The co-evolution of herbivore impact: field estimates of impact in older and newer gallmaker–Solidago interactions

Zoryana Shibel¹, Stephen D. Hendrix² and Stephen B. Heard¹

¹Department of Biology, University of New Brunswick, Fredericton, New Brunswick, Canada and ²Department of Biology, University of Iowa, Iowa City, Iowa, USA

ABSTRACT

Background: Novel associations of specialist insect herbivores with host plants are frequently assembled. If herbivore and host co-evolve following their first interactions, we would expect increases over evolutionary time in plant tolerance to attack, and usually decreases over time in insect virulence (because most insect herbivores depend for resources on continued survival of their hosts). The result should be a pattern of decreased herbivore impact through time. This is the ‘evolution-of-impact hypothesis’.

Organism: Two specialist gallmakers of the goldenrods Solidago altissima L. and S. gigantea Ait.: Rhopalomyia solidaginis/capitata (Diptera: Cecidomyiidae), and Eurosta solidaginis (Diptera: Tephritidae).

Site of experiments: Tallgrass prairie at two sites in Iowa, USA.

Predictions: For both gallmakers, S. gigantea is the novel host and S. altissima the older one. We therefore predict a higher impact of both gallmakers on S. gigantea than S. altissima. However, we predict a similar impact on each species of mechanical damage, with which both hosts have long evolutionary experience.

Methods: We measured the aboveground biomass of galled and ungalled ramets from the field to assess herbivore impact by each gallmaker on each species. We used stem width–biomass relationships for ungalled ramets, applied to stem widths of galled ramets, to estimate the potential biomass of galled ramets in the absence of attack. We calculated impact as potential minus actual biomass. We also re-analysed similar data for a third gallmaker of the same hosts, Gnorimoschema gallerae-solidaginis (Lepidoptera: Gelechiidae) from Heard and Kitts (2012). We assessed impacts of mechanical damage with field and greenhouse experiments of clipping and defoliation damage imposed on ramets of each species.

Results: All three gallmakers had a higher impact on the novel host (S. gigantea) and, combined across all datasets, the pattern of higher impact on the novel host was highly significant. Impacts of mechanical damage did not differ between host species.

Conclusions: Patterns in impacts of goldenrod gallmakers on their hosts are consistent with the evolution-of-impact hypothesis.

Keywords: co-evolution, gallmakers, herbivore impact, plant–insect interactions, Solidago.
INTRODUCTION

Most plants experience attack by herbivorous insects – attack that can have severe consequences for plant fitness and, at least sometimes, for plant population dynamics (Ancheta and Heard, 2011). The ecological and economic importance of such damage has motivated a vast body of work on insect–plant interactions, including the study of the mechanisms and extent of herbivore-caused damage and of the ability of plants to grow and reproduce despite that damage (e.g. for the Solidago we study here: Goldberg, 1988; Carson and Root, 1999; Hakes and Cronin, 2011a, 2011b, 2012; Bode and Kessler, 2012; March et al., 2013).

Most insect herbivores feed on restricted sets of host plants (Bernays and Graham, 1988), and many are extreme specialists – attacking, in many cases, a single organ of a single host plant species at a single phenological stage. Diet specialization is typically achieved by means of sophisticated sensory and behavioural mechanisms, and this might lead us to expect a high degree of constancy in herbivore–host associations. However, over evolutionary time and even in the shorter term, herbivore–host associations can be quite labile, with herbivores attacking new plant species and plants coming under attack by new herbivores. New herbivore–host associations can arise via range expansion by either insects or plants. Invasions also lead to new herbivore–host associations, either as invasive insects attack native plants, or as invasive plants come under attack by native insects in their new ranges. Finally, even without conspicuous distributional changes, herbivorous insects may make host shifts that add new plant species to their diets [often leading to host-associated genetic differentiation and the origin of new diet specialists (e.g. Stireman et al., 2005)].

The frequent reassortment of herbivores and hosts means that some herbivore–host pairs are evolutionarily well acquainted, while others are strangers recently met. Older and newer herbivore–host associations are likely to differ in many ways, since herbivore and host will exert selective pressure on each other and co-evolution should often result. In particular, the impact of herbivore attack on plant fitness might be expected to differ between older and novel herbivore–host pairs. This impact is determined in part by herbivore virulence (the amount of damage done to plant tissue per attacking herbivore), and in part by plant tolerance (the extent to which the plant can maintain fitness despite of damage done). Each has an evolutionary optimum (defined with respect to the appropriate species): optimal virulence maximizes insect survival and reproduction in the available plant landscape, while optimal tolerance minimizes the impact on plants of expected insect attack. Both virulence and tolerance are likely to change as herbivores and hosts co-evolve. In particular, each is likely to begin away from its fitness optimum in a new association, and then evolve towards the optimum as selection acts on both insect and plant as they interact. This suggests change in herbivore impact as a new association ages.

Predictions for the evolution of virulence and tolerance in plant–insect systems can be derived from host–parasite co-evolution theory (Heard and Kitts, 2012). Although this body of theory has been largely developed and tested for diseases and parasites of animals (Ewald, 1994; Read, 1994; Frank, 1996), nearly all plant-feeding insects are in fact classical 'parasites' in the sense that they feed on living host tissue without directly killing the host, and so depend on host survival and performance while they feed. Host–parasite theory thus motivates our work and provides the context for the hypothesis we test.

Selection should, all else being equal, favour increased plant tolerance (Roy and Kirchner, 2000; Agrawal and Fishbein, 2008). One important way that all else may not be equal is that tolerance may be costly in the absence of herbivore attack, in which case tolerance might decline when
attack is rare (Strauss and Agrawal, 1999). However, the costs of tolerance seem modest for many plants (e.g. Núñez-Farfán et al., 2007; Hakes and Cronin, 2011b; but see Stinchcombe, 2002). Selection on virulence may be more complex, with fitness either increasing or decreasing with virulence (Ewald, 1994; Frank, 1996; Alizon et al., 2009). Optimal virulence can be high when an individual host is attacked by many parasite individuals, when host debility or death favours parasite growth or dispersal, or when each parasite attacks many individual hosts throughout its lifetime. Each of these circumstances applies to some herbivore–host interactions. Lepidopteran pests of trees may attack in huge numbers; such attack can have severe impacts including, at least in conifers, tree mortality. Many bark beetles need to overwhelm tree defences before securing access to phloem tissue in a tree rendered moribund by this attack. Individual migratory locusts may feed on many individual hosts as they travel over long distances. However, cases like these are likely to be exceptions rather than the rule. Most insect herbivores have life histories in which their own fitness depends on survival and continued performance of the host under attack (Heard and Kitts, 2012). In the extreme, leafminers, gallmakers, and other internal feeders normally spend their entire larval period feeding in a single host individual, and are unable to leave even if their first host is too damaged to survive. Optimal virulence for such parasites should be low.

At least for most specialist herbivores, then, following association with a new host, selection should drive increases in plant tolerance and reductions in herbivore virulence. Both of these co-evolutionary trends will work to lower the impact of herbivory on the host plant, and we therefore predict that herbivore impact should be lower in older associations than in evolutionarily novel ones. This is the ‘evolution-of-impact’ hypothesis.

In contrast to co-evolved specialist insect–plant interactions such as those involving gallmakers, all plants also experience generalized types of damage in which portions of the plant are removed by abiotic events (wind storms, freezes, trampling, etc.) or by generalist herbivores. Most species will have extensive evolutionary experience with such damage. Host–parasite theory predicts no consistent differences in impact between species for such ancient forms of damage, which therefore provide a control against which to test the evolution-of-impact hypothesis.

Heard and Kitts (2012) were the first to test (and support) the evolution-of-impact hypothesis for herbivorous insects, but they reported data for only a single herbivore–host pair: the lepidopteran gallmaker *Gnorimoschema galleasolidaginis* on the goldenrods *Solidago altissima* and *S. gigantea*. Here, we extend this test to two additional gallmakers on the same hosts: the dipteran gallmakers *Eurosta solidaginis* and *Rhopalomyia solidaginis*/*capitata*. Congruent with predictions from host–parasite theory, we find a consistent pattern across field sites and years: a greater impact of gallmaking on the novel host than on the older host.

**METHODS**

**Study system**

We studied gallmaker impacts on the goldenrods *Solidago altissima* L. and *Solidago gigantea* Ait. (Asteraceae). These goldenrods range across much of eastern and central North America, where they often co-occur at high densities in prairies, old fields, and disturbed habitats (Abrahamson and Weis, 1997). *Solidago altissima* and *S. gigantea* are perennials that reproduce sexually through seed and asexually via rhizomes (which persist through
winter and produce new shoots the following spring). These species play host to many herbivores, including the gallmaking specialists *Eurosta solidaginis* (Fitch) (Diptera: Tephritidae), *Gnorimoschema galleasolidaginis* (Riley) (Lepidoptera: Gelechiidae), and *Rhopalomyia solidagnis* (Loew) (Diptera: Cecidomyiidae). Each of these gallmakers has undergone a host shift from *S. altissima* to *S. gigantea*, thus making those goldenrods the older and novel hosts, respectively, for each species (Stireman et al., 2005). The *Rhopalomyia* host shift is the oldest, dating to at least 2 million years ago, and the host-specialist forms are cryptic species (with the name *R. capitata* (Felt) applicable to flies from *S. gigantea*). The *Eurosta* races, in contrast, experience appreciable gene flow (Itami et al., 1998) and the *Eurosta* host shift is much more recent, dating to at most 200,000 years ago. *Gnorimoschema*’s host shift is intermediate in age (Stireman et al., 2005).

All three gallmakers induce stem galls, although they vary in timing of attack, gall morphology, and impact. *Gnorimoschema* galls (Miller, 2000) are initiated very early in the growth of new ramets as overwintered caterpillars chew down from the growing tip, often damaging or destroying the terminal meristem as they pass. Galls are spindle-shaped, hard-walled, and hollow; larvae feed within them until pupation and emergence in the fall. *Eurosta* galls (Abrahamson and Weis, 1997) are initiated when females oviposit directly into stems in mid-season, usually just below the apical meristem (with meristems often sustaining incidental damage). The result is a ‘ball gall’ of largely solid tissue, 2–4 cm in diameter, in which the larva feeds and overwinters. *Rhopalomyia* galls (Gagné, 1989) are initiated when females oviposit one to perhaps a dozen eggs into a plant’s terminal bud, again in mid-season (McEvoy, 1988; Wise et al., 2006). Larval feeding provokes formation of a ‘bunch gall’ of many closely clustered leaflets providing multiple chambers for larvae; this typically aborts further stem growth and usually suppresses flowering. *Rhopalomyia* may also have an earlier-season generation with single-chambered galls, but our collections focused on the multi-chambered gall.

**Defining and measuring herbivore impact**

There are many ways of measuring the impact of herbivores on plants. Given that our central hypothesis is an evolutionary one, we should be measuring impact on plant fitness. This is a major challenge in plant ecology, however, because fitness has three components: male fitness (through pollen), female fitness (through seed), and asexual fitness (in *Solidago*, through rhizomes) – and none is straightforward to measure. Many studies with *Solidago* have attempted to measure female fitness, usually indirectly by measuring inflorescence mass (e.g. Hakes and Cronin, 2011a, 2011b) or estimating seed production from capitula counts (e.g. March et al., 2013). Fewer have estimated asexual fitness. Because rhizomes are exceptionally difficult to excavate in mature turfs, most rhizome data come from greenhouse studies (e.g. Shibel and Heard, 2016), from outplants in tilled soil (e.g. Goldberg, 1988), or from recently colonized fields (e.g. Wise et al., 2006). All these methods are tedious, and therefore constrain sample sizes, in field situations. Fortunately, aboveground biomass is much easier to measure and there is good theoretical reason to expect it to correlate with fitness: it constitutes the photosynthetic tissue that supports production of all plant tissues. Empirically, in most plants aboveground biomass correlates well with more direct estimates of fecundity (Younginger et al., 2017), and this has proven true for *Solidago* as well (e.g. Goldberg, 1988; Hakes and Cronin, 2011b, 2012). We therefore measured aboveground (dry) biomass in both greenhouse and field experiments.
Impacts of mechanical damage: greenhouse and field experiments

We begin with our experiments measuring impacts of mechanical damage to stems and leaves (clipping and defoliation). Our Solidago species have extensive and presumably equal evolutionary experience with these ancient types of damage (simulating effects of trampling, storms, mammalian grazing, etc.), and the evolution-of-impact hypothesis predicts no difference in impact between species. We conducted two greenhouse clipping studies, and one field study involving both clipping and defoliation. ‘Clipping’ consisted of removing the terminal one-quarter of a plant’s stem with pruning shears. Because it removes apical meristems, clipping is similar in some ways to gallmaker attack, which frequently damages those meristems; but it is both more rapid and more severe. ‘Defoliation’ consisted of removing every other leaf on the plant’s main shoot. This amount of damage is not unusual for outbreaks of oligophagous insect herbivores such as Trirhabda spp. [Coleoptera: Chrysomelidae (McBrien et al., 1983; Carson and Root, 1999); both S. altissima and S. gigantea experience such defoliation (Messina, 1982).

We ran greenhouse clipping experiments in 2012 and 2013. In 2012, we dug up 45 S. altissima and 49 S. gigantea rhizomes from a field in Fredericton, New Brunswick, in October around the time of senescence of most ramets in the field. In older stands of Solidago, genets are likely intermixed, so we cannot know how many ramets were of unique genotypes. However, our collection area spanned >100 m, so we assume that we had members of many different genets. We stored rhizomes briefly (about one week) at 4°C, before planting them in 13 × 13 × 15 cm pots in ‘Pro-mix BX’ potting mix (Premier Tech, Rivière-du-Loup, Québec). We randomized the pots to bench positions and added 5 mL of 14-14-14 slow-releasing fertilizer (Nutricote) per pot. After each plant reached 50 cm in height, we assigned it to either the clipping treatment or a control group (in alternation; this procedure ensures that variation in plant vigour arising from variation in size of planted rhizomes is not confounded with treatment). At the time of treatment, we also measured basal stem width (approximately 5 cm above soil) of each ramet to the nearest 0.01 mm. By day 105, most plants were beginning to senesce, and we re-measured stem width and then collected aboveground biomass by clipping the stems at soil level. Collected ramets were dried for 48 hours at 70°C. We measured dry mass for each ramet to the nearest 0.01 g and once again measured basal stem width, this time on the dried stem. In the summer of 2013, we repeated this experiment with 43 individuals of S. altissima and 35 individuals of S. gigantea. We tested for effects of species, treatment, and their interaction on aboveground biomass using two-way analysis of variance (ANOVA); a significant interaction would indicate a difference between species in tolerance to damage.

We ran our field clipping + defoliation experiment in 2013, using 53 ‘triplets’ of S. altissima and 53 of S. gigantea located and tagged in an old field in Fredericton, New Brunswick in early summer. A triplet consisted of three neighbouring, similar-sized ramets. After ramets reached 50 cm in height, we randomly chose one member of each triplet for clipping and one for defoliation (the third ramet acting as an undamaged control). At day 105 after treatment, we clipped all ramets at ground level and processed them as for the greenhouse experiments. For each damage treatment, we calculated the difference in aboveground biomass between damaged and control plants within each triplet. We then used one-way ANOVA to compare the impact of damage between plant species.
Measuring gallmaker impact in field samples

We measured gallmaker impact by comparing aboveground biomass of naturally galled and ungalled ramets in the field. Experiments manipulating galling would, of course, be preferable. Such experiments have been conducted with *Eurosta* (e.g. Abrahamson et al., 1989b; Walton et al., 1990), although they are logistically very difficult and have been executed primarily in the greenhouse or via mass release of flies into field cages. For *Rhopalomyia*, no techniques for manipulating gallmaker oviposition are available. We used observational methods for both gallmakers because this let us achieve useful sample sizes for multiple field sites in multiple years. Heard and Kitts (2012) did the same for *Gnorimoschema* impact, for the same reasons [experiments manipulating gall formation have been attempted without success (G.C. Cox and S.B. Heard, unpublished data)].

We present new estimates of gallmaker impact for *Eurosta solidaginis* and *Rhopalomyia solidaginoides* and compare them with previously published estimates for *Gnorimoschema gallaesolidaginis* (Heard and Kitts, 2012). Our analysis of gallmaker impact involves several steps, with later steps depending on results of earlier ones. For clarity, we summarize intermediate results in this section when they are needed only for our impact estimates (the intermediate results themselves can be found at evolutionary-ecology.com/data/3094Appendix.pdf). All statistical analyses were conducted in R v.3.3.1 ‘Bug in Your Hair’ (R Development Core Team, 2016).

Collections

We collected galled and ungalled goldenrods of each species at two sites in Iowa, USA. At the first site, Kent Park (23 km west-northwest of Iowa City, IA; 41.73N, 91.729W), *S. altissima* dominated a dry hillside prairie, while *S. gigantea* occurred mostly in an adjacent moist swale. At the second site, Iowa Lakeside Laboratory (about 100 km northeast of Sioux City, IA; 43.38N, 95.18W), the goldenrod species were well interspersed in a mesic old field. We collected at Kent Park in late August 2012, and at both sites in September 2013 and 2014. For each species and each gallmaker, in each year, we collected approximately 45 goldenrod triplets, with each triplet consisting of a galled ramet and the two closest ungalled ramets of the same species. An analysis of triplet neighbours removes local environmental and some genetic variation since the plants of a triplet share a common habitat (and if they belong to a single genet, also the same genotype).

We chose ramets for collection by walking through a site and inspecting the area for galls. Once we sighted a galled ramet, we collected the nearest other galled ramet (along with its ungalled neighbours). This procedure avoided collecting bias for larger galls or for galls on larger plants, which are more evident when searching. We did not collect ramets with more than one gall. Ramets were clipped at ground level, returned to the laboratory, and processed for aboveground biomass data as described for our mechanical damage experiments, including measurement of post-drying basal stem width (we did not measure fresh stem width for our field collections).

We did not score survival or parasitism of gall inhabitants. While both gallmaker survival and parasitism may differ between *S. altissima* and *S. gigantea*, neither host provides a consistent advantage (Heard et al., 2006). In any case, once a gall is established, most of the damage is done regardless of the subsequent fate of the gallmaker larva.
A complication: size-selectivity in gallmaker attack

The simplest analysis of gallmaker impact would directly compare biomasses of galled and ungalled plants. However, such an analysis is problematic if gallmakers choose ramets for attack non-randomly with respect to plant size (Crutsinger et al., 2008). Preferences for attacking larger ramets or modules are widespread among insect herbivores, including gallmakers (Cornelisson et al., 2008), and they exist in our system: both *Eurosta* and *Gnorimoschema* females prefer larger ramets (Walton et al., 1990; Heard and Cox, 2009). For our third gallmaker, *Rhopalomyia*, ramet-size preferences have not been measured but are likely. Because of these preferences, simple comparisons would underestimate herbivore impact, as galled plants will have started larger. If size preferences were strong enough, galled ramets could even be larger than ungalled ones, even after growth reduction from attack; but we would not conclude from this that gallmakers increase plant growth.

If size preferences are equally strong for gallmakers attacking the two host species, impact estimates will be inaccurate for either species, but the difference between impacts on older and novel hosts will be correctly estimated. This was the case for *Gnorimoschema* in Heard and Kitts’ (2012) analysis, and we use those comparisons here without further adjustment. However, if the *altissima* and *gigantea* host forms exert different size preferences, then neither individual-species impact nor the difference between impacts on older and novel hosts will be estimated correctly. We cannot rule out such differences for either *Eurosta* or *Rhopalomyia*, but for these species we can correct for the gallmaker preference problem using data on stem width. Our stem-width correction depends on the hypothesis that early-season stem width predicts both late-season stem width and late-season aboveground biomass, for both galled and ungalled ramets. If this hypothesis holds, we can use these relationships to estimate ‘potential biomass’, by which we mean the aboveground biomass to which a galled ramet would have grown in the absence of attack. Testing this hypothesis involves combining stem-growth data from our mechanical-damage experiments (to represent the potential effect of gallmaker attack on growth in stem width) with stem width data from our field impact collections. This is a three-step process.

We start with a technical detail: we need to know how drying changes stem width. This is because we assessed stem growth in our mechanical-damage experiments from fresh stem widths before and after damage – but we needed to apply this result to our gallmaker-impact collections, for which we measured stem widths after drying. In the mechanical-damage experiments, we measured dried as well as fresh stem width at harvest, and fresh stem width predicted dry stem width extremely well (experiments separately, all \( r^2 > 0.82 \); experiments pooled, \( r^2 = 0.91 \); all \( P < 0.0001 \); Appendix, Table A1). We used the pooled regression equation, \((\text{dry width}) = 0.79 (\text{fresh width}) + 0.07\), to convert fresh to dry stem width in all subsequent calculations.

Next, we asked whether, for ungalled plants from our Iowa field collections, stem width at harvest was a good predictor of aboveground biomass. Because mass is likely to scale non-linearly with diameter, we regressed \( \log_{10}(\text{biomass}) \) against \( \log_{10}(\text{stem width}) \), finding strong relationships for all ten datasets \((0.73 < r^2 < 0.89\) and all \( P < 0.0001 \); Appendix, Table A2).

Finally, we asked whether gallmaker attack might reduce growth in stem width over the season. This is important because our procedure uses end-season stem width to estimate a galled ramet’s potential mass in the absence of attack – and so any loss of stem width due to attack needs to be corrected for in the estimation. The impact of gallmaking on stem-width
growth could be estimated directly with pre- and post-attack diameter data for galled ramets, but we were unable to gather pre-attack data. Instead, we had pre- and post-treatment diameters for our mechanical damage experiments, in which we used clipping as an analogue of gallmaker damage. Gallmakers often destroy or severely damage apical meristems, while clipping removes them completely (see also MacDonald and Kotanen, 2010; Spirko and Rossi, 2015); thus, clipping should yield something of a worst-case estimate for effects of gallmaker attack.

We calculated stem growth as the difference in stem width for each ramet between the day of treatment and the day of harvest. We pooled data from the three clipping experiments and compared stem growth between control and clipped plants and between *S. altissima* and *S. gigantea* using two-way ANOVA. Stem growth was reduced by clipping, but the impact of clipping on growth did not differ between species (non-significant treatment × species interaction; 3094Appendix, Table A3 and Fig. A1). The mean effect of clipping was a 31% reduction in stem-width growth.

To avoid underestimates of potential biomass for our galled ramets, we applied a stem-width correction to the observed stem width of all galled stems from our field collections. To implement this correction, we regressed end-season on pre-treatment stem widths for clipped and for control ramets in our mechanical-damage experiments. We ran separate regressions for the two *Solidago* species, but pooled data across years. We then corrected each galled stem width from our field collections using the ‘clipped’ and ‘control’ regression coefficients. In particular, we took the end-season stem width for the galled ramet, and estimated its likely early-season width using the ‘clipped’ regression. We then used the ‘control’ regression to estimate its late-season width had it not been galled (in other words, its potential, unattacked, stem width). Potential stem widths were on average ∼6% larger than observed widths. The stem-width correction is outlined in more detail in 3094Appendix (Table A4 and associated text).

**Estimating gallmaker impact**

Our estimate of gallmaker impact is based on the difference between each galled ramet’s potential aboveground biomass, had it not been attacked, and its actual biomass. For each galled ramet, we calculated potential stem width using the stem-width correction, and from that calculated potential biomass using the appropriate (for the year and field site) ungalled stem width–biomass regression (from 3094Appendix, Table A2). This estimate can be improved, however, because of our triplet design. Each galled ramet has two ungalled nearest neighbours, and the residuals of those ungalled neighbours from their ungalled stem width–biomass regression line reflect influences on their growth of genotype and microsite. A positive residual, for example, might derive from a robust genotype or favourable microsite conditions. Galled ramets are likely to share these factors with their ungalled neighbours (genotype because goldenrod is clonal, and microsite because neighbours were usually just a few centimetres apart). Therefore, we add the average residual for the two ungalled neighbours to our regression estimate of the galled ramet’s potential biomass. Our impact estimate for the galled ramet, finally, is the difference between the ramet’s actual biomass and this neighbour-corrected potential biomass.

**Comparing impact estimates**

We calculated gallmaker impacts for each of our 10 datasets (10 gallmaker species × site × year combinations). For each dataset, we then asked whether impact differed significantly
between the older host (*S. altissima*) and the novel host (*S. gigantea*), using one-way ANOVA. We also re-estimated herbivore impact for *Gnorimoschema gallaesolidaginis*, using data from Heard and Kitts (2012). That study used paired (rather than triplet) galled and ungalled ramets. Because *Gnorimoschema*’s ramet-size preference appears not to differ between host species (Heard and Cox, 2009), the stem-width correction is unnecessary. Instead, we simply calculated the difference between aboveground dry masses for galled and ungalled ramets within pairs, and used a one-way ANOVA to compare this impact estimate between goldenrod species.

To test the more general hypothesis that impact differs between older and novel interactions, we combined results from multiple datasets using the weighted (by df) $Z$-method for combining $P$-values (Whitlock, 2005) using the R package ‘metap’ (Dewey, 2016). We converted our two-tailed $P$-values to one-tailed values before combining them, and converted the result back to a two-tailed $P$-value (see Whitlock, 2005), although omitting this complication would not change our interpretation of the results.

**RESULTS**

**Impact of mechanical damage**

We ran three clipping experiments (two in the greenhouse, one field) and one defoliation experiment (in the field), and in none of them could we detect any difference in the impact of damage between goldenrod species (Table 1; non-significant species × clipping interactions for greenhouse experiments, both $P \geq 0.08$, and species main effects for field experiments, both $P \geq 0.43$).

**Impact of gallmaker attack**

The impact of each gallmaker was greater on the novel host (*S. gigantea*) for all site × year combinations (Fig. 1; grey boxes are above white boxes in each panel). The difference was individually significant in four cases (two each for *Eurosta* and *Rhopalomyia*; Table 2; entries with $P$-values in bold). Our re-analysis of Heard and Kitts’ (2012) data for *Gnorimoschema* confirmed their finding of a significantly higher impact on *S. gigantea* in both site × year combinations (Fig. 2, Table 3). The magnitude of the impact difference varied considerably from site to site, year to year, and gallmaker species to species, but was often substantial (Figs. 1, 2; compare offsets between grey and white boxes). Combining our results over gallmakers, years, and sites using the weighted $Z$-method (three gallmakers, 12 datasets) provides very strong evidence for greater herbivore impact on *S. gigantea* ($P = 3.2 \times 10^{-14}$).

**DISCUSSION**

We compared the impact of three specialist gallmakers on their evolutionarily older and (relatively) novel *Solidago* host plants. As predicted by the evolution-of-impact hypothesis, gallmaker impact was consistently greater on the more novel host, *S. gigantea* (Figs. 1, 2). While our results strongly suggest that, in the *Solidago* system, herbivore impact is more severe in novel plant–herbivore interactions than in older ones, we recognize that there are
Table 1. Impact of mechanical damage compared between *S. altissima* and *S. gigantea*

(A) Greenhouse clipping experiments

<table>
<thead>
<tr>
<th>Year</th>
<th>Factor</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>Species</td>
<td>1</td>
<td>0.02</td>
<td>0.001</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Clipping</td>
<td>1</td>
<td>42.0</td>
<td>1.74</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Species × Clipping</td>
<td>1</td>
<td>77.0</td>
<td>3.19</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>71</td>
<td>24.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>Species</td>
<td>1</td>
<td>0.2</td>
<td>0.006</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>Clipping</td>
<td>1</td>
<td>386</td>
<td>12.7</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>Species × Clipping</td>
<td>1</td>
<td>35.1</td>
<td>1.15</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>89</td>
<td>30.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(B) Field clipping and defoliation experiments

<table>
<thead>
<tr>
<th>Damage type</th>
<th>Factor</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clipping</td>
<td>Species</td>
<td>1</td>
<td>0.0108</td>
<td>0.275</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>99</td>
<td>0.0395</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defoliation</td>
<td>Species</td>
<td>1</td>
<td>0.0409</td>
<td>0.632</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>93</td>
<td>0.0647</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: In the greenhouse experiments (A), the test for species differences in impact of clipping is the species × clipping interaction. In the field experiments, the test for species differences in the impact of clipping or defoliation is the species main effect.

a Two-way ANOVAs on aboveground biomass.

b One-way ANOVA of (log-transformed) damage impact (control minus damaged biomass).

Table 2. Impact of herbivory compared between older (*S. altissima*) and novel (*S. gigantea*) hosts, for *Eurosta* and *Rhopalomyia* gallmakers

<table>
<thead>
<tr>
<th>Gallmaker</th>
<th>Year/site</th>
<th>Factor</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eurosta</em></td>
<td>2012 Kent</td>
<td>Species</td>
<td>1</td>
<td>49.8</td>
<td>6.58</td>
<td><strong>0.0125</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>70</td>
<td>7.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2013 Kent</td>
<td>Species</td>
<td>1</td>
<td>6.46</td>
<td>1.58</td>
<td>0.212</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>87</td>
<td>4.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2014 Kent</td>
<td>Species</td>
<td>1</td>
<td>1.20</td>
<td>0.739</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>87</td>
<td>1.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2013 Lakeside</td>
<td>Species</td>
<td>1</td>
<td>0.225</td>
<td>0.320</td>
<td>0.573</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>86</td>
<td>0.703</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2014 Lakeside</td>
<td>Species</td>
<td>1</td>
<td>23.5</td>
<td>38.0</td>
<td><strong>1.7 × 10^-8</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>95</td>
<td>0.617</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Rhopalomyia</em></td>
<td>2012 Kent</td>
<td>Species</td>
<td>1</td>
<td>4.89</td>
<td>0.481</td>
<td>0.491</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>55</td>
<td>10.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2013 Kent</td>
<td>Species</td>
<td>1</td>
<td>1.72</td>
<td>0.334</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>85</td>
<td>5.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2014 Kent</td>
<td>Species</td>
<td>1</td>
<td>3.71</td>
<td>1.50</td>
<td>0.225</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>87</td>
<td>2.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2013 Lakeside</td>
<td>Species</td>
<td>1</td>
<td>12.6</td>
<td>16.0</td>
<td><strong>1.35 × 10^-4</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>88</td>
<td>0.792</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2014 Lakeside</td>
<td>Species</td>
<td>1</td>
<td>25.4</td>
<td>20.0</td>
<td><strong>2.28 × 10^-5</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>89</td>
<td>1.27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Tests reported are one-way ANOVAs on the estimated effect of gallmaking on aboveground biomass.
Fig. 1. Impact of herbivory by *Eurosta* (top panels) and *Rhopalomyia* (bottom panels) gallmakers on *Solidago altissima* and *S. gigantea*. Herbivore impact is the inferred reduction in aboveground (dry) biomass caused by gallmaker attack (for details, see Methods – Estimating gallmaker impact). Centre line is median; box shows 25th and 75th percentiles; whiskers show 10th and 90th percentiles, and dots show 5th and 95th percentiles. For bottom left plot, the *altissima* 5th percentile (not shown) is at −9.8.

Fig. 2. Impact of herbivory by *Gnorimoschema* gallmakers on *Solidago altissima* and *S. gigantea*. Data re-plotted from Heard and Kitts (2012) using uncorrected differences in (dry) aboveground biomass (for details, see Methods – Comparing impact estimates). Boxplot conventions as in Fig. 1.
reasons for caution in interpreting and generalizing our results. We comment here on two such reasons.

First, our three gallmakers all made their host shift in the same direction (with *S. gigantea* the novel host for each). It is therefore conceivable that *S. gigantea* experiences higher impact not because it is the novel host, but simply because it is less tolerant of damage of any kind – regardless of mechanism and, especially, of evolutionary experience. However, our field and greenhouse mechanical-damage experiments found no differences between the species in their response to damage by defoliation or clipping. This result is inconsistent with the general-intolerance hypothesis but consistent with the evolution-of-impact hypothesis. An obvious next step would be to add data for herbivores that made the opposite host shift (from *S. gigantea* to *S. altissima*). Unfortunately, no such herbivores are yet known. Analogous tests should, therefore, be pursued for specialist herbivores attacking other host–plant pairs.

Second, for both *Eurosta* and *Rhopalomyia*, the difference between impact on older and novel hosts varied considerably among sites and years (Fig. 1). We expect that substantial spatiotemporal variation exists for *Gnorimoschema*, too, although we could not test for it because Heard and Kitts (2012) provided data for only two sites in one year. This kind of variation in interaction strength is, of course, no surprise to ecologists (Stinchcombe and Rausher, 2002; Thompson, 2005), and it is not incompatible with the evolution-of-impact hypothesis. However, high spatial and temporal variation make extrapolation difficult. Our 2012 data (for Kent Park; Fig. 1, left panels) stand out in particular, with much larger variance among triplets in estimates of impact, and with negative estimates for substantial numbers of triplets (that is, estimates of potential ungalled biomass smaller than actual galled biomass). This may reflect drought conditions: 2012 was one of the driest and hottest years on record for Iowa (Hillaker, 2013), with precipitation 228 mm (24%) below average and temperatures 2.1°C above average. Heat and drought were particularly intense in June and July, when gallmakers and goldenrods both grow rapidly. The effects of drought are likely to be quite variable across microhabitats. Interacting with this is the fact that for *S. altissima* (but not for *S. gigantea*) drought stress acts synergistically with clipping to reduce plant growth (Shibel and Heard, 2016). However, an alternative analysis that omits the 2012 Kent Park dataset still shows strong overall evidence for a higher impact on *S. gigantea* (weighted Z on 10 datasets instead of 12 gives $P = 5.4 \times 10^{-13}$).

While *S. gigantea* is the more recent host for all three gallmakers, it has nonetheless been under attack for some time. The host shift from *S. altissima* to *S. gigantea* was most recent

<table>
<thead>
<tr>
<th>Site/year</th>
<th>Gallmaker</th>
<th>Factor</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fredericton 2003</td>
<td><em>Gnorimoschema</em></td>
<td>Species</td>
<td>1</td>
<td>74</td>
<td>4.67</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>134</td>
<td>15.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toronto 2003</td>
<td><em>Gnorimoschema</em></td>
<td>Species</td>
<td>1</td>
<td>79.7</td>
<td>7.03</td>
<td>0.0088</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>159</td>
<td>11.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Data from Heard and Kitts (2012), re-analysed as separate one-way ANOVAs by site (raw galled minus ungalled dry masses).
for *Eurosta* (less than 200,000 years), and substantially longer ago for *Gnorimoschema* and especially *Rhopalomyia* – for the latter, likely at least 2 million years (Stireman et al., 2005). The apparent long persistence of higher impact raises interesting questions about selection and about responses to selection by both insects and plants. In particular, it suggests that selection and/or response may be relatively weak, such that a novel association takes a long time to reach the levels of tolerance and virulence characterizing the corresponding older association.

Studies measuring selection on tolerance for *S. altissima* suggest that while selection varies in space and time, it is often weak (Hakes and Cronin, 2011b, 2012). This might be so for at least three reasons. First, if tolerance to different stresses is achieved via similar genetic or physiological pathways, host shifting by one herbivore may not greatly affect overall selective pressure on tolerance. Müller-Schärer et al. (2004) speculated that such pathway sharing might account for the apparent maintenance of high tolerance by several invasive plants freed from their specialist herbivores (including *Solidago* invasive in Europe). Second, the goldenrod system features an extremely diverse set of herbivores, parasitoids, and other interactors, and diffuse and/or fluctuating co-evolution may result and may slow adaptation to selection based on pressure from any single herbivore. Third, goldenrods reproduce clonally and gallmaker attack is relatively uncommon – often less, and sometimes much less, than 3% of available ramets (McCrea and Abrahamson, 1987; Abrahamson et al., 1989a; Root and Cappuccino, 1992; Crutsinger et al., 2008; S.B. Heard, unpublished data). Even if herbivore attack sharply reduces fitness of a single ramet (the scale at which tolerance is normally measured), selection pressure at the genet level may be weak. Alternatively, our goldenrods might experience strong selection on tolerance but not respond quickly if heritability is low. This is true for tolerance of leaf-chewing by *S. altissima* (Hakes and Cronin, 2011b) and in at least some other plant–enemy systems (Roy and Kirchner, 2000; Carr et al., 2006).

Selection for decreased insect virulence might be expected to be strong, since individual gallmakers depend for development on the single ramet they are attacking. However, the outcome of selection may depend on the reasons that gallmaker attack damages the host. Selection should always favour reduction in incidental damage to the plant associated with attack. However, if damage is primarily associated with losses of plant resources to gall and gallmaker tissues, then selection to reduce damage is likely to be much weaker. For each of our gallmakers, there is damage of each kind. Disruption or destruction of the apical meristem is incidental damage, while nutrition for the growing larva and construction of gall tissue have fitness value for the gallmaker. Unfortunately, we cannot currently quantify this distinction in our system. The incidental-damage perspective suggests an additional hypothesis: selection for reduced virulence, and thus the novel-impact pattern, might be stronger for stem-borers and gallmakers (which cause incidental damage) than for leaf-chewers (which sequester tissue but don’t force construction of herbivore-associated plant structures, and don’t usually damage meristems). It would be rewarding to test this hypothesis, although doing so would require work with novel and older herbivores representing multiple feeding guilds.

The kind of host shifts that we studied are not the only way that novel host–herbivore pairs are formed. They are also a product of plant and insect invasions, and there are some data from the invasive-species literature that are consistent with the evolution-of-impact hypothesis. On the whole, suppression of plant populations by herbivores is stronger in evolutionarily novel associations [exotic insects on native plants, or native insects on exotic plants (Parker et al., 2006)], although this result combines possible differences in plant resistance
(likelihood of successful attack) with those in plant tolerance and insect virulence. Plant tolerance is very low to some exotic herbivores, but high to others (Chun et al., 2010; Fornoni, 2011). These results are difficult to interpret, however, because studies of invasive-native interactions confound associational novelty with novelty in geographic distribution (which is why we studied herbivore impact in the native Solidago system). The same theoretical perspective that motivates our work would suggest that non-native plants accepted by native herbivores may experience (initially) very high impact (e.g. Parker et al., 2006; but see Müller-Schärer et al., 2004), which could help suppress their spread. Under this scenario, some plant species that are established but not invasive outside their native range actually represent future threats, as co-evolution will act over time to reduce herbivore impact. Along similar lines, new biocontrol agents for invasive plants might lose effectiveness over time as they co-evolve with their hosts. There is extensive evidence for this kind of co-evolution in systems where pests are attacked by microbial agents, including cases of increasing tolerance and decreasing virulence (Roderick and Navajas, 2003). In contrast, there is (so far) little evidence of such co-evolution between plants and insect control agents, despite the fact that genetic variance for tolerance can exist in invasive populations (e.g. Garcia-Rossi et al., 2003). Among the possible explanations for this surprising result is that enemy–victim co-evolution does indeed occur in insect–plant systems, but is generally too slow to be detected on typical monitoring time scales (Holt and Hochberg, 1997). Our data are consistent with this hypothesis, although we are quick to caution that our Solidago results should not be extrapolated without similar work in many other systems. In the Anthropocene, invasions, biocontrol introductions, and climate-driven range expansions have all become frequent, will likely continue to increase, and carry enormous ecological and economic costs and benefits. It is critically important that we understand not just present-day interactions in the novel associations we are assembling but also their likely evolutionary trajectories.

That herbivore–host associations are evolutionarily dynamic should surprise no one. Our data suggest that one facet of that dynamism is likely to be host–parasite co-evolution driving changes over time in herbivore impact. Such changes may be both predictable and important, both in natural systems and in applied contexts, including biocontrol and impacts of invasive species.

ACKNOWLEDGEMENTS

We thank Iowa Lakeside Laboratory and the Johnson County (Iowa) Conservation Board for permission to collect. For field and lab assistance we thank Alex Alder, Mischa Giasson, Eric Merion, Matt Methven, Jacob Milroy, and Bethany Nordstrom. Les Cwynar, Charles Fox, Linley Jesson, Dave MacLean, several anonymous reviewers, and members of the Heard lab provided helpful comments on the manuscript. S.B.H. thanks his Twitter followers for feedback on ‘strangers recently met’. This work was funded by Discovery Grants to S.B.H. from the Natural Sciences and Engineering Research Council (Canada).

REFERENCES


Evolution of herbivore impact


