals, similar to procedures described in[10,21,25]. An increase in fat levels after feeding is indicative of active fat synthesis; equal or lower fat levels suggest fat synthesis did not take place. Each population was analysed separately[21] for each host species using one-way ANOVAs with treatment as a fixed factor.

Fat synthesis quantification using a familial design and GC-MS analyses

To tease apart the effect of wasp genotype and host environment, we used a split-brood design where the offspring of each mother developed on lean *D. simulans* or fat *D. melanogaster* hosts. Two experiments were performed, one in which mothers were reared on *D. melanogaster* (experiment 1) and one in which mothers were reared on *D. simulans* (experiment 2). In both experiments, mothers were allowed to lay eggs in ~200 2nd to 3rd instar host larvae of one species for four days, after which ~200 host larvae of the other species were offered during four days. The order in which host larvae were presented was randomized across families. Following offspring emergence, daughters were allocated into two treatment groups: a control where females were fed a mixture of honey and water (1:2 m/m) or a treatment group fed a mixture of honey and deuterated water (Sigma Aldrich)(1:2 m/m; stable isotope treatment) for 7 days. Samples were prepared for GC-MS as described[26]. Incorporation of up to three deuterium atoms can be detected, but percent incorporation is highest when only 1 deuterium atom is incorporated. As incorporation of a single atom unequivocally demonstrates active fat synthesis, we only analysed percent incorporation (in relation to the parent ion) for the abundance of the m+1 ion. Percent incorporation was determined for five fatty acids, C16:1 (palmitoleic acid), C16:0 (palmitate), C18:2 (linoleic acid), C18:1 (oleic acid), and C18:0 (stearic acid), and the internal standard C17:0 (margaric acid). Average percent incorporation for C17:0 was 19.4 (i.e. baseline incorporation of naturally occurring deuterium) and all values of the internal standard remained within 3 standard deviations of the mean (i.e. 1.6). Percent incorporation of control samples was subtracted from treatment sample values to correct for background levels of deuterium (i.e. only when more deuterium is incorporated in treatment compared to controls fat is actively being synthesized). For statistical analyses, percent incorporation was first summed for C16:1, C16:0, C18:2, C18:1 and C18:0 to obtain overall incorporation levels, as saturated C16 and C18 fatty acids are direct products of the fatty acid synthesis pathway (that can subsequently be desaturated).

Identification of functional *acc* and *fas* genes in distinct parasitoid species

To obtain *acc* and *fas* nucleotide sequences for *L. clavipes, G. legneri, P. maculata* and *A. bilineata*, we used *D. melanogaster* mRNA ACC transcript variant A (NM_136498.3 in Genbank) and FASN1-RA (FBtr0077659 in FlyBase) and blasted both sequences against transcripts of each parasitoid (using the blast function available at parasitoids.labs.vu.nl[43,44]. Each nucleotide sequence was then entered in the NCBI Conserved Domain database[45] to determine the presence of all functional protein domains. All sequences were then translated using the Expasy translate tool (https://web.expasy.org/translate/), where the largest open reading frame was selected for further use. Protein sequences were then aligned using MAFFT v. 7 to compare functional amino acid sequences between all species[46].

Acknowledgements

We are grateful to Jerry Coyne, Frietson Galis, Georges Lognay, Philippe Vernon, Claude Remacle, René Rezsohazy, David Denlinger, Majeed Askari and Jérome Casas for helpful suggestions on an earlier draft of this manuscript. We would like to thank Christophe Pels for maintaining hosts and parasitoids. This work was supported by the Fonds de la Recherche Scientifique - FNRS under grant n˚ 24905063 and 29109376. This is publication BRC 427 of the Biodiversity Research Centre.

Conflict of interest disclosure

The authors of this preprint declare that they have no financial conflict of interest with the content of this article. Bertanne Visser and Caroline M. Nieberding are recommenders of the Peer Community in Evolutionary Biology, Bertanne Visser is a founder of the Peer Community in Zoology, Caroline M. Nieberding is in the managing board of the Peer Community In.

Data and materials availability

All data are available on
https://visserlab.be/download/visser-et-al-2020-pci-evol-biol.zip

References

1. Dollo, L. *Bull. la Société belge géologie, paléontologie d'hydrologie* **7**, 164–166 (1893).
2. Gould, S. J. *J. Hist. Biol.* **3**, 189–212 (1970).
3. Ellers, J. *et al. Ecol. Lett.* **15**, 1071–1082 (2012).
4. Collin, R. *et al. Trends Ecol. Evol.* **23**, 602–609 (2008).
5. Esfeld, K. *et al. Curr. Biol.* **28**, 3776-3786. (2019).
6. Lahti, D. C. *et al. Trends Ecol. Evol.* **24**, 487–496 (2009).
7. Tan, L. *et al. Phys. Rev. Lett.* **198102**, 1–4 (2011).
8. Wagner, A. (Princeton University Press, 2007).
9. Visser, B. *et al. J. Insect Physiol.* **54**, 1315–1322 (2008).
10. Visser, B. *et al. Proc. Natl. Acad. Sci.* **107**, 8677–8682 (2010).
11. Turkish, A. R. *et al. Am. J. Physiol. Endocrinol. Metab.* **297**, E19–E27 (2009).
12. Amiar, S. *et al. Cell Rep.* **30**, 3778-3792.e9 (2020).
13. Jenke-kodama, H. *et al. Mol. Biol. Evol.* **22**, 2027–2039 (2003).
14. Maier, T. *et al. Science (80-.).* **321**, 1315–1323 (2008).
15. Maier, T. *et al. Q. Rev. Biophys.* **43**, 373–422 (2010).
16. Bukhari, H. S. T. *et al. Struct. Des.* **22**, 1775–1785 (2014).

17. Peters, R. S. *et al. Curr. Biol.* **27**, 1013–1018 (2017).
18. Moiroux, J. *et al. Ecol. Entomol.* **35**, 727–736 (2010).
19. Eijs, I. E. M. *et al. Ecol. Entomol.* **23**, 133–138 (1998).
20. Le Lann, C. *et al. Oecologia* **174**, 967–977 (2014).
21. Visser, B. *et al. Ecol. Evol.* **8**, 7355–7364 (2018).
22. Prager, L. *et al. Insect Biochem. Mol. Biol.* **115**, 103256 (2019).
23. Godfray, H. C. J. (Princeton University Press, 1994).
24. Ament, S. A. *et al. J. Exp. Biol.* **214**, 3808–21 (2011).
25. Visser, B. *et al. Genome Biol. Evol.* (2012).
26. Visser, B. *et al. R. Soc. Open Sci.* **4**, 160815 (2017).
27. Abu-Elheiga, L. *et al. Proc. Natl. Acad. Sci. U. S. A.* **102**, 12011–12016 (2005).
28. Fleury, F. *et al. Adv. Parasitol.* **70**, 3–44 (2009).
29. Lue, C. *et al. Environ. Entomol.* **47**, 1096–1106 (2018).
30. Hoffmann, A. R. Y. A. *et al. Heredity (Edinb).* **83**, 637–643 (1999).
31. Whiting, M. F. *et al. Nature* **421**, 264–267 (2003).
32. Domes, K. *et al. Proc. Natl. Acad. Sci.* **104**, 7139–7144 (2007).
33. Collin, R. *et al. Proc. R. Soc. B Biol. Sci.* **270**, 2551–2555 (2003).
34. Stone, G. *et al. Curr. Biol.* **13**, PR436-R438 (2003).
35. Goldberg, E. E. *et al. Evolution (N. Y).* **62**, 2727–2741 (2008).
36. Christin, P.-A. *et al. Trends Ecol. Evol.* **25**, P403–P409 (2010).
37. Galis, F. *et al. Evolution (N. Y).* 1–11 (2010).
38. Hall, B. K. *Biol. Rev.* **59**, 89–124 (1984).
39. Zhang, C.-X. *et al. Annu. Rev. Entomol.* **64**, 297–314 (2019).
40. Parker, D. J. *et al. Mol. Biol. Evol.* **36**, 350–364 (2018).
41. Tvedte, E. S. *et al. Curr. Opin. Insect Sci.* **31**, 77–83 (2019).
42. Hanschen, E. R. *et al. Evolution (N. Y).* **72**, 386–398 (2017).
43. Kraaijeveld, K. *et al. G3 Genes Genomes Genet.* (2019).
44. Kraaijeveld, K. *et al. Genome Biol. Evol.* **8**, 3685–3695 (2016).
45. Marchler-Bauer, A. *et al. Nucleic Acids Res.* **45**, D200–D203 (2017).
46. Katoh, K. *et al. Mol. Biol. Evol.* **30**, 772–780 (2013).

Supplementary table 1: Fat synthesis is induced plastically based on host fat content

We analysed the data presented in Figure 2 in the main text statistically by means of a linear mixed effects model (GLMM, lme4 package) with host (lean *D. simulans* and fat *D. melanogaster*) as fixed effect, population (Japan, United Kingdom 1 and 2, Belgium 1 and 2), family, and experiment (each experiment was conducted twice) as random factors, and percentage of incorporation of stable isotopes as dependent variable (n = 138). Non-significant terms (i.e. population and experiment) were sequentially removed from the model to obtain the minimal adequate model as reported in the table. When referring to "families," we are referring to the comparison of daughters of singly inseminated females, which (in these haplodiploid insects) share 75% of their genome.

Fixed effects		Estimate	Std. error	t-value	p-value
Intercept		2.073	0.258	8.053	**0.001**
Host *D. simulans*		1.129	0.243	4.636	**0.001**
Random factors		Variance	Std. error		p-value
Family (intercept)		0.544	0.738		**0.002**

Supplementary table 2: Fat synthesis is plastically induced in lean *D. melanogaster* hosts
Mean absolute fat amount ± 1se (in µg) was quantified in adult wasps from field-caught *L. heterotoma* populations raised on lean *D. melanogaster* hosts at two adult stages (Emerged: just after emergence; Fed: having fed for 7 days after emergence). Lean *D. melanogaster* hosts were produced by rearing the larvae on a medium containing 100 times less sugar than usual. P-values based on t-tests (with log-transformed data in case of the Belgium 2 population noted by ^) reveal whether 7 days of feeding led to a significant increase in fat content, indicating the occurrence of fat synthesis (in bold).

Population	Sample size	Emerged	Fed	p-value
Belgium 2	31	27.40 ± 1.82	38.08 ± 3.50	**0.018 (^)**
UK 1	33	25.09 ± 2.51	40.50 ± 2.66	**<0.001**
UK 2	35	27.25 ± 2.60	38.62 ± 2.70	**0.006**
Japan	31	34.70 ± 2.72	34.36 ± 3.69	0.954

Supplementary text 1: Acetyl coenzyme A carboxylase (ACC) amino acid sequence alignment for *D. melanogaster*, *P. maculata*, *L. clavipes*, *G. legneri* and *A. bilineata* :

```
                                            10        20        30
40        50        60        70        80        90       100

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  --------------------------------
----------------------------MLITILGTL---
AAFLAFLLTLIFGRGQKRSKPVQSS--
Pmac ACC protein Pmac maker-sc  --------------------------------
-----------------------------------------------------------------
--
Gleg ACC protein augustus_mask
MTDQEDKSEPNLREERAAPQRKIRHRGIVEREILFPNVTESLEVAVVEDTLSILRITFEAALT.A
..LA.LAC...GVV.SRVWNASASS.GNTGYTEND
Lclav ACC protein scf718000516  --------------------------------
-----------------------------------------------------------------
--
Abil ACC protein Abil maker-sc  --------------------------------
----------------------------------------------MAKQLNR.NS..-------
--

                                            110       120       130
140       150       160       170       180       190       200

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  ------------------
AATSATTATTGDSDNGNTNHNSSV-
IAATATATTTSSKPPIAPAAPPSAVKASDKRFPKACIKKVQFSSESLRSDVDELCDQ
Pmac ACC protein Pmac maker-sc  --------------------------------
-----------------------------------------------------------------
--
Gleg ACC protein augustus_mask  NSRIDRDSEQEEALNAERM.EAPVSFVV..P.-
--ADP.EELE.EDSFPNE.SD.NIQMQQTIA.GLLE---R.------RRLR--------------
--
Lclav ACC protein scf718000516  ------------------MTET.VSFVL..P--
---DPKEELE.EDSFPE--PEANDR.QQPIL.GL.E---R.------.RLR.-------------
--
Abil ACC protein Abil maker-sc  -------------------------
FVI..E.V.S.DN---------.P.DES.------.FTI..I.E---------------------
----------

                                            210       220       230
240       250       260       270       280       290       300
```

```
....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  LKDSALSNLSNDNIRIACHQNNNN---
SSINKNQNNNSIDISISISKMSETNESNDTAAQSAEGERPSFLVGDEIDERAAEAGEACDEFPLK
MQNDVRQN
Pmac ACC protein Pmac maker-sc  --------------------------------
-----------------------------------------------------------------
--
Gleg ACC protein augustus_mask  --------------------------------
-----------------------------------------------------------------
--
Lclav ACC protein scf718000516  ----V.AER.EK.---
.AWKSSD.LLAGIVRM.SE..EDN---EVF.REPSI.TEK..G----------------------
-------------------
Abil ACC protein Abil maker-sc  ------.H.QD.D-----------------
.L.G.G.GY-----..-------------------------------------------
.VW..QARLLK

                                            310       320       330
340       350       360       370       380       390       400

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  GDISERRKRLRPSMSRGTGL--
GQDRHQDRDFHIATTEEFVKRFGGTRVINKVLIANNGIAAVKCMRSIRRWAYEMFKNERAIRFVV
MVTPEDLKANAEY
Pmac ACC protein Pmac maker-sc  -------------......--
....Y......V........K...N....R....................S....Q....V....
............
Gleg ACC protein augustus_mask  -----------
....Q..VMIQA.S.L.EK..TV..P....H.............................S...
......V................
Lclav ACC protein scf718000516  ---------
.T....Q..VMIQA.S.QLEK..TV..P....R.....K.......................S.
........V................
Abil ACC protein Abil maker-sc  PGV.-----.T....Q..VI--
M.N.LHE...TV..P....R.....K....................V...S........SV....
............

                                            410       420       430
440       450       460       470       480       490       500

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
IKMADHYVPVPGGSNNNNYANVELIVDIALRTQVQAVWAGWGHASENPKLPELLHKEGLVFLGPP
```

```
ERAMWALGDKVASSIVAQTAEIPTLPWSGSDLKAQ
Pmac ACC protein Pmac maker-sc
.......................................................N.......
....................D.........E...H
Gleg ACC protein augustus_mask
.....Q.......T...............I.........................NNIC.I..S
..........I.........DV........E...H
Lclav ACC protein scf718000516
.....Q...................I...I.........................NNIC.I..S
..........I.........DV........E....
Abil ACC protein Abil maker-sc
.............T...........I...V.C.......................NNIA.I...
DK........I.........D..............

                                          510       520       530
540       550       560       570       580       590       600

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
YSGKKIKISSELFARGCVTNVEQGLAAVNKIGFPVMIKASEGGGGKGIRRVDTTEEFPGLFRQVQ
AEVPGSPIFVMKLARGARHLEVQLLADQYGNAISL
Pmac ACC protein Pmac maker-sc
..........D..N........E.............................A...SAM.....
..I.......................M.......
Gleg ACC protein augustus_mask
.............KK...ST..EC...A........V............K.ENA..L.V......
..I......I....KC...I....V..N.......
Lclav ACC protein scf718000516
.............KK...STI.EC.T.A........V............K.ENA..L.A......
T.I......I....KC...........N.......
Abil ACC protein Abil maker-sc
..........D.......HTA...QQ.AQ...................K..NP.D.AAS.....
..............KC..................I

                                          610       620       630
640       650       660       670       680       690       700

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
FGRDCSIQRRHQKIIEEAPAIVAQPEVFEDMEKAAVRLAKMVGYVSAGTVEYLYDPEGRYFFLEL
NPRLQVEHPCTEMVADVNLPAAQLQIGMGIPLYRL
Pmac ACC protein Pmac maker-sc
.................................................A..K......
....................C.......V..F..
Gleg ACC protein augustus_mask
```

```
....................VI.K.D...E.........................TS...Y....
.............IS...........A..L..HHI
Lclav ACC protein scf718000516
....................VI.K..I..E.........................S...Y....
..............S...........A..L..HHI
Abil ACC protein Abil maker-sc
....................VI..QS.............................TS.H.Y....
..............S...........A..LS.HKI

                                                  710       720       730
740       750       760       770                 780       790       800

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
KDIRLLYGESPWGSSVIDFENPPNKPRPSGHVIAARITSENPDEGFKPSSGTVQELNFRSSKNVW
GYFSVAASGGLHEFADSQFGHCFSWGENRQQAREN
Pmac ACC protein Pmac maker-sc
.........N...T.I.................................................
...................................
Gleg ACC protein augustus_mask
.............D.Q...DQ.RH..Q.W.................................
.....G....................D.H.....
Lclav ACC protein scf718000516
.............D.A...DQ.RH..Q.W.................................
.....G....................D.......
Abil ACC protein Abil maker-sc
.........T...D.....DQ.KH..Q.W.................................
.............................E.....

                                                  810       820       830
840       850       860       870                 880       890       900

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
LVIALKELSIRGDFRTTVEYLITLLETNRFLDNSIDTAWLDALIAERVQSEKPDILLGVMCGSLH
IADRQITESFSSFQTSLEKGQIQAANTLTNVVDVE
Pmac ACC protein Pmac maker-sc
.............................................................A..
......S.A.....................I...
Gleg ACC protein augustus_mask
...........................ES.QQ.N.......L......R.D...V..A.T..A..
....T..AA.TG...A........S.D.E.I....
Lclav ACC protein scf718000516
..V.........................EC.QQ.C.......V......R.D...V..A.T..A..
....T..AA..G...A........S.D.N..M...
```

```
Abil ACC protein Abil maker-sc
.........................KS.Q..T.......I..S..M.......M...I..A..
...KT.STA.NE..I...R....GS...NHTM...

                                          910       920       930
940       950       960       970       980       990       1000

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
LINDGIRYKVQAAKSGANSYFLLMNSSFKEIEVHRLSDGGLLISLEGASYTTYMKEEVDRYRIVI
GNQTCVFEKENDPSLLRSPSAGKLINMI-------
Pmac ACC protein Pmac maker-sc
...G.N............T......N....V.........M.M.F....................
.........................LL-------
Gleg ACC protein augustus_mask
.....YK..I......P.....V..G.Y..V...........L..D...V....R........I.
.....I...D...............SFL-------
Lclav ACC protein scf718000516
.....YK.....T...L.T...V..N.Y..VDI.........L..D...F....R..........
.......D.D................FLVEDGGHV
Abil ACC protein Abil maker-sc
.....HK.R...T.....T...V..G......L.......I.L.VD...F...............
.....I.D.....T...........GFL-------

                                          1010      1020      1030
1040      1050      1060      1070      1080      1090      1100

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  ------
VEDGAHVSKGQAYAEIEVMKMVMTLTSQEAGTVTFVRRPGAVLDAGSLLGHLELDDPSLVTKAQP
FKGQFLQPE--NAPVPEKLNRVHNTYKSI
Pmac ACC protein Pmac maker-sc  ------
.......N...................L.....................................
Y....P...--.P.L.........I....
Gleg ACC protein augustus_mask  ------
....G..DA............I..V..S...S.FY.K......E..T.IA.............SE
YT...PPAA--APAI.....HL.TK.RTA
Lclav ACC protein scf718000516
DAGQAS....G..DA................V.AG...SIFY.K......E..T.IA.........
.....EYL....A.V--TPA......HL.AK.RAA
Abil ACC protein Abil maker-sc  ------
....G..YR..................AG...S.SY.K.A........IIAT.....A.......L
YTSP.PDLDVSHPLAS....HI..S....

                                          1110      1120      1130
```

```
1140      1150      1160      1170      1180      1190      1200

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
LENTLAGYCLPEPFNAQRLRDIIEKFMQSLRDPSLPLLELQEVIASISGRIPISVEKKIRKLMTL
YERNITSVLAQFPSQQIASVIDSHAATLQKRADRD
Pmac ACC protein Pmac maker-sc
.....................V.........................................
...............................T...
Gleg ACC protein augustus_mask
...........D.YHLP....L.....N.................T.................S.
..................A...G.....S..S...
Lclav ACC protein scf718000516
.......F...D.YHLP...EL.....N.................T...................
..................A...G.....S......
Abil ACC protein Abil maker-sc
.....Q.F...D.Y..K...EV.....A.......................PA.........S.
..........................SM....E..

                                          1210      1220      1230
1240      1250      1260      1270      1280      1290      1300

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
VFFLTTQSIVQLVQRYRNGIRGRMKAAVHELLRQYYDVESQFQYGHYDKCVGLVREHNKDDMQTV
VNTIFSHSQVAKKNLLVTLLIDHLWANEPGLTDEL
Pmac ACC protein Pmac maker-sc
N......G..............................C...H............R........
...................................
Gleg ACC protein augustus_mask
.......A.................T..........T......Q.......SALI.QY...VA..
TGM....N..T...V...M................
Lclav ACC protein scf718000516
.......T...........T.....T......K...N......Q.......SALIDQF...KTM.
TS........T...V...M......T.........
Abil ACC protein Abil maker-sc
N...A.DG.....................D......N......L.......SAL.DKH....SM.
TQI...........M...M......S.........

                                          1310      1320      1330
1340      1350      1360      1370      1380      1390      1400

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
```

```
ANTLSELTSLNRAEHSRVALRSRQVLIAAHQPAYELRHNQMESIFLSAVDMYGHDFHPENLQRLI
LSETSIFDILHDFFYHSNRAVCNAALEVYVRRAYT
Pmac ACC protein Pmac maker-sc
.................................................................
...................................
Gleg ACC protein augustus_mask
SS..T.......T........A.........................................K..
...................................V
Lclav ACC protein scf718000516
.S..T.......T........A.........................................K..
....C..............V..............I
Abil ACC protein Abil maker-sc
.A..N.......S........A.................................E.......K..
V...............T..................

                                          1410      1420      1430
1440      1450      1460      1470      1480      1490      1500

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
SYELTCLQHLELSGGLPLVHFQFLLPTAHPNRLFSRMSSPDGLDQAAAESLGNSFVRTGAIAAFD
SFEHFEMYSDEILDLLEDFVSPAMVNAKVLEAVEA
Pmac ACC protein Pmac maker-sc
..D..............I................L..A.-
EAATE.GTDN..T.Y....CM.........D............LA.ST..S.........
Gleg ACC protein augustus_mask  ..............EI..........NN....-
---------QN.STVNH------
....M...QDL.Q.SQ.A..V......LS..SS.S..I......
Lclav ACC protein scf718000516  ..............EI.......M..NN....-
---------QN.SLVNH------
....M...QDL.Q.NQ....V......LS..NS.S.........
Abil ACC protein Abil maker-sc
..DI.........AEV..I.......PS.....VTLD.I---
.EETEPAKVFD..Q...CM...E..QQ..S.A...F..I...AN..TIS..D.NML.S

                                          1510      1520      1530
1540      1550      1560      1570      1580      1590      1600

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  ADSISDSR-HSTSINVSLSDPVTRANAAEEA-
KSTEPIHIVSVAVRETGELDDLQMAQIFGNYCQEHNEELFQRRIRRITFAALKKRQFPKFFTFRA
RDK
Pmac ACC protein Pmac maker-sc  ....G.G.-L...........IS........-
........I.........M..V..........KQ.RD.......................Y..
..N
```

```
Gleg ACC protein augustus_mask  .G--.E..-........I.-
TAEPSTTI.RGERPS..V..L.I..Q.IDNQ..TAL.RM..DW.ANNKD..IS.G...V......
R.........Q..G
Lclav ACC protein scf718000516  VG--.E..-
.........TT.EGNAQN.SGDDPA..F..L.I..IDK.NQ..AT..RV..DW.ALNKD..IA.
GV..V..L........L....Q..G
Abil ACC protein Abil maker-sc  G---G...TN........IDGQ.QITEDSN-
.VC.....LHIG.KDK.DE..ST.SR...SF.ER.RQD.ET.G.........HK.......Y..
..G

                                          1610      1620      1630
1640      1650      1660      1670      1680      1690      1700

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
FTEDRIYRHLEPASAFHLELNRMKTYDLEALPTANQKMHLYLGKAKVSKGQEVTDYRFFIRSIIR
HSDLITKEASFEYLQNEGERVLLEAMDELEVAFSH
Pmac ACC protein Pmac maker-sc
YE..........C.YQ.........................R..........F........
.................................
Gleg ACC protein augustus_mask
.V.........GC..Q......R........S........Q...A...Q............
..........D..H...................
Lclav ACC protein scf718000516
.F...V......GC...I.....R........S........Q...A...Q............
..........D..H...................
Abil ACC protein Abil maker-sc
.K..........C..Q......R..N......S............AP.H.............
.................................

                                          1710      1720      1730
1740      1750      1760      1770      1780      1790      1800

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
PHAKRTDCNHIFLNFVPTVIMDPAKIEESVTKMIMRYGPRLWKLRVLQAELKMVIRQSPQSPTQA
VRLCIANDSGYFLDISMYTEQTEPETGIIKFKAYG
Pmac ACC protein Pmac maker-sc
.F.......................................................L...N..A...S
...................H.DK...V...M...
Gleg ACC protein augustus_mask
.L....E........A.N......R......S.VL...........R...I..T..PA.GK..TN
I.........SI.LHL...A.D.K....R.ESFP
Lclav ACC protein scf718000516
.L....................R......S.VL...Q.......R...I..T..PA.GK..SN
```

```
.....S.....SI.LHL...AIDQK....R.ES.-
Abil ACC protein Abil maker-sc
.QSR...........I.......S....A..S.V...................T..S..T...TT
..........Y...N....VVNVD....R.E...

                                          1810      1820      1830
1840      1850      1860      1870      1880      1890      1900

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  E--------
KQGSLHGHPISTPYMTKDFLQQKRFQAQSNGTTYVYDVPDMFRQMTERHWREFSKARPT-
VDIRTPDKILIECKELVLEGDN---LVEMQR
Pmac ACC protein Pmac maker-sc  .--------
.......L.............................I..........L.K.Y.M....-
....I.E......V....D...---....K.
Gleg ACC protein augustus_mask
SQNPNNPNPRI.PM..L......L...Y..A.......A.......L......QL.KT.AKYIDE
.SAIEP.TM.NPVM-DSV...V..E.---...LK.
Lclav ACC protein scf718000516  SGSANNSN-
RP.PM..L......L...Y..A.......S.......L......QV.KS.K..IDE..S-
EV.TI.NPLI-.IV....D.-D---...LK.
Abil ACC protein Abil maker-sc  T--------
...P...L......LA..Y.........QS.......Y.......VDLL.KQY.QE.MN-
EVVVI.E.VM-D.I....DPE.ESR...QK.

                                          1910      1920      1930
1940      1950      1960      1970      1980      1990      2000

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
LPGENNCGMVAWRIVLATPEYPNGREIIVIANDLTYLIGSFGIKEDVLFAKASQLARQLKVPRIY
ISVNSGARIGLAEEVKAMFKIAWEDPEEPDKGFKY
Pmac ACC protein Pmac maker-sc
.....T.......I.......E...M........FF......P..IV.H.......SR......
....................V..............
Gleg ACC protein augustus_mask
.....DV......FT.Y...C.T..DV.L.G..I.HM.....PR..I..YR..ER.....I....
FAA..............L.......EM..E.....
Lclav ACC protein scf718000516
.....DV......LT.Y...C.T..D..L......H.......P..I..F...ER....GI..V.
F.A......A.......L.R.....EA..E.....
Abil ACC protein Abil maker-sc
V.....V......LT.Y.....A..I.......I.F.M...APR..KV.GL..E...N.......
.AA..............LY....D..N...R..R.
```

```
                                                  2010      2020      2030
2040      2050      2060      2070      2080      2090      2100

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  LYLSTEDYAQVANLNSVRAILIED-
EGEQRYKITDIIGKDDGLGVENLRYAGLIAGETSQAYEEIVTIAMVTCRTIGIGSYVVRLGQRVI
QIDNSHIILT
Pmac ACC protein Pmac maker-sc  ...T....SR..K..........-
...P..........E....................D.....S...........L.......
..E.......
Gleg ACC protein augustus_mask
I..TPD...RL.P....K.S....PA..S..........Y.I.....K...M......K..D.V
...SI.S..A......L.........E.......
Lclav ACC protein scf718000516  I..TPD...RLSP....K.S....-
G..S..R...............K.........K....V...SV.S..A..............
..E.......
Abil ACC protein Abil maker-sc  ...TP....K.SAW.....V....-
...S............F.....Q.............D....S..S..A....A.L.......
.........

                                                  2110      2120      2130
2140      2150      2160      2170      2180      2190      2200

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
GYAALNKLLGRKVYASNNQLGGTQIMFNNGVTHKTEAIDLDGVYTILDWLSYIPAYIGCDLPIVL
P-NDRIERPVDFMPTKSPYDPRWMLGGRVNPVNAN
Pmac ACC protein Pmac maker-sc
....................V...Y.........L.......................E
.-....D................A.....G..S
Gleg ACC protein augustus_mask
.NR...AV...E.........V...H....S.S.DVR.....A.A.K....V.KAK.AP...LP
.LL.P...E.MYT........F..D...S.SDP.
Lclav ACC protein scf718000516
..R...TV...E.........I...H...IS.AI.PR.....E.V.R....M.KSK.AP...IE
S-I.P.D.EIG.V...A.......E.K--Q..DH
Abil ACC protein Abil maker-sc
..S.......E.........I...Y....S....PR....I....K......KDKLSGV..LP
.-T.PYT.EIGY....A.......A..Q..NSPA

                                                  2210      2220      2230
2240      2250      2260      2270      2280      2290      2300

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
```

```
Dmel ACC protein transcript va
DWENGFFDRDSWSEIMASWAKTVVTGRARLGGVPVGVIAVETRTVEVEMPADPANLDSEAKTLQQ
AGQVWYPDSSYKTAQAIKDFGREELPLIVFANWRG
Pmac ACC protein Pmac maker-sc
E...............P.............I..............................
............S............MI......
Gleg ACC protein augustus_mask
V..S.....N..Q...KP..Q...........I.C...........LHL............IS.
.........A..........H.....FI......
Lclav ACC protein scf718000516
T..S.....G..Q...KP..Q...........I.C.I........LHL............VS.
.....F...A..........NK.....FI......
Abil ACC protein Abil maker-sc
E..A....K......QP..Q...........I............LK.............VS.
.....F...A......Q...K.D...FI......

                                          2310      2320      2330
2340      2350      2360      2370      2380      2390      2400

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
FSGGMKDMYEQIVKFGAYIVDGLREYKKPVLIYLPPNAELRGGAWAVLDSLINPRYMETYADPEA
RGGVLEPEGIVEIKYKEKDLVKTIHRLDPTTIALK
Pmac ACC protein Pmac maker-sc
...........................I...........................S......
...........................A..----
Gleg ACC protein augustus_mask
............M............TR.I.V.I...G........V.PT...D...MF..NTS
.......DA.....F.TR.TL..M..V.HIIQK..
Lclav ACC protein scf718000516
............M............TR.IVV.I..YG........V.PM....H..MF..HTS
..............FRN..I...M..N.SVIHN..
Abil ACC protein Abil maker-sc
...........V............K.R..II..I...G........V.PF..S....M......
...I...........RK...L..M..I.A.LMQ.D

                                          2410      2420      2430
2440      2450      2460      2470      2480      2490      2500

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va  KELDEANASGDKVRAAQ----------------
---
VDEKIKARIAVLMHVYHTVAVHFADLHDTPERMLEKECISEIVPWRDSRRWLYWRLRRLLLEDA
Pmac ACC protein Pmac maker-sc  --------------------------------
------------------------------------------------------------
```

```
--
Gleg ACC protein augustus_mask  EQ.SNTSSP-------EERSQ------------
---
IEA...E.EQQ.EPM.RQI................INA.ND.....R..KL.......R.F.EE
Lclav ACC protein scf718000516  EK.ASCGSA-------EERAS------------
---
.ERE.HD.ECH.ESM..Q..................NT.Q...A..SA..I.......R....R
Abil ACC protein Abil maker-sc
EK.KML..ANVPIEILERRGSVTQTPERKKTPEIIA.EKE.VE.ENY.LPM..Q...N.......
....H..GT.LD.....K..TI..........Q.R

                                          2510      2520      2530
2540      2550      2560      2570      2580      2590      2600

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel ACC protein transcript va
YIKKILRAQDNLSVGQAKQMLRRWLVEEKGATEAYLWDKNEEMVSWYEEQINAE---
SIVSRNVNSVRRDAIISTISKMLEDCPDVALDAVVGLCQGLTP
Pmac ACC protein Pmac maker-sc  ---------------------------------
-----------------------------------------------K..SF----------
--
Gleg ACC protein augustus_mask
IRSEV.ST.PG.DIR.VGA.....FI.D..T..S....QD.TAAR.L.N.L.D.--
N.V....IAC.KK.TVVTR.KES..AY.E.R.N.MLEIVHR.HS
Lclav ACC protein scf718000516
IRSE.IST.PG.D.R.VDA.....F..D.....S....QD.VVAT.L.A.CEN.--
S.V.M..ISC.KN.S.VTRVKEA..V..E.RF...LEIVNR.Q.
Abil ACC protein Abil maker-sc
V.TQL.ETNS..GI..GEA.....F......S.G.K..N..AV.E.L.K.MSV.NEN.ML...LH
A.KK..L.QK.KNSI.....L......EIL.K.ND

                                          2610      2620
                                  ....|....|....|....|....|...
Dmel ACC protein transcript va    VNRGVVVRTLAQMQLNEETSNSNQG---
Pmac ACC protein Pmac maker-sc    ----.QY.CF------------------
Gleg ACC protein augustus_mask    TE.AELL...S.IEASGQEHHNDSNVSS
Lclav ACC protein scf718000516    AEIAELQ.....LESTSQENHNDSSASS
Abil ACC protein Abil maker-sc    NQKAE.I...S.V.PET.S---------
```

Supplementary text 2: Fatty acid synthase (FAS) amino acid sequence alignment for *D. melanogaster*, *P. maculata*, *L. clavipes*, *G. legneri* and *A. bilineata*:

```
                                                10        20        30
40        50        60        70        80        90       100

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein    MPARF-------
AEEVITAEPAQRAAPQLDLGGGHYVPRQQHLNDEIAITGFSGRLPESSTIEEFKQNLFDGVDMVN
DDPRRWERGLYGLPDRIGKLKDS
Lclav FAS transcript protein   .....ESVNAPIVRDQ.IPNGT------------
-----TYSID.D.V............N......K....I.L.T..E...PS..H...T.T....-
-
Gleg protein FAS               ...Q.ESMNTPIVR.PVVTNGSR-----------
-----SFSMD.D.V.S.........EN....QK...E.I.L.T..E...PS..H...T.T....-
-
Abil protein FAS               ....-----VPEMR-----NGE------------
-----CHQA...VV...L.......Q.......Q......L.T..E...TA..H...T.T..I.-
-
Pmac protein FAS               .....-------
VDG..NENGGRN.PHD.QERF.LHSSGPMTVS.-
..............N.....K...E.............K........L..I..E

                                            110       120       130
140       150       160       170       180       190       200

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
DLENFDQQFFGVHQKQAECMDPLLRMLLELTHEAIIDAGLNPSDLRGSRTGVYIGVSNSETEQHW
CSDADRVNGYGLTGCARAMFANRISFTFDFKGPSY
Lclav FAS transcript protein
.IAS..AT.....A...HV...Q................I...EI...K...F....D..SDEF.
TA.P.M........C....P....Y....T...F
Gleg protein FAS
.IAS..AT.....A...NV...Q..LM..A.....V...F..T.V..T....F....S..SDDF.
TRSPET........C....P....Y....N....
Abil protein FAS
.ITH..AT.....A...HL...Q...........L....I...EF...N...FV...E..SSEQ.
TR.P.AI...S....C....P....Y...L.....
Pmac protein FAS
...H......S..........QM.....C.Y.....S.V...EV............A.....Y.
TA.Q...................V.Y........

                                            210       220       230
240       250       260       270       280       290       300
```

```
....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
SIDTACSSSLYALEQAFSDMREGKVDNALVAGAGLILKPTMSLQFKRLNMLSPDGSCKAFDESGN
GYVRSDGCVVLLLQRTSAARRVYASILNVRTNTDG
Lclav FAS transcript protein
AV.......M..MH..VAA..T.QC.S.I.G.VN.V....S....H......M..A.....A...
.....EAA..IF..KARNS.....TVV.SK.....
Gleg protein FAS
A.........F.MH..IAA..S.EC.A.I.G.VN.C....A....HK.S...ME.A.....A...
.....EAV..VF..KSKD......TVV.SK..V..
Abil protein FAS
A........MF..Q..YAAIKS.QC.S.I.G.VN.L....N....H..G...AE.K.....A..S
....AEAA..IV..KA........T..GAK.....
Pmac protein FAS
.V.......F...........D..C.S.V...L.V....................M.........
..........IF..K..QSK....TV.........

                                                 310       320       330
340       350       360       370       380       390       400

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
FKEQGITYPIGKMQNRLIRETYEEIGLNPADVVYVEAHGTGTKVGDPQEVNSITDFFCKDRTTPL
LIGSVKSNMGHSEPASGVCSVAKILIAMEEGVIPG
Lclav FAS transcript protein
N.VE...F.S.A...K.M..V.A.V.VD.V.......................A.L.....KK..
....I............L..I..M.L...A.....
Gleg protein FAS
Y.DL.....S.S...K.M..I.D.C.V..S..T....................A.L.....KN..
.L.................AI..L...L.T.M..P
Abil protein FAS
S......F.S.QV.....N.V.ADS.IE.N.......................A.....N.K...
.L...............L..I..M.....S.Q..A
Pmac protein FAS
Y........D.R.............N.D.NE.A.........................N.....
....................I..V..........A

                                                 410       420       430
440       450       460       470       480       490       500

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
NLHYNKPNPDLYGLVDGRLKVVDRNLPWNGGIIGLNSFGFGGANAHVILKSNPKPKALTPK--
DGALKVVLASGRTFEAVEQLLESASTNADDDEYLQLI
```

Lclav FAS transcript protein
...FKS..K.IPA.S....Q....SM.....LVAI.....LV.R.....IA-
.VLDVNVP.I.PV....DD..NLF.DRIKEHEK...FTSMV
Gleg protein FAS
...FQN..K.IPA.S...IQ..TQPTAYK.NLMAV.....ILVRGHS...LS-
.VMDR.VP.L.AV....N...NVM.DKIKEHHR...FIA..
Abil protein FAS
...FKN..T.IPA.C...I...AT.EK.T...V.V.....R.....ENW.V--
EQLPRL.VV....ED..NHF.DKIKEQSH.E.FYAML
Pmac protein FAS
....KN.....M.....K....E.....I.....TI..T--
V.PP.M.VC.....D..QE...D.TSHR.....A..

```
                                        510       520       530
540       550       560       570       580       590       600

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
```

Dmel FAS transcript protein
NEIHSKAIPNHFFRGYGVVSSKGTHQREVIESNDDKRPIWYIYSGMGSQWASMAKDLMKIEAFAK
TIQRCADVLKPEGVDLIDVLTRSTDKSFENILNSF
Lclav FAS transcript protein QDL.ANN.TG.GY..FQILGDVN.--
..IDQVGSE.....F.....SG.GRA.FC.DT.QSA.R...EA.....I...NLILNG.EE
..Q.VV...
Gleg protein FAS H...N.N..G.N....QILGGED.--
..IL.NHSA.....FVF.....PG.G.E.LHLDV.NRSLR...EA.RS.....M.IIQNG.NE
T....I...
Abil protein FAS .N..A.N.TG.NY..FA.LGDNEI--
.D.SMVGNE.K...FVF.....PG..R.....DL.QQ..K.A.QA.N.Y....E.I.LN..EE
TLT.VR...
Pmac protein FAS
.D....N..L.YY...C.MDT..SL....L.F..EN..V.....QF.V..N
S.H...KA.R...I..V.....L..D.....

```
                                        610       620       630
640       650       660       670       680       690       700

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
```

Dmel FAS transcript protein
ISIAAMQVALTDLLSSLGIHPDGIVGHSVGELGCAYADGCFTPEQTVLAAYWRGKSILDTQLAKG
KMAAVGLSWEDAHSRVPSDCFPVCHNSEDNCTISG
Lclav FAS transcript protein
V....I.IG.V.V.NLI..Q....I...I.....G....TM..S.....T...GN.PP.
A.....E.QK.C.PEIVLA...AA.SV....
Gleg protein FAS
V....I....V...T.....T.....I....L...A.AESD.PA.

C.....TKA.C.P.VV.A...AA.SV....
Abil protein FAS
....SI.IG.LEI.K..N.E....L...I..V.....TL.L...IQL.WA..TA..ESD.PP.
A.....T..ECKK.C...I.....SV....
Pmac protein FAS
.....T..N.K.....S.....Q..K.PP.
...SI..D..E..K.M.A.....A.....

```
                                          710       720       730
740       750       760       770       780       790       800

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
```

Dmel FAS transcript protein
PEASIEALVAKLNAEGVFAKAVNSSGYAFHSKYIAEAGPKLRKSLEKIIPNAKNRTARWISTSIP
ESAWNTPVAKQSSAAYHVNNLLSPVLFHEALQHVP
Lclav FAS transcript protein
.PEPLAKF.EE.KSQEI...Q.H...C.....SV.....TI.....P.Q.SS....S...
.T....SL.QL..P..Y.....Q...A...
Gleg protein FAS
.TG.V.KFIEE.KK.EI.....K.N.I.....S.....AG....LT.P.Q.SS....S...
....G..L.QL..P.....Q...A...
Abil protein FAS
.P...DKFT.E.TK..I...K.....F.....A.....A.DT..Q.P.A.S.....
....G..L.Q..N.....Y....TK.I.
Pmac protein FAS
.....D....Q.SS.....K.....D.....R.....NK.....
.....I.....Y.....I.

```
                                          810       820       830
840       850       860       870       880       890       900

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
```

Dmel FAS transcript protein
KNAISVEIAPHGLLQAILKRALGPDATNLSLVKRGHENNVEFFLTNVGKLFAAGAQPQVLTLVRP
ISYPVGRGTPMLNSKVGWDHTQKWLVAKF-GKETS
Lclav FAS transcript protein
D...VI.....C.....R.SFPSTV..IG.H..D.SD.LA.L...I...YV.....ILSK.YP.
VT.....MIL...STQ.S..D.S..SGN
Gleg protein FAS
ED..AI.....C.....R.S.PKTV..I..H..D.TD.MNYL.S.....YC.....ISK.YP.
.NF.....I..MIK...SVQ.G..NYAQNSAR
Abil protein FAS
D...VI.....G.SK.CA.I.....D.AR.L.S.I.RI.N..G..NIAN.YH.
V.F.....A.MIE...STE.S..NYCD.NDR
Pmac protein FAS

```
....AI....T.................V............L.M.I...Y....K.....MFK.
...................S...N.P.Y-....T

                                          910       920       930
940       950       960       970       980       990       1000

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
SGETIVEVDLSKEDDAFLAGHTIDGRILFPATGYMTLAWQTFAKMQGSEFHKTPVVMENLVFHRA
TILNKNAVVKFGINFFDGTGAFEICESGSLAVSGK
Lclav FAS transcript protein
..QSV..F.....S.SYI...C............L.IV.K....LR..SYE.....F.DVQ.L..
..MP.EGS...I..I.E.......S....ST.I..
Gleg protein FAS
...S...I...T.S..Y................L.IV......LHNE..NRM.IIL..VQ....
..MP.EGK...L..I.E...D........I..T..
Abil protein FAS
..QFVIDI.....EHKY.V...............L..V.K.....RNQD.EQL..II.DVK.M..
..MP.EGS...L..I.E.S.E........V.....
Pmac protein FAS
....VI.IN.G..E.S.F....................M.....K.M.YQ.C..I...I.....
.....EG.....L.......N.....G........

                                          1010      1020      1030
1040      1050      1060      1070      1080      1090      1100

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein    ITIPESIDNEELPLEEQTPSAVAKE-----
LGTNDVYKELRLRGYDYGGIFRGIVRSDTVASTGKLQWVDNWISFMDTMLQFSILSKNLRELYLP
TRIER
Lclav FAS transcript protein   .R.S.D.EKDQ.N.---
PIPVTCN.PDLLE.K......D........S...Q..KS..NR.I..N.A.NND...Y.......
A..G..T.D.F....LQY
Gleg protein FAS               .RRA.Y.E..Q.N.---
PIPVLR..ENILD.N...I..D........S...Q..KSA.NRGII...T.SND...........
...G..T.D.F....LQY
Abil protein FAS               .YVA.EPEKQF.T.---
PKHTLI..KDILD.N.P.I...........D......HSA.NYGLV...K.EQ.....I......
N.....T........LQ.
Pmac protein FAS               .S...D.EM.....DALPA.TLG..-----
.N...............S.......K.....N..Q...AE..V....................
....K

                                          1110      1120      1130
1140      1150      1160      1170      1180      1190      1200
```

```
....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
AVINPAKHFELLSALTKEEQVETGLPVQWYSDINVIKSAGVELRGLKANLAQRRPGTQAPPTLER
YQFVPNINTTDLNENSEKARLHALDVAIQVIIENS
Lclav FAS transcript protein    .A...ER.IQ.VEK.QEN.----
NI..FH..NVGIV..G.....M.SSI.P..QQ...D.K..K.S.I.YE..QA.V.DP..SK...
.TSLL..VR..I
Gleg protein FAS                .A....L.MH.V.G.KSD.----
....YS..N.GIL..G.I....M..S..P..QQA....KH...T...YETNNA.V.DPQ..KV..
MC.LF.I.C..M
Abil protein FAS                .I...VE.IR-----NAK.----
HVT.SM.R..D....G.........S..P..QQS.SA....Q...L.YL.MNQVVDEQI-----
T.T..S.IAL...
Pmac protein FAS
........LATV.K.SE.YLTLN....YM.G.......G...M.....S..SK...S.N......
.T.L..V.YAE.H.....S..Q..T..L.T.M...

                                        1210      1220      1230
1240      1250      1260      1270      1280      1290      1300

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
SGAVKLKGVELANGRNPDVLVANRLLQIIEGEPVLTGDVAVVTSNNNEETITAALG-
DSGVRVVSKDVLKEPVEQNCHFVFGIDVLSRPDTKTLENSIAS
Lclav FAS transcript protein    G.-
I.I.AI.TTME...EA.LTPIV.D.LLS..M.AV.LKLA.T--VPDNY.PIME-
QCN.KTTVV.IHSS..G.DMQLIITA.IMNNQMIAAVK.LE..
Gleg protein FAS                GS-
M...II.V.GE.SAES.L.PTVMDVLYS..LMSV.IQIA.T--TP..YN..ME-
QYN.KT.VR..NSN.AG.DL.T.IAP....NKNVNM.K.IA..
Abil protein FAS
G..L.M.V...QGSK.IEQ.LIPKVQG.L.CQ.M..VE-SILV.--Q.NID....E-
EKSIK.SR..PSADAF...A...LMS...AYNKSEV.T.AFK.
Pmac protein FAS
Q..I.I.............M.VK........TI.A....A.Y......AST..T.....II.
.NI.E........LYAL.......MI....K.T

                                        1310      1320      1330
1340      1350      1360      1370      1380      1390      1400

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
IRENGFLILEETLPTYTKTGRALLTKFGFVAVQEQSLGATRVLVLARKAVDLKTRKSVVVVATEQ
```

NFNWVDDLKAALATAATEEQYVYVVCQGEELFGAV
Lclav FAS transcript protein .KPG..ILT..ATEI----
DESI.KGSSLIVIGK.VVPG-KSYI.LK.K-EEMD-
VPL.IKV..K..S..N...V..KKSE..G.K.L..S....AL.L.
Gleg protein FAS LKNG..A.....GAV----
DMK..NGT.LLYAGK.ISAG-KTYI.LK.R-ED.K-
EPIIIQI..R..S.LEGV....KKSE..G.E.LL.S.....L.L.
Abil protein FAS LKPG..VLF..SSNF---
SDYS.F.SQELEI.YQ.RTPM-KIYI....Q.QVAQ-
DAIIIEV..NTYS..EPI.Q.MKESE.NNRKI.LIV.....S.L.
Pmac protein FAS
.KD.....F..STTS.G.SS.D..H.Y.LIV.T..VI.GS....M...P....Q.DA...HV..A
..D.LE...E...K..EI.R.....

1410 1420 1430
1440 1450 1460 1470 1480 1490 1500

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
GLMTCIKNENGGKLARLVFVQDAKAEKFSLTSTLYRQQLEKDLISNVLKNGAWGTFRHLKLETQQ
--ATLQVEHAYVNALVKGDLASLKWIEAAQADTAA
Lclav FAS transcript protein
.....VRQ.A..MNV.YF.I..VN.SA...DDAF.AK.FD.QVMA....G.Q..SY...R.DK.S
DIPS.....I...TR...S..R...GPLCYYEP
Gleg protein FAS
.F....RR.P..MN..Y..I..KN.P..G..TPF.AD..S.Q.A.....G.Q..SY...R.DQ.N
DASS.....I.T..R...S..R...GPLSYYQP
Abil protein FAS
.MVN.L.Q.P..VNM.A.LI..T...T.N.S.KFFVD..Q...VH.....I..N....S.SMEK
--S.....I.T.TR.....GPLGYYNN
Pmac protein FAS
.F.N.....M..I..KN.....NK..AE..S....N.....SV.....DV
--.....T.....P.L-CS.

1510 1520 1530
1540 1550 1560 1570 1580 1590 1600

....|....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
TVDKNLETCTVYYAPINFRDVMLTSGKLAADALPGDLAEQDCVLGLEFAGRDTQGRRVMAMVPAK
SLATTCVASKRMMWQIPEKWTMEEASTVPCVYSTV
Lclav FAS transcript protein
EKFVGQ.F.D.....L....I..AT...PP.....S...I.....SS.....G..A.R
G....LL.DPGF..EV.D...L...A.I.V..A.S
Gleg protein FAS

EKYP.T.M.S.....L....I..AT...PP.....G...I.....S....K....IG.IA.R
G....VL.DPGFL.EV.D...L...A.I.V..A.S
Abil protein FAS DD-
P.A.L.S.....L....I..AT...PP.....G...I.....S...SR.....G..A.....
..VL.DPGFL.EV....SL...A.I.V..G.S
Pmac protein FAS
LL.Q...L.....S....S.....Q.....SK...I.....
.....N...E..DN.....A..

```
                                        1610      1620      1630
1640      1650      1660      1670      1680      1690      1700

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
```

Dmel FAS transcript protein
YYALVVRGQMKKGEKILIHAGSGGVGQAAISVALAHGLTVFTTVGSKEKREFLLKRFPKLQERNI
GNSRDTSFEQLVLRETKGRGVDLVLNSLSEEKLQA
Lclav FAS transcript protein
.....RLRP..S.....S.AI..HA.C.....S.....K.T..Q.TDK..
.....T..G.....
Gleg protein FAS
....F...SLRP..S.....T.....S.AI..H..CK.....K.T..Q.TD...
.....T..N.....V.....A.....
Abil protein FAS
....I...GLRP..S..V...T.....S.AI..HM.CK.....S.QA..D..K....Q.TDNQ.
.....I.TQ.....AG.Q...
Pmac protein FAS
.....S...I..H.....K....Q.KDSH.
.....C....MIM.Q.Q.....A.....

```
                                        1710      1720      1730
1740      1750      1760      1770      1780      1790      1800

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
```

Dmel FAS transcript protein
SIRCLGLNGRFLEIGKFDLSNNSPLGMSVFLKNTSFHGILLDSVMEGEEEMQNQVVSLVAEGIKT
GAVVPLPTSVFNDQQVEQAFRFMASGKHIGKVVIK
Lclav FAS transcript protein
.V...AKD.....A....F.....ALFDTNGPEKKE..R..Y....S
...R...AT..TE..I..G.....A.....LL.
Gleg protein FAS
.V...ANG.....N.....T.....ALFDTDCPEKRE..KI.N....N
...R...STI..EN.I..G..Y..T.....LL.
Abil protein FAS
.V...ANG...C....V.....A....L.....T.....ALF.SDCSEKKE.MR..S...AN
...Q...ST.YGET.A.....LL.

```
Pmac protein FAS
.V.............................................D..............S
...R........EH...S...............V.

                                          1810      1820      1830
1840      1850      1860      1870      1880      1890      1900

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
VRDEEAGKKALQPKPRLINAIPRTYMHPEKSYILVGGLGGFGLELTNWLVTRGARYIVLTSR---
----SGVKTGYQGLMIRRWQERGVKVVIDTSDVTT
Lclav FAS transcript protein   I.....-
Q.ITR.AMKTVA.......N.D...V.........M..A..MI....KN......XXXXXXXX.I
R....A.C....R.M.ITIQ.S.C....
Gleg protein FAS               I....P-
N.IVPYS.KTVP.......N.....V.........M..A..MI....K..I....-------
..IR....S.C....M.M..NIHVC.H....
Abil protein FAS               I....S-
R.T.I.AIKTVT...K...DT..V..................I....KK......-------
..I.....SMC....RSQ..T.L.S.A.A.K
Pmac protein FAS
.....D.R.T.K.SS..V...............I................S...K.L..S..---
----..I.....S............L.......S

                                          1910      1920      1930
1940      1950      1960      1970      1980      1990      2000

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
AAGAKKLLENSNKLALVGGIFNLAAVLRDALIEDQTAKDFKTVADPKVTATKYLDQFSRDICTEL
DYFICFSSVSCGRGNIGQTNYGLANSAMERICEQR
Lclav FAS transcript protein
EV..DS..KEA....P...............M.NLEEDH..V.TL...NG.RN..AS.KKF.P..
.F.V..........M...............M...
Gleg protein FAS
LS..EE..KVC.RI.P.................NLDEGQ..A.VA..I.G..N..TA..SL.PS.
...VV.........M...............V.S.
Abil protein FAS
PE..RQ..NE.A..GPI.A...........FM.NLSEA..N..CK...D......AA...LAAN.
.H.V....I......A..S.......V......M.
Pmac protein FAS
DK.CQQ..I.A..F.................L....V...Q..CES..QG......Y..AM.S..
...V.............................L.

                                          2010      2020      2030
```

```
2040      2050      2060      2070      2080      2090      2100

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
QVSGFPGTAIQWGAIGDTGLVLENLGDNDTVIGGTLPQRMPSCLQTIDLFLQQPHPVVASMVVAE
KRKSD-QSAGVSLIATIANILGLRDTKNIQDGASL
Lclav FAS transcript protein
.AV.L..L.........V..I..TM.N.E.EV........A...M.M.S........L....L..
RQ.AGDS.SQ.N.LDAVG....IK.V.TVNMNN..
Gleg protein FAS
.GI.L..L.........V..I.DTM.N...EV........W...S.M.T........L....L..
.N.PTDSANQI..VDAV.....IK.....NVNN..
Abil protein FAS
.SV.L..L.........V..I..TM.G...EV......KIS..MA.M.I......A......L..
....G-GDNQ.K.TDAV.....IK...TVPAI...
Pmac protein FAS
.A.....L.............I......N..........T.......F........L.......
.....-..G.....SC................S...

                                                 2110      2120      2130
2140      2150      2160      2170      2180      2190      2200

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
ADLGMDSLMSAEIKQTLERNFDIVLSAQEIRQLTFGALKAMDGGADVKP--
AAAAPAAAAGVPEANITSGGSSRTASPMGDGTQVVFT-TSLIPTEAIVQ
Lclav FAS transcript protein
.........GT........Y.L........A....K.LELSS.SAEANEV.SQS..NSS-----
---------LTETDP.EFLFQCSG.EIV.PKSLI.
Gleg protein FAS
.........GT........G..L...P....N....K.MELSSD-------VSI..STS------
-----------ESQPENLLFQYSSNEIV.F.PL.K
Abil protein FAS
.........G..........Y.L..N.....A...AR.VELES.G------.ST....S------
------------DN..KNL.Q.V-DE.M..QLL.N
Pmac protein FAS
.........G..............M.P........IQ..QLS...ESSD--...S..SPV-----
-----RR.PSP..F.....M..S-.E.M..Q...R

                                                 2210      2220      2230
2240      2250      2260      2270      2280      2290      2300

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein
```

```
LDTKAPANSKQSPIFFISPIEGFASALEPLAKRLEVPAYGLQYTEAVPSDSLESAAKFFIKQLRT
VQPKGPYKLAGYSFGCLLTYVMAGILEETNEVANV
Lclav FAS transcript protein    .---
QSTSE.GE...V.HA...VV.S.KS..SE..R.VW...C.KDA.L..IPNL.TYY.QEMK..KKQ
...SII.....ACVAFE..LQ..KAG.T.EL
Gleg protein FAS                F---
DSIGIGKP...MVHA...SVAG.KL..SA.NITVF.I.C..D..L..IPEL.AHYV.LMTS..KV
...R.F.....ACVAFE..LQM.AIGHKLDL
Abil protein FAS
MNNVDSTE..KT...ILH....AVTI.KKF.QEIQA.V..I.C...A.LS.INDL..YY.E.IK.
M......T.I.....ACVGFE.GIQ..AM..KVKL
Pmac protein FAS
.QSA...E..KR.L.VV.......D..KQ..S..DC.V....C.AEANLE.IDTL.D.YL..I..
..AR...AI....Y.A.VG..IVLH..KMK.N.RL

                                          2310      2320      2330
2340      2350      2360      2370      2380      2390      2400

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....|....|....|
Dmel FAS transcript protein     IMLDGAPSYVNWYTSSFKQRYTDGTNADNDNQ--
SYGLAYFGIVL-ANIDYKALVRLLIVIPTWEEKLERFAELMSNEITQPVE---------------
-
Lclav FAS transcript protein
TL...S.DFIKLHSQTIGKQTNQVSRDLAS.DGFQKAI.F.ARQ.NSD.SFIKAYEI.RGSKSED
.T.NKMI..IG.T-PFKS.----------------
Gleg protein FAS                .L...S.EFITLHSTLINKQVSPDNSELQT-
DGCRKS..F.IKQFNR..N.TNAYKS.QEVKDE-.IFDKMI..IGPT-SLDID------------
----
Abil protein FAS                FLI..S.T..ATH.GKAS..IQP.NT.AEH----
TEA.LF.MHQF-KEV.QQKTAAE.MALK.LD.RAKLTTQIIGDACPF.K.---------------
-
Pmac protein FAS                V......K......TN....L---NTS.DQ.E--
A.......M.V-
.....SLVAKV.LN....DS.VAKC..IVAA..N..TDLVSNIFGKNKKKKRST

                                          2410      2420      2430
2440      2450      2460      2470      2480

....|....|....|....|....|....|....|....|....|....|....|....|....|
....|....|....|....|....
Dmel FAS transcript protein     ---
TIKKSATLFYKKLELADGYQPTLKLKTNVTLVKPTDNSAKLDEDYRLKEVCTKPVEVHTVEGNHR
TFLIEDQSLKTIQSILKRLFN
Lclav FAS transcript protein    ---
DL.IAGI......RA.NM.KASN.YNGPI..I.AK..FVS.NN..G.S.I.RQT.RIEELP....
SI.-SGE.V.KMATLV.T---
```

```
Gleg protein FAS                ---
DL.MAGY.L....RA.NL.R.SG.F.GP.Q.I.AN.AFIHMS...G.SQ----------------
---------------------
Abil protein FAS                ---
Q.TAA.KS..Y..KA..M.K.AS.FNG.II.A.AN..YVQGES..G.SN.SL-------------
---------------------
Pmac protein FAS
KNVM.VQA..S.....LA..K.V.SI.VSCD.......E.Y...E...G.N..---N.DL-----
------------------------
```