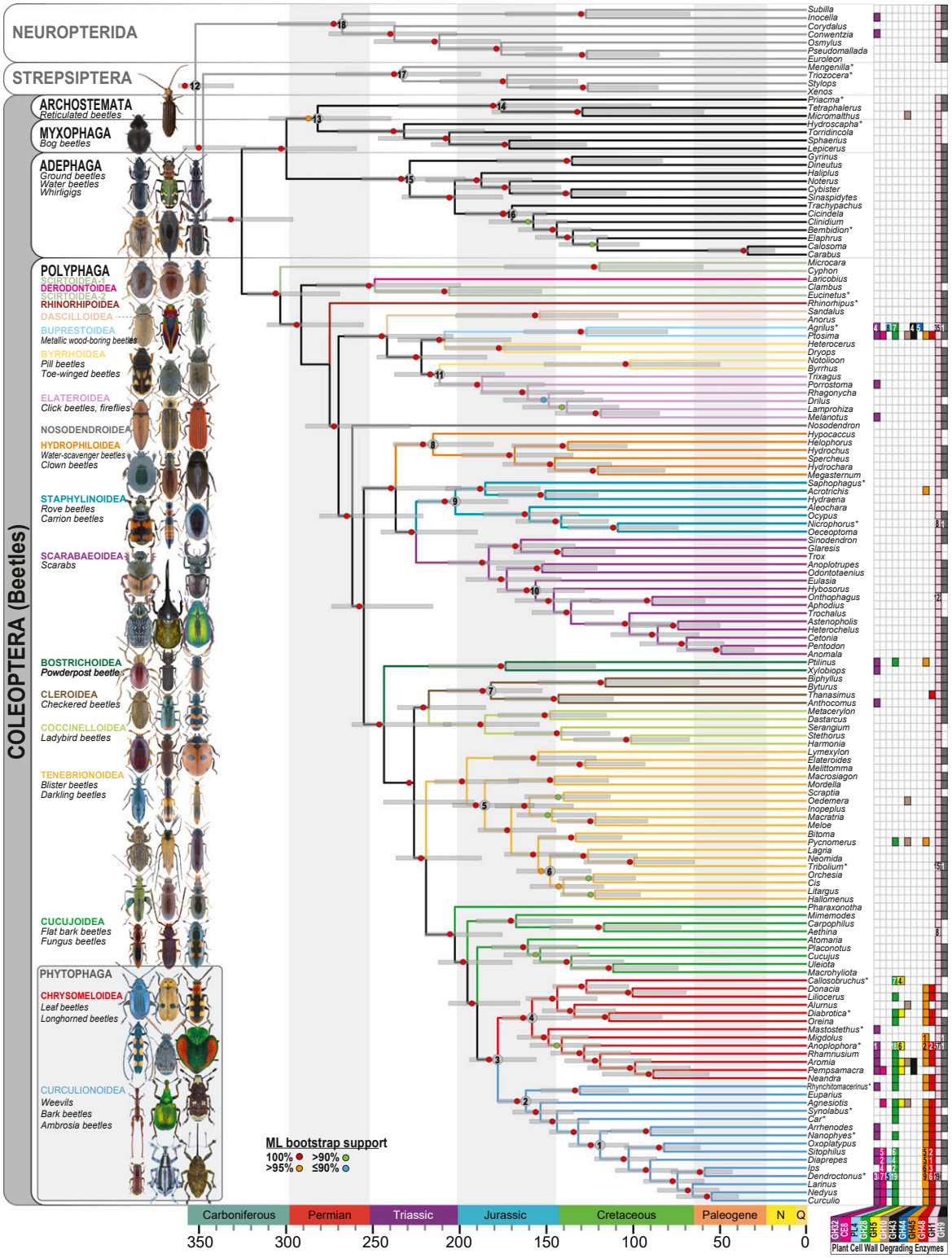

Fig. 1. A dated phylogeny of beetles showing the distribution of putative PCWDEs, inferred from 4,818 nuclear genes (*Methods* and *SI Appendix*). Branches in suborder Polyphaga are color-coded by superfamily. Numbers indicate nodes constrained by fossil priors (*SI Appendix*, Table S5). Filled squares indicate the presence of putative PCWDEs and GH32 invertases (color-coded by gene family) based on analyses of whole-genome (asterisk) or RNA-Seq data. GH1 and GH9 have known ancient origins in metazoans (14, 47) and were expected to occur in most species. Asterisks denote results from the analysis of WGS (versus RNA-Seq) data. Numbers of homologs are indicated in each box when previously published (*SI Appendix*, Table S1). Note that *Rhinorhipus* was added after the initial analyses were completed based on a new ML tree search, which recovered the same topology. Bootstrapping was not conducted due to computational constraints. However, its placement was the same in the 521-taxon tree, where it had 100% ML bootstrap support. All higher taxa shown to illustrate morphological diversity (but not all species) were sampled. *Cupes* image courtesy of Matthew Bertone (North Carolina State University, Raleigh, NC). All other photos courtesy of Udo Schmidt (photographer).

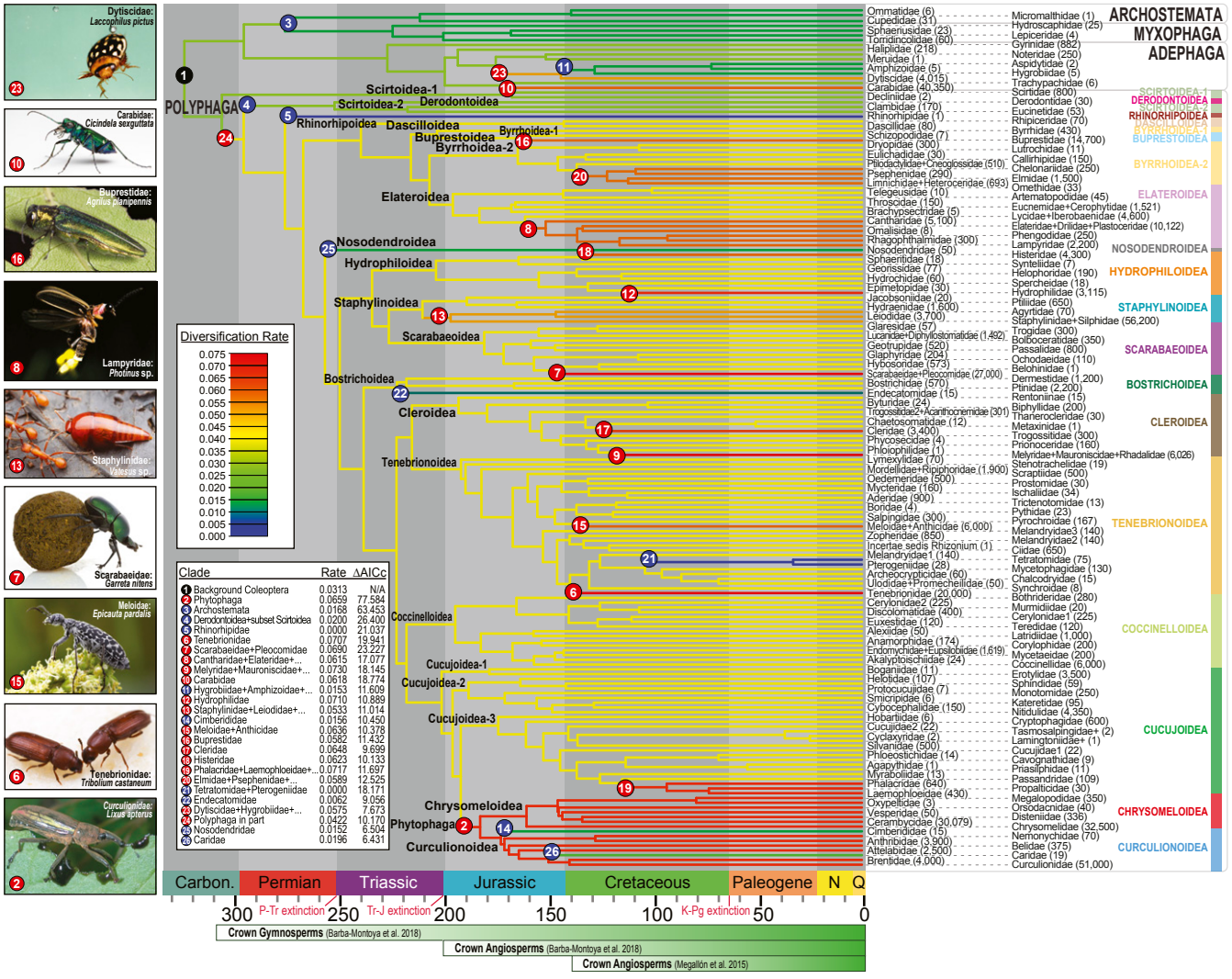

Fig. 2. Timing and rates of beetle diversification. Family-level net diversification rates and rate shifts for 188/190 beetle families (full tree; *SI Appendix*, Fig. S10 and Tables S5–S7). Branch colors indicate net diversification rates. Numbered circles indicate the locations of significant net diversification rate shifts (red/increases; blue/decreases). Approximate numbers of described extant species are indicated next to family-level taxon names (76). Note that the right column entries are interdigitated between the left ones, as connected by dashed lines. The beetle suborders and polyphagan beetle superfamilies are labeled on the right side of the tree. Outgroups are not shown. The timeline indicates mass extinctions and significant events in the history of seed plants (39, 77). Photos show select beetle groups that experienced significant diversification rate increases. *Lixus* image courtesy of D.D.M. *Garretes* image courtesy of Piotr Naskrecki (Harvard University, Cambridge, MA). *Agrilus* image courtesy of David Cappaert (photographer). All other photos courtesy of Alex Wild (photographer).

Pharaxonothinae), appeared by the late Jurassic, well before bees and butterflies (21, 42). They were likely among the first insect pollinators of gymnosperms and early angiosperms. Furthermore, our results are consistent with fossil evidence in suggesting that pollenivory was a transitional state between detritivory, mycophagy, and saprophagy (Cucujoidea) and specialized herbivory (Phytophaga) (43–45) (Fig. 1 and *SI Appendix*). The apparent prevalence of transitions in Coleoptera from generalized diets, such as detritivory and saprophagy, toward more specialized diets, such as mycophagy and herbivory, is consistent with the high rate of such transitions across insects (6).

Comparative Genomics of Beetle-Encoded PCWDEs. We studied putative PCWDEs encoded in 154 transcriptomes or genomes corresponding to the 147 taxa in Fig. 1 (*SI Appendix*, Figs. S15–S27, and Tables S1, S2, and S8). We also studied GH32 invertases, which catalyze the conversion of sucrose—the primary form of photoassimilated carbon in plant vascular tissues—to glucose and fructose. Like PCWDEs, invertases have played a potentially important role in the evolution of specialized herbivory (e.g., see ref. 46).

GH1 and GH9 have ancient origins in animals (14, 47) and were nearly ubiquitous in our study. The other gene families we studied were found almost exclusively in Buprestoidea and Phytophaga, which encoded an expansive and remarkably similar array of PCWDEs (Figs. 1 and 3). Buprestoidea and Phytophaga are the most species-rich and most specialized radiations of herbivorous beetles (1). Their feeding habits collectively include chewing, mining, and boring of virtually all kinds of plant tissues (living or dead) and plant taxa.

Outside of Buprestoidea and Phytophaga, 10 beetle species scattered widely across the phylogeny had matches to 1 PCWDE gene family (other than GH1 and GH9), and 2 beetle species each matched 3 gene families (Figs. 1 and 3). These included Bostrichidae (*Xylobiops,* GH32), Cleridae (*Thanasimus,* GH48), Elateridae (*Melanotus,* GH32), Lycidae (*Porrostoma*, GH32), Melyridae (*Anthocomus,* GH32), Micromalthidae (*Micromalthus,* GH10), Oedemeridae (*Oedemera,* GH10), Ptiliidae (*Acrotrichis,*

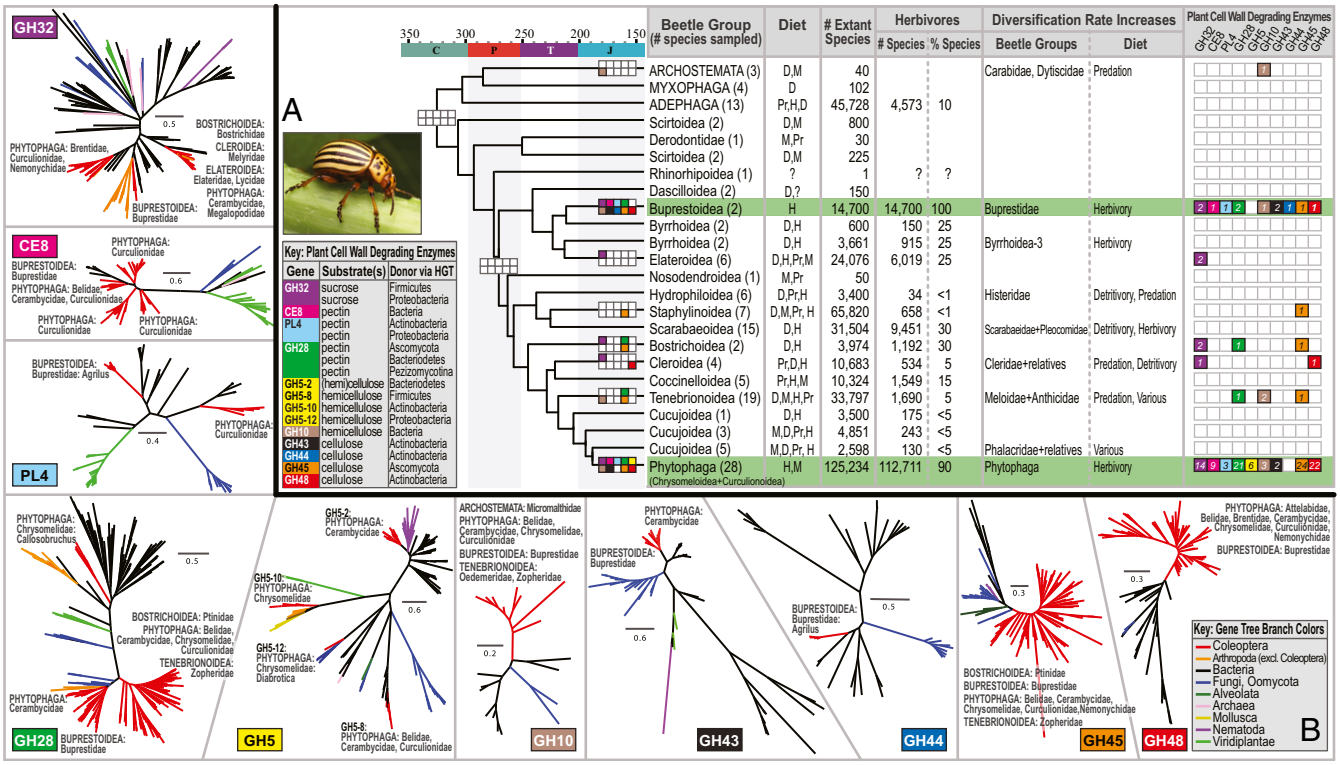

Fig. 3. Adaptive radiation of specialized herbivorous beetles after the acquisition of PCWDEs from microbes. (*A*) Summary of beetle time tree showing 2 major independent origins of novel PCWDEs from bacteria and fungi and coincident net diversification rate increases among specialized herbivorous beetles (green background; summarized from Figs. 1 and 2): 1) along the stem of Buprestidae (Buprestoidea) and 2) along the stem of Phytophaga (Chrysomeloidea + Curculionoidea). Colorized boxes indicate the presence of candidate PCWDEs. The number of beetle species sampled that contain at least one PCWDE gene family member is indicated in each box. Empty/white boxes indicate a gene family was not observed. Beetle diets: detritivory (D), mycophagy (M), predation (Pr), herbivory (H), and unknown (?). Most data were obtained from ref. 6 and are ordered by decreasing prevalence. Percent herbivores >1% is shown to the nearest 5% and was estimated based on our collective knowledge of these beetle groups. (*B*) Unrooted phylogenetic trees for PCWDE gene families illustrating the taxonomic origins of beetle-encoded genes (*SI Appendix*, Figs. S15–S26 and Datasets S1–S3). The beetle groups represented in each gene tree are labeled. *Leptinotarsa* image courtesy of the USDA Agricultural Research Service/Scott Bauer, licensed under CC BY 3.0.

GH45), Ptinidae (*Ptilinus*, GH28, GH32, and GH45), and Zopheridae (*Pycnomerus,* GH10, GH28, and GH45). In contrast, Buprestoidea and Phytophaga had species with matches from up to 7 families of PCWDEs, often with multiple apparent homologs from each family. Independent losses and reacquisitions of PCWDEs are known in Phytophaga (10) and were observed in this study.

Overall, we documented putative endogenous PCWDEs (including GH32 invertases, excluding GH1 and GH9) from 22 families of beetles. Previous to this study, they were known from only 5 beetle families (Figs. 1 and 3 and *SI Appendix*, Tables S1 and S2, and Fig. S27). Within Buprestoidea, we report GH10, GH45, GH48, and CE8, in addition to previously reported genes. Within Phytophaga, we report GH43, in addition to previously reported genes. Also within Phytophaga, we document PCWDEs from the families Attelabidae, Belidae, Brentidae, Caridae, Megalopodidae, and Nemonychidae, in addition to the families of Phytophaga from which these genes have been previously reported. Thus, we significantly expand knowledge of the phylogenetic distribution of putative PCWDEs encoded in the genomes of Coleoptera, while at the same time establishing that they are particularly diverse in the 2 lineages of specialized herbivorous beetles—Buprestoidea and Phytophaga.

Microbial Donors of Beetle PCWDEs via HGT. The inferred last common ancestors and potential donors of beetle PCWDEs were bacteria and fungi, including taxonomic groups that are today quintessential degraders of lignocellulose and other complex polysaccharides in plant and soil detritus (Fig. 3, *SI Appendix*, Figs. S15–S26, and Datasets S1–S3) (48). Some of these groups of bacteria and fungi are also found in beetle guts (49). Beetle-derived PCWDEs nonetheless formed well-supported clades distinct from microbial taxa in our phylogenies. Moreover, they were largely placed within the same clades as their homologs derived from other insect genomes and transcriptomes, including those derived from high-quality draft genomes. Within these clades, some gene families contained clusters of closely related sequences from the same beetle higher taxa or species, consistent with lineage-specific gene duplications post-HGT.

Physical incorporation of genes encoding PCWDEs into the genome of one or more beetle species has been documented for Buprestoidea-derived GH28, GH32, GH43, GH44, and PL4 and for Phytophaga-derived GH5-2, GH5-8, GH10, GH28, GH32, GH45, GH48, and CE8 (16, 50, 51) (*SI Appendix*, Table S1). Enzyme product functionality (metabolic integration) has been demonstrated for Buprestoidea-derived GH43, GH44, and PL4 and for Phytophaga-derived GH5-2, GH5-8, GH10, GH11, GH28, GH32, GH45, and CE8 (16, 50, 51). For GH5-2, GH28, GH45, and GH32, a similar gene has been independently horizontally transferred to a plant-feeding organism outside of Insecta (16).

Evidence from the available high-quality draft genomes of Buprestoidea and Phytophaga further indicates that these genes are encoded in beetle genomes and are not the result of contamination (8, 17). The emerald ash borer beetle (*Agrilus planipennis*; Buprestoidea) genome encodes GH28, GH32, GH43, GH44, and PL4 (each represented by multiple copies in the genome; Fig. 1), all of which have been PCR-amplified from adult *A. planipennis* elytra and legs—tissues not known to contain symbionts