Exaggeration and cooption of innate immunity for social defense

Mayako Kutsukake a,1, Minoru Moriyama a,b, Shuji Shigenobu c, Xian-Ying Meng d, Naruo Nikoh e, Chiyo Noda d, Satoru Kobayashi f, and Takema Fukatsu a,g,h,1

aBioproduction Research Institute, National Institute of Advanced Science and Technology, 305-8566 Tsukuba, Japan; bComputational Bio Big Data Open Innovation Laboratory, National Institute of Advanced Science and Technology, 305-8566 Tsukuba, Japan; cCore Research Facilities, National Institute for Basic Biology, 444-8585 Okazaki, Japan; dDepartment of Liberal Arts, The Open University of Japan, 261-8586 Chiba, Japan; eOkazaki Institute for Integrative Bioscience, National Institute for Basic Biology, 444-8787 Okazaki, Japan; fLife Science Center for Survival Dynamics, Tsukuba Advanced Research Alliance, University of Tsukuba, 305-8777 Tsukuba, Japan; gDepartment of Biological Sciences, Graduate School of Science, University of Tokyo, 113-0033 Tokyo, Japan; and hGraduate School of Life and Environmental Sciences, University of Tsukuba, 305-8572 Tsukuba, Japan

Edited by Nancy A. Moran, University of Texas at Austin, Austin, TX, and approved March 14, 2019 (received for review January 17, 2019)

Social insects often exhibit striking altruistic behaviors, of which the most spectacular ones may be self-destructive defensive behaviors called autothysis, “self-explosion,” or “suicidal bombing.” In the social aphid Nipponaphis monzeni, when enemies damage their plant-made nest called the gall, soldier nymphs erupt to discharge a large amount of body fluid, mix the secretion with their legs, and skillfully plaster it over the plant injury. Dozens of soldiers come out, erupt, mix, and plaster, and the gall breach is promptly sealed with the coagulated body fluid. What molecular and cellular mechanisms underlie the self-sacrificing nest repair with body fluid for the insect society? Here we demonstrate that the body cavity of soldier nymphs is full of highly differentiated large hemocytes that contain huge amounts of lipid droplets and phenoloxidase (PO), whereas their hemolymph accumulates huge amounts of tyrosine and a unique repeat-containing protein (RCP). Upon breakage of the gall, soldiers gather around the breach and massively discharge the body fluid. The large hemocytes rupture and release lipid droplets, which promptly form a lipidic clot, and, concurrently, activated PO converts tyrosine to reactive quinones, which cross-link RCP and other macromolecules to physically reinforce the clot to seal the gall breach. Here, soldiers’ humoral and cellular immune mechanisms for wound sealing are extremely up-regulated and utilized for colony defense, which provides a striking case of direct evolutionary connection between individual immunity and social immunity and highlights the importance of exaggeration and cooption of preexisting traits to create evolutionary novelties.

Author contributions: M.K. and T.F. designed research; M.K., M.M., S.S., X.-Y.M., and C.N. performed research; S.S. and S.K. contributed new reagents/analytic tools; M.K., M.M., S.S., and N.N. analyzed data; and M.K. and T.F. wrote the paper.

The authors declare no conflict of interest.

Supporting Information online includes four figures and two tables. This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1900917116/-/DCSupplemental.

Significance

Conventionally, immunology has focused on molecular and cellular mechanisms against pathogens and parasites to ensure survival of individuals. Recently, the notion of social immunity has emerged, which highlights the mechanisms in social animals to combat against pathogens, parasites, and other enemies to ensure survival of their society as a whole. Conceptually, social immunity is analogous to but distinct from individual immunity. However, we discovered that, in the social aphid Nipponaphis monzeni, molecular and cellular immune components of soldier individuals are extremely up-regulated and massively excreted via “body eruption” upon gall breakage, and the “hyperclotting” body fluid repairs the damaged gall for colony defense, which uncovers unexpected molecular, cellular, and evolutionary commonalities across individual immunity and social immunity.

www.pnas.org/cgi/doi/10.1073/pnas.1900917116

PNAS Latest Articles | 1 of 10
as galls (9, 10). The social aphid *Nipponaphis monzeni* forms large galls on the tree *Distylium racemosum* (Fig. 1A) in which hundreds to thousands of insects proliferate by sucking plant sap from the inner wall (Fig. 1B and C) (11). In *N. monzeni*, all nymphs exhibit an extended first instar stage, perform social tasks as soldiers, and then grow and reproduce (11, 12). In spring, young galls of *N. monzeni* are often attacked by lepidopteran enemies (13), whose larval tunnel and damage gall tissues and also consume inhabiting aphids. Upon invasion of the predator, morphomorphic first-instar soldiers attack the enemy by stinging with their stylets (Fig. 1D). Then, many soldiers gather around the hole on the gall wall and erupt to discharge a large amount of whitish body fluid from their cornicles on the abdominal tip (Fig. 1E and F and Movie S1). The shriveled soldiers actively mix the secretion with their legs and skillfully plaster it over the plant injury, and the secretion promptly solidifies (Fig. 1G and Movie S1). Dozens of soldiers come out, erupt, mix, and plaster, and the hole is completely plugged by the coagulated body fluid (Fig. 1H and I and Movie S1) (12, 14). Here, the soldier nymphs of *N. monzeni* perform not only a soldier-type aggressive task of attacking enemies but also a worker-type housekeeping task of repairing their plant-made nest, which entails a series of unique self-destructive behaviors. The molecular and cellular mechanisms that underlie the self-sacrificing nest repair with the use of body fluid for the insect society is of great interest, which prompted us to investigate the biochemistry, physiology, cell and molecular biology, and developmental and evolutionary aspects of the gall-repairing body fluid produced by soldier nymphs of *N. monzeni*.

**Results and Discussion**

**Darkening upon Clotting, Large Globular Cells, and a Few Major Proteins in Body Fluid of Soldier Nymphs.** The soldiers’ secreted body fluid turned black as its solidification proceeded (Fig. 1 I–L). Mechanical stimulation of soldier nymphs in a saline solution elicited the body fluid secretion (Fig. 1M and Movie S2), which demonstrated that the body fluid contains numerous peculiar large cells, called the large globular cells (LGCs) (14), that fill up the soldier’s body cavity (Fig. 1 N and O). Histological examination of the hole-filling plugs identified nuclei and granules derived from LGCs (Fig. 1P). SDS/PAGE of the soldiers’ body fluid detected only a few major protein bands (Fig. 2A, Left), suggesting that these proteins may play some roles in the self-sacrificing gall repair by soldier nymphs.

**Phenoloxidase as a Predominant Protein in Body Fluid of Soldier Nymphs.** The darkening of the coagulated plugs (Fig. 1 I–L) suggested progressive melanin formation in the secreted body fluid, where the melanin synthesis pathway driven by PO may be activated (3, 15). Immunoblotting identified one of the predominant proteins in the soldier’s body fluid as PO (Fig. 2A, Center, band 5). An enzymatic activity assay detected drastically high PO activities in the soldier nymphs of *N. monzeni* (Fig. 2 B and C). Molecular cloning of the PO cDNA assisted by protease digestion and peptide sequencing of the no. 5 band protein yielded a full-length cDNA sequence 2,355 bp in size containing an ORF encoding a protein of 685 aa residues (SI Appendix, Fig. S1A). Putative copper-binding histidine residues, which are generally conserved in arthropod POs (16), were present, whereas a conserved proteolytic cleavage site for pro-PO activation, namely arginine-phenylalanine at approximately 50 residues from the N terminus (15), was not identified (SI Appendix, Fig. S1A). Molecular mass fingerprinting of the no. 5 band protein identified an N-terminal preprotease fragment of typical insect pro-PO, indicating that the PO protein is mainly in a proenzyme form in vivo. RNA sequencing (RNA-seq) analysis of *N. monzeni* (SI Appendix, Tables S1 and S2) detected that, in addition to the soldier-specific PO that exhibited extremely high expression levels preferentially in the soldier nymphs, four additional PO genes were expressed at low levels in *N. monzeni* (SI Appendix, Fig. S2 B–E). On the phylogenetic tree, the five PO genes of
N. monzeni clustered with two PO genes of the pea aphid *Acyrthosiphon pisum* (SI Appendix, Fig. S1B), suggesting the possibility of PO gene duplications and diversification in the evolutionary course of aphids. Of these, the soldier-specific PO gene exhibited an outstandingly high expression level, LGC-specific up-regulation, and the most basal and elongated branch in the phylogeny (SI Appendix, Figs. S1B and S2 A–E and Table S2). The soldier-specific PO gene showed a significantly accelerated molecular evolutionary rate in comparison with the other PO genes (SI Appendix, Fig. S3), which may be relevant to the evolution of molecular evolutionary rate in comparison with the other PO genes (SI Appendix, Fig. S3), which may be relevant to the evolution of novel biological function for the self-sacrificing gall repair.

**Repeat-Containing Proteins as Abundant and Highly Polymorphic Proteins in Body Fluid of Soldier Nymphs.** Next, we attempted to identify the other abundant proteins in the soldier’s body fluid (Fig. 2D, bands 4 and 6). Molecular cloning assisted by protease digestion and peptide sequencing of the no. 6 band protein yielded mixed cDNA sequences whose 5′- and 3′-end regions were conserved but whose middle region was polymorphic. Immunoblotting using an antibody against the N-terminal peptide of the no. 6 band protein revealed but whose middle region was polymorphic. Immunoblotting using an antibody against the N-terminal peptide of the no. 6 band protein revealed conserved region (Fig. 2A). Structurally, the polymorphic regions were explained by partial insertions/deletions of amino acid residue is replaced as indicated. Dashed lines highlight deduced insertions/deletions of the motifs during the evolution of RCP.

Fig. 2D shows SDS/PAGE and immunoblot profiles of the body-fluid proteins of the soldier nymphs derived from eight galls (designated as galls a–h) collected at the same locality. Whereas the most abundant PO protein (Fig. 2D, band 5) was constant in size, the next most abundant proteins (Fig. 2D, bands 4 and 6) exhibited peculiar patterns: one or two bands were detected from each gall colony, and the size of the protein bands differed among the gall colonies (Fig. 2D), uncovering that the no. 4 and no. 6 band proteins show striking size polymorphism even in the same population of *N. monzeni*. To identify the basis of the size polymorphism, insects from five gall colonies (designated as galls i–m) were subjected to full-length cDNA sequencing of the polymorphic proteins. RT-PCR amplification of RCP cDNA sequences from five sympatric gall colonies i–m. Two sequences were amplified from colonies i, j, and m, whereas only one sequence was obtained from colonies k and l. (Fig. S1B) Schematic diagram showing amino acid sequences of RCP alleles deduced from complete cDNA sequences obtained from colonies i–m. Colored boxes depict seven types of repeat motif, each 8 aa in size: gray, GSGGQGYS; green, EHEQG[S/H]T[N/I]; blue, EHGQG[S/H]T[N/I]; orange, EPEEQG(Y/T); purple, EHKQG[S/H]; pink, GPGGQG[S/H][T/N]; and yellow, GHEGQ[S/H][Y/T]/N]. Daggers indicate repeat motifs in which the first amino acid residue is replaced as indicated. Dashed lines highlight deduced insertions/deletions of the motifs during the evolution of RCP. “S” indicates the N-terminal signal sequence. (G) Attributes and sequence accession numbers for the RCP alleles identified from gall colonies i–m.

![Fig. 2](image-url)

**Fig. 2.** PO and RCP, major proteins in the *N. monzeni* soldier’s body fluid. (A) SDS/PAGE of soldier’s body fluid: (Left) general protein staining with Coomassie Brilliant Blue (CBB), (Middle) immunoblotting against PO, and (Right) immunoblotting against RCP. Numbers 1–6 show major protein bands. Asterisks indicate nonspecific signals. (B) Measurement of PO activity in soldier’s body fluid. (C) Measurement of PO activity in whole body extract of *N. monzeni* and allied nippo-naphnid aphids *N. distyliicola* and *N. yanonis* that are incapable of gall repair with the use of their body fluid. In B and C, letters a and b indicate statistically significant differences (Tukey’s honestly significant difference test, P < 0.05). (D) SDS/PAGE and RCP-targeted immunoblotting of soldier’s body fluid proteins from eight sympatric gall colonies a–h: general staining with CBB (Left) and immunoblotting against RCP (Right). Asterisks indicate nonspecific signals. (E) RT-PCR amplification of RCP cDNA sequences from five sympatric gall colonies i–m. Two sequences were amplified from colonies i, j, and m, whereas only one sequence was obtained from colonies k and l. (F) Schematic diagram showing amino acid sequences of RCP alleles deduced from complete cDNA sequences obtained from colonies i–m. Colored boxes depict seven types of repeat motif, each 8 aa in size: gray, GSGGQGYS; green, EHEQG[S/H][T/N/I]; blue, EHGQG[S/H][T/N/I]; orange, EPEEQG(Y/T); purple, EHKQG[S/H]; pink, GPGGQG[S/H][T/N]; and yellow, GHEGQ[S/H][Y/T]/N]. Daggers indicate repeat motifs in which the first amino acid residue is replaced as indicated. Dashed lines highlight deduced insertions/deletions of the motifs during the evolution of RCP. “S” indicates the N-terminal signal sequence. (G) Attributes and sequence accession numbers for the RCP alleles identified from gall colonies i–m.
as a repeat-containing protein (RCP). The RCP gene showed an outstandingly high expression level and LGC-specific up-regulation in soldier nymphs of *N. monzeni* (SI Appendix, Fig. S2F and Table S2).

Oddly, SDS/PAGE estimated the molecular mass of RCP as 60–80 kDa (Fig. 2D), which was discrepant with the actual molecular mass determined by molecular cloning (36–45 kDa for galls I–m; Fig. 2C). MALDI-TOF MS analysis of the soldier’s body fluid proteins supported the latter molecular mass of RCP (27–43 kDa for galls a–h; SI Appendix, Fig. S3C), corroborating the unusual electrophoretic mobility of RCP. We speculate that the atypical structure of the large repeat region might be responsible for this peculiar feature of RCP.

**Localization of PO and RCP in LGCs and Body Cavity of Soldier Nymphs.** In situ hybridization and immunohistochemistry visualized signals of PO and RCP in LGCs that occupied the body cavity of soldier nymphs and soldier-to-be embryos in the maternal body (Fig. 3 A–C and SI Appendix, Fig. S4 A–C). In early soldier embryos, LGCs had a large nucleus and ribosome-rich cytoplasm (Fig. 3D). As the embryonic and nymphal development proceeded, LGCs accumulated lipid droplets in the cytoplasm (Fig. 3 E and F). In mature soldier nymphs, LGCs were full of lipid droplets with an atrophied nucleus and some cytoplasm remaining between the lipid droplets (Fig. 3G). By immuno-EM, PO was preferentially localized to the cytoplasm of LGCs (Fig. 3H and SI Appendix, Fig. S4D), whereas RCP was detected mainly in the hemocoele outside LGCs (Fig. 3I and SI Appendix, Fig. S4E). The extracellular localization of RCP is consistent with the presence of the secretion signal sequence at the N terminus of the RCP gene (SI Appendix, Fig. S3A).

**Up-Regulation of PO-Related Melanization Pathway Genes in LGCs of Soldier Nymphs.** In insects, PO-mediated melanin formation is involved in a variety of biological functions such as innate immunity, hemolymph clotting, and wound sealing (1–3, 15, 17). Fig. 4A shows the melanin synthesis pathway in insects: PO or tyrosine hydroxylase converts tyrosine to dihydroxyphenylalanine (DOPA); DOPA decarboxylase (DCC) catalyzes the reaction from DOPA to dopamine; PO and multicopper oxidase 2 (MCO2) catalyze the reaction from DOPA or dopamine to dopaquinone or dopamine-quinone, respectively; dopaquinone and dopamine-quinone spontaneously become dopachrome and dopaminochrome, respectively; dopachrome and dopamine-quinone are converted to 5,6-dihydroxyindole (DHI)-2-carboxylic acid (DHICA) and/or DHI by dopachrome conversion enzyme (DCE; encoded by yellow family genes); and DHICA or DHI is converted to melanochrome by PO and MCO2 and polymerized to form melanin. Our RNA-seq data revealed that, in addition to PO, DDC and DCE were conspicuously up-regulated in LGCs (SI Appendix, Fig. S2 J and J and Table S2), indicating that enzymes constituting the PO cascade are generally and drastically up-regulated in LGCs of soldier nymphs (Fig. 4A). It is also notable that a serine protease gene and a serpin gene, which are known to be involved in the proteolytic cascade of PO activation (18), were highly expressed in LGCs (SI Appendix, Fig. S2 K and L and Table S2), although their actual involvement in the PO activation cascade should be verified in future studies.

**Extraordinarily Abundant Tyrosine in Body Fluid of Soldier Nymphs.** As depicted in Fig. 4A, tyrosine is the main and starting substrate for the PO-mediated melanization pathway. Notably, tyrosine was extraordinarily predominant among free amino acids in the body fluid of soldier nymphs of *N. monzeni*, measuring 31.7 ± 5.5 mM and accounting for 75.9% of total free amino acids (Fig. 4B). We found that (i) tyrosine was the most abundant free amino acid in the body fluid of *N. monzeni* irrespective of developmental stage, (ii) the level of tyrosine was more than twice as high in soldier nymphs than in *N. monzeni* adults, and (iii) such high tyrosine titers are not observed in nonsocial aphids like *A. pisum* (19) (Fig. 4 B–D). In soldier nymphs of *N. monzeni*, tyrosine was detected not only from hemolymph (0.62 ± 0.21 nm per insect, n = 9) but also from LGCs (0.14 ± 0.08 nm per insect, n = 9). DOPA was negligible in soldiers’ freshly secreted body fluid (0.02 ± 0.03 mM, n = 7) but became detectable 15 min after secretion (2.3 ± 1.8 mM, n = 7), indicating conversion of tyrosine to DOPA in the solidifying body fluid.

**In Vitro Clotting Assay Using PO, RCP, and Tyrosine.** In the PO-mediated melanin synthesis pathway, tyrosine is converted to highly reactive quinone molecules (Fig. 4I), which kill parasites and pathogens through free-radical reactions (20) and cross-link proteins for hemolymph clotting and cuticular tanning (21, 22). The extreme up-regulation of PO, RCP, and tyrosine in soldier nymphs of *N. monzeni* prompted us to hypothesize that these molecular components may play pivotal roles in the body fluid coagulation and melanization upon self-sacrificing gall repair (Fig. 1 and Movie S1). To test this hypothesis, we attempted to reconstruct the chemical reactions by using defined molecular components. Recombinant PO protein was produced by using SF9 insect cells and baculovirus expression system and successfully purified and activated by an addition of 2-propanol (23 (SI Appendix, Fig. S4 F–H)). Recombinant RCP protein was produced and purified by using *Escherichia coli* and pET system (SI Appendix, Fig. S4I). In the soldier’s body fluid, estimated concentrations of PO, RCP, and tyrosine were 15.3 ± 5.6 μg/μL (n = 3), 5.3 ± 0.6 μg/μL (n = 3), and 31.7 ± 5.5 mM (n = 9), respectively. As these protein concentrations are too high to be handled practically, we performed the reconstruction experiments at a 1/10 concentration scale, whereby the basic reaction mixture (6 μL) contained approximately 1.5 μg/μL recombinant PO, 0.5 μg/μL recombinant RCP, and 3.3 mM tyrosine. In the presence of 20% 2-propanol, PO was immediately activated, and the solution turned brownish within 1 min and then blackish within 10 min. Concurrently, RCP was cross-linked and polymerized, causing disappearance of the RCP band and accumulation of a high–molecular-mass protein band at the upper end of the SDS/PAGE separating gels (Fig. 5A). After 1 h, a lot of black clots formed in the solution (Fig. 5D), and, after 16 h, all RCP protein was consumed and black precipitates formed (Fig. 5A). In the absence of tyrosine, by contrast, neither RCP consumption nor clot formation was observed (Fig. 5B and E), corroborating the idea that PO-mediated conversion of tyrosine to reactive quinones should be responsible for protein cross-linking and clotting (Fig. 4I). When RCP was replaced by BSA, cross-linking and clotting certainly occurred, but the levels of protein consumption and cross-linking of BSA were much less efficient than those of RCP (Fig. 5 C and F), indicating that RCP is a more susceptible substrate for PO/tyrosine-mediated cross-linking in comparison with BSA. Notably, the black clots formed not only with RCP (Fig. 5D), but also with BSA (Fig. 5F), and even without both RCP and BSA (Fig. 5G). Considering that PO is the most abundant (1.5 μg/μL) protein component in the reaction mixture and PO bands also quickly disappeared as the reaction proceeded (Fig. 5 A–C), PO itself must be cross-linked and contribute to the clot formation. Note that activated PO is known as a very “sticky” protein that is easily lost in experimental handling (15). When quantities of PO and tyrosine were reduced, the clot formation was suppressed conspicuously (Fig. 5 H–K), whereas the quantities of black clots were apparently not affected by increase or decrease of RCP (Fig. 5 L and M). These results strongly suggest that the up-regulated production of PO, RCP, and tyrosine underlies the melanization and hardening of the secreted body fluid upon self-sacrificing gall repair by soldier nymphs.

**Abundant Lipids in Clotted Body Fluid and LGCs.** Here it should be noted that soldiers’ excreted body fluid is full of LGCs (Fig. 1 M–O), and LGCs are full of lipid droplets (Fig. 3G). Chloroform
extraction of hole-filling plugs collected from repaired galls revealed that lipids accounted for 82.5 ± 3.1% (n = 12) of their dry weight, indicating that lipids quantitatively constitute the major component of soldiers’ coagulated body fluid. Cornicles are a pair of tube- or pore-like defensive organs specifically found on the abdominal tip of aphids (24, 25). When stimulated or threatened, aphids secrete sticky liquid from the cornicles, which contain waxy substances and alarm pheromones, to deter predatory attacks and to elicit escaping behavior of colony mates (26–28). Previous studies reported that triglycerides are among the major components of the aphid’s cornicle secretion (29–31). Our lipid analysis revealed that (i) the lipids in the secretion were mostly triglycerides whose estimated total concentration was as high as 528 ± 85 mM (n = 7), (ii) the major components were triglycerides consisting of sorboyl and dimyristic acids (C6:2, C14:0, C14:0; 35.0%) and sorboyl, myristic, and palmitoyl acids (C6:2, C14:0, C16:0; 28.7%), (iii) the lipid composition did not change before and after body fluid coagulation, and (iv) the fatty acids separated by alkaline hydrolysis of soldiers’ excreted body fluid were almost the same as those constituting the triglycerides (SI Appendix, Fig. S5).

Two-Step Model for the Mechanism of Gall Repair: Immediate Formation of Lipidic Soft Clot Followed by PO-Mediated Clot Hardening. On the basis of these results, we propose a model for the molecular and cellular mechanisms underlying the self-sacrificing gall repair by soldier nymphs of N. monzeni. In the soldier’s body cavity, inactive PO and lipid droplets are stored in LGCs, whereas tyrosine and RCP accumulate in the hemolymph (Fig. 6A, Left). Upon discharge of the body fluid, LGCs rupture and release lipid droplets, which promptly form a lipidic soft clot, and PO is concurrently activated, presumably by hemolymphal enzymes, and initiates the PO cascade, by which tyrosine is converted to reactive quinones (Fig. 6A, Middle). Subsequently, proteins and other macromolecules are
cross-linked by the reactive quinones, and the lipidic clot is physically reinforced and melanized (Fig. 6A, Right).

On Developmental and Evolutionary Origin of LGCs. LGCs are peculiar and enigmatic cells found in soldier nymphs of *N. monzeni*, being very large in size, occupying the soldier’s body cavity, highly expressing PO and RCP, and accumulating a large amount of triglycerides. In insects and crustaceans, PO is generally produced by hemocytes and released and activated by cell lysis (2, 3) (Fig. 6B). Here we hypothesize that LGCs may be derived from the aphid’s hemocytes (32). On the contrary, considering the large size, localization in the body cavity, and accumulation of lipids, LGCs look reminiscent of fat body cells, which favors the hypothesis that LGCs may be derived from the aphid’s fat body cells (33). RNA-seq analysis of *N. monzeni* revealed that, besides PO and RCPs, lipid- and sugar-related genes tend to be highly expressing in LGCs (Fig. S2). Gene Ontology (GO) analysis confirmed that GO terms related to lipid and sugar metabolisms are enriched in the LGC-dominant genes (Appendix S1). These gene-expression patterns are concordant with the histochemical detection of accumulated lipids and polysaccharides in LGCs (Fig. 3). In an attempt to evaluate which of these hypotheses better applies, we obtained RNA-seq data from fat body cells, hemocytes, gut, and whole body of the nonsocial aphid *A. pisum* (Appendix S1) and analyzed them with RNA-seq data of LGCs from soldier nymphs of *N. monzeni*. In *A. pisum*, two PO genes were preferentially expressed in hemocytes (Appendix S2), whereas sugar- and lipid-related genes and RCP-like genes (to be detailed later) were mainly expressed in fat body cells (Appendix S4). Principal-component analysis and cluster analysis using 9,825 orthologous genes across *N. monzeni* and *A. pisum* revealed that the gene-expression pattern of LGCs is similar to neither hemocytes nor fat body cells of *A. pisum* (Appendix S2, Fig. S5). On the basis of these results, we conclude that LGCs constitute a unique cell type specific to soldier nymphs of *N. monzeni*. Considering the high PO expression in aphid hemocytes (Appendix S2, S5), it seems plausible that LGCs represent highly differentiated hemocytes with fat body-like features.

Exaggeration and Cooption of Wound Sealing and Defense Mechanisms at Individual Level for Social Defense at Colony Level. When an insect is injured, the epidermal breach is promptly sealed by body fluid clotting, in which PO activation, melanization, protein cross-linking, and hemocyte aggregation are involved (2, 3). Here we hypothesize that these biochemical and cellular mechanisms of body fluid clotting at the individual level are extremely up-regulated in soldier nymphs of *N. monzeni*, and, by recruiting the aphid-specific defensive trait of cornicle secretion behavior, the superclotting body fluid is collectively excreted by soldier nymphs in a highly coordinated manner, thereby promptly accomplishing the gall repair at the colony level in a social context (Appendix S2, C and D).

On Polymorphism and Evolution of RCP and Other Body Fluid Proteins. The remarkable size polymorphism of RCP among the sympatric gall colonies of *N. monzeni* is striking. Plausibly, the repeat number variation in RCP (Fig. 2D–G) has been generated through replication slippage and/or unequal crossing-over, as observed in spider and lepidopteran silk proteins (34, 35). In general, gall-forming...
aphids exhibit cyclical parthenogenesis, in which sexual females and males appear and mate in autumn to produce overwintering eggs. The fertilized eggs hatch in spring as gall-forming parthenogenetic females called fundatrices or stem mothers (36). It should be noted that, after sexual reproduction, the fundatrices are expected not to produce males, even from the same mother. As N. monzeni forms a completely closed gall (11, 37), clonal mixing across the gall colonies resulting from intergall migration, as observed in social aphids forming open galls (38, 39), is unlikely to occur in N. monzeni. Hence, different gall colonies of N. monzeni likely represent different genotypes. The conspicuous repeat number variation may be attributable to functional constraints imposed on the histidine-rich amino acid composition rather than on the repeat number (Fig. 2F). Note that, in insect cuticular sclerotization, histidine and other amino acid residues of cuticular proteins react with quinone molecules, by which cross-linking and hardening of cuticle proceed (40).

BLAST searches against the DNA databases retrieved no genes with significant sequence similarity to RCP. Meanwhile, cDNA cloning of the RCP gene also identified a smaller gene (654–702 bp in size and encoding an ORF of 217–233 aa residues), designated as RCP small (RCP-S), whose N- and C-terminal regions exhibited sequence similarities to those of RCP, with the middle region consisting of only two to four repeat units (SI Appendix, Fig. S3A and E). RCP-S gene also exhibited soldier-specific high expression and LGC-specific up-regulation (SI Appendix, Figs. S2G and S3D and Table S2). It seems likely that RCP-S also participates in the body fluid coagulation upon gall repair, although its exact function is to be established in future studies. In the transcriptome data of the pea aphid A. pisum, we identified five genes (ACYPI005732, ACYPI069330, ACYPI086030, ACYPI081606, ACYPI088785) presumably related to RCP in that they possess a signal sequence at the N terminus (75–80% identical to that of RCP) followed by a highly repeated sequence region, although the region showed no remarkable sequence similarity to RCP. Of these, two genes (ACYPI086030 and ACYPI005732) exhibited up-regulation in fat body cells (SI Appendix, Figs. S2 U and V), whereas three genes (ACYPI081606, ACYPI088785, and ACYPI069330) were expressed in hemocytes at low levels (SI Appendix, Fig. S2 W–Y). On the basis of these results, we speculate that RCP has experienced dynamic evolutionary trajectories in aphids, which must have entailed gene duplications, repeat acquisitions, and amplifications. The original biological function of RCP-like genes in aphids is still an enigma and is to be addressed in future studies.

Molecular cloning assisted by protease digestion and peptide sequencing of the no. 3 band protein identified a large 7,170-bp gene, which encodes a fatty acid synthase of 2,389 aa residues with an expected molecular mass of 259 kDa as observed on SDS/ PAGE and immunoblotting (SI Appendix, Fig. S3F), although the gene exhibited no conspicuous expression levels in LGCs and soldier nymphs (SI Appendix, Fig. S2H). As protease digestion and peptide sequencing of no. 1 and no. 2 band proteins failed, their molecular identity has not yet been determined. On the basis that the no. 1 and no. 2 band proteins reacted to anti-RCP antibody (Fig. 2A and D), although speculative, these bands may represent some protein complexes containing RCP as a component. Functional aspects of these large-sized body fluid proteins of soldier nymphs deserve future studies.

On the Origin of Tyrosine: Attributable to Bacterial Symbiont? The origin of the huge amount of tyrosine in the soldier’s body fluid (Fig. 4B), which plays a pivotal role in the gall repair by soldier nymphs (Fig. 5), is of great interest. Initially, we suspected that the bacteriocyte symbiont Buchnera may play a primary role in the overproduction of tyrosine because the symbiont is well known to provide the host with essential amino acids (41, 42), and some insect symbionts are known to be specialized for the provision of tyrosine (43). However, our sequencing of the 0.59 Mb Buchnera Nmo genome (Fig. 7A) uncovered no genomic features specialized for tyrosine production. The tyrosine synthesis genes were neither amplified nor encoded on plasmids, but looked like the same as those found in other Buchnera genomes, retaining most of the shikimate pathway genes for converting erythrose-4-phosphate and phosphoenolpyruvate into phenylalanine via chorismate and phenylpyruvate, but lacking the final step gene for conversion of phenylalanine to tyrosine (44, 45) (Fig. 7B). RNA-seq analysis of bacteriocytes of N. monzeni detected no conspicuous up-regulation of these symbiont genes (SI Appendix, Table S5), plausibly reflecting the paucity of transcriptional regulators in the streamlined Buchnera genome (46, 47). On the contrary, phenylalanine-4-monoxygenase, the host enzyme catalyzing the conversion of phenylalanine to tyrosine, was drastically up-regulated in bacteriocytes but not in LGCs (SI Appendix, Figs. S2M and S7A), highlighting the host’s regulation of tyrosine overproduction in the symbiotic cells, as reported in weevils Nardenella endosymbiosis (43). It is also notable that transaldolase and phosphoenolpyruvate carboxykinase (PEPCK), the host enzymes to generate erythrose-4-phosphate and phosphoenolpyruvate, respectively, were up-regulated in bacteriocytes and LGCs (SI Appendix, Figs. S2 N and O and S7 E and F).
suggesting the speculative possibility that the starting substrates for the tyrosine synthesis pathway are overproduced by the host cells. Taken together, the massive tyrosine accumulation in soldier nymphs of *N. monzeni* may be mainly attributable to up-regulation of the host’s metabolic genes, whereas the tyrosine synthesis pathway is constituted by the host genes and the symbiont genes (Fig. 7B). In addition, how tyrosine and its precursors are transported across the host and symbiont cellular compartments, in which some specific transporter molecules should be involved (48, 49), is to be established in future studies.

**Conclusion and Perspective.** In conclusion, the body fluid of soldier nymphs of *N. monzeni* consists of two distinct components: the cellular component, LGCs, as highly differentiated, PO-producing, and lipid-accumulating hemocytes; and the humoral component, hemolymph, that accumulates huge amounts of tyrosine and RCP. The hemocytes and the PO/tyrosine-mediated cross-linking cascade are normal components underpinning innate immunity and wound sealing in insects. In soldier nymphs of *N. monzeni*, these mechanisms are extremely up-regulated, and, in combination with the aphid-specific defensive behavior of cornicle secretion and the highly coordinated social behavior of soldier nymphs, the superclotting body fluid is massively excreted outside the insect body and used to seal the gall breach. In this way, molecular and cellular mechanisms of innate immunity at the individual level are directly coopted for colony defense at the group level in an ecological context, thereby realizing an amazingly unique style of social immunity.

To gain further insight into the evolutionary process leading to the molecular and cellular specializations in soldier nymphs of *N. monzeni*, it is necessary to better understand the innate immune mechanisms, in particular those concerning body fluid clotting and PO-expressing hemocytes, in allied nonsocial aphids. Currently, however, even in the model aphid *A. pisum*, very little is known about humoral and cellular aspects of innate immunity (32) except for the genomic and transcriptomic revelation of the aphid’s exceptionally reduced immune system, in which many conserved immune components, including IMD pathway genes, peptidoglycan recognition protein genes, and antimicrobial peptide genes, are lacking (50). In *Drosophila melanogaster*, cross-linking enzymes, PO and transglutaminase, and structural proteins, hemoclitin and fondue, were reported to be involved in hemolymph.
clotting (51–54). In the transcriptomic data of *N. monzeni*, in contrast to the extremely up-regulated PO ([SI Appendix, Fig. S2](#)), transglutaminases and hemolectin were expressed at marginal levels ([SI Appendix, Fig. S2](#)) and fondue was not detected, which reflects the diversity of clotting mechanisms among the different insect groups.

Social insects often exhibit remarkable altruistic behaviors (55). Among them, the most spectacular ones may be self-destructive defensive behaviors called autotomy, “self-explosion,” or “suicidal bombing” in some ants, termites, and aphids (56). Initially, these striking phenomena were regarded as the subject of curiosity in the field of natural history, and then reinterpreted in the context

---

**Fig. 7.** Tyrosine synthesis genes of *N. monzeni* and bacteriocyte symbiont *Buchnera*. (A) The *Buchnera* Nmo genome consisting of a chromosome and two plasmids. In total, 439 protein-coding genes (blue on + strand, pink on – strand), 31 tRNA genes (green), and 2 rRNA genes (orange) are identified. (B) A hypothetical model of tyrosine synthesis metabolic pathways operating in soldier nymphs of *N. monzeni*. Genes and metabolic pathways of *Buchnera* are shown in blue. Up-regulated host genes of interest are highlighted in red. In the *Buchnera* cell, the shikimate pathway genes are constitutively expressed, thereby converting the starting substrates, erythrose-4-phosphate (E4P) and phosphoenolpyruvate (PEP), into phenylpyruvate (PHP) and phenylalanine. In the bacteriocyte, the phenylalanine-4-monooxygenase gene is highly and specifically up-regulated ([SI Appendix, Figs. S2M and S7A](#)), by which phenylalanine is converted to tyrosine. Aspartate aminotransferase genes are not up-regulated in the bacteriocyte ([SI Appendix, Fig. S7 B–D](#)), suggesting that the tyrosine synthesis pathway via PHP may be less important in the host cytoplasm. The transaldolase (Tal) gene, which constitutes the pentose phosphate pathway and produces E4P, is constitutively expressed, whereas the PEPCK gene, which constitutes the gluconeogenesis pathway and produces PEP, is up-regulated in the LGC ([SI Appendix, Figs. S2 N and O and S7 E and F](#)), which suggests the possibility of host supply of starting substrates for the tyrosine synthesis pathway in the symbiotic system. Note that the enolase (Eno) gene, which constitutes the glycolysis pathway and produces PEP, is not up-regulated in the bacteriocyte and the LGC ([SI Appendix, Fig. S7G](#)).
of evolutionary ecology and sociobiology. In this study, we un-
covered the immunity-derived molecular and cellular bases of the self-sacrificing colony defense in a social aphid, thereby high-
lighting the importance of exaggeration and cooption of preex-
isting traits in the creation of evolutionary novelties and pointing to a way in which molecular and cellular biology can contribute to our understanding of ecology and evolution from a viewpoint of immunity and defense.

Materials and Methods

Galls of N. monzeni were collected from D. racemosum trees at Shin-Kiba, Tokyo, Japan. Insects collected from the galls were subjected to experiments immediately or preserved in an ultracold freezer until use. Galls of related aphid species, Nipponaphis distyliicola were also collected from D. racemosum trees at the same locality. The pea aphid A. pisum strain Apl, which was used for RNA-seq analysis, was collected and established in Sapporo, Hokkaido, Japan. Further details on the study ma-
terials and methods are provided in SI Appendix, SI Materials and Methods.

ACKNOWLEDGMENTS. We thank Keigo Uematsu and Harunobu Shibao for aphid collection; Tsunaki Asano, Hiroaki Sato, and Tomohiro Tamura for protein analysis; Ryouichi Koga and Kazuhiro E. Fujimori for histology; Kazutoshi Yoshitake and Ryo Futahashi for bioinformatics; and Ryo Futahashi comments on the manuscript. This work was supported by Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research JP18K06373, JP23770278, JP18770222 (to M.K.), and JP17H06388 (to T.F.); and a Hayashi Memorial Foundation for Female Scientists Research Grant (to M.K.).

1. Lemaire B, Hoffmann J (2007) The host defense of Drosophila melanogaster. Annu Rev Immunol 25:697–743.
2. Cerenius L, Söderhäll K (2011) Coagulation in invertebrates. J Innate Immun 3:3–8.
3. Eleftherianis I, Reven C (2011) Role and importance of phenoloxidase in insect hemostasis. J Innsicr Immu 3:28–33.
4. Jravaranipaisal P, Lee BL, Söderhäll K (2006) Cell-mediated immunity in arthropods: Hematopoiesis, coagulation, melanization and opsonization. Immunobiology 211: 213–236.
5. Strand MR (2008) The insect cellular immune response. Insect Sci 15:1–14.
6. De Gregorio E, Spellman PT, Tzou P, Rubin GM, Lemaire B (2002) The Toll and irmd pathways are major regulators of the immune response in Drosophila. EMBO J 21: 2568–2579.
7. Cremer S, Armitage SAO, Schmid-Hempel P (2007) Social immunity. Curr Biol 17: R637–R702.
8. Stroeymeyt N, Casillas-Pérez B, Cremer S (2014) Organisational immunity in social insects. Curr Op Insect Sci 5:1–15.
9. Stern DL, Foster WA (1996) The evolution of soldiers in aphids. Biol Rev Camb Philos Soc 71:27–79.
10. Aoki S, Kurosu U (2010) A review of the biology of Cerataphidini (Hemiptera, Aphidiidae). Entomol Zool (Jpn) 21:73–88.
11. Kurosu U, Aoki S (2009) Extremely long-closed galls of a social aphid. Jpn J Appl Entomol Zool 53:575–586.
12. Uematsu K, Kutsukake M, Fukatsu T, Shimada M, Shibao H (2010) Altruistic colony defense by menopausal female insects. Curr Biol 20:1182–1186.
13. Ito Y, Hattori I (1983) Relationship between Nola innocua Butler (Lepidoptera: Nolidae), A. pisum also collected from Curr Op Insect Sci 4:47–50.
14. Rahbé Y, et al. (2002) Metabolic and symbiotic interactions in amino acid pools of the pea aphid, Acrithosiphon pisum strain Apl. J Insect Physiol 48:1937–1948.
15. Callow RK, Greenway AR, Griffiths DC (1973) Chemistry of the secretion from the coriscles of various species of aphids. J Insect Physiol 19:733–748.
16. Greenway AR, Griffiths DC (1973) A comparison of triglycerides from aphids and their coriscle secretions. J Insect Physiol 19:1649–1655.
17. Alfares S, Hijaz F, Killiny N (2016) Chemical composition of coriscle secretion of the cowpea aphid Acanthoscyphus cowpeae. Physiol Entomol 41:36–47.
18. Schmitz A, et al. (2012) The cellular immune response of the pea aphid to foreign intrusion and symbiotic challenge. PLoS One 7:e42114.
19. Arrese EL, Soulages JL (2010) Insect fat body: Energy, metabolism, and regulation. Annu Rev Entomol 55:207–225.
20. Mita K, Ichimura S, James TC (1994) Highly repetitive structure and its organization of the silk fibroin gene. J Mol Evol 38:583–592.
21. Hayashi CY, Levins RV (2000) Molecular architecture and evolution of a modular spider silk protein gene. Science 287:1477–1479.
22. Moran NA (1992) The evolution of aphid life cycles. Ann Rev Entomol 37:321–348.
23. Kutsukake et al. (2010) An insect-induced novel plant phenotype for sustaining social life in a closed system. Nat Commun 3:11187.
24. Abbot P, Wittbrodt J, Moran NA (2011) Conflict and conditional altruism in social aphid colonies. Proc Natl Acad Sci USA 98:12068–12071.
25. Johnson PCD, Whitfield JA, Foster WA, Amos W (2002) Clonal mixing in the soldier-producing aphid Pemphigus spyrothecae (Hemiptera: Aphididae). Mol Ecol 11: 1525–1531.
26. Andersson SO (2012) Cytotoxic scelerotization and tanning. Insect Molecular Biology and Biochemistry, ed Gilbert LI (Academic, Cambridge, MA), pp 167–192.
27. Wilson AC, et al. (2010) Genomic insight into the amino acid relations of the pea aphid, Acrithosiphon pisum, with its symbiotic bacterium Buchnera aphidicola. Insect Mol Biol 19:249–258.
28. Hansen AK, Moran NA (2011) Aphid genome expression reveals host-symbiotic co-operation in the production of amino acids. Proc Natl Acad Sci USA 108:2849–2854.
29. Abbot P, Wittbrodt J, Moran NA (2011) Small genome symbiont underlies cuticle hardness in beetles. Proc Natl Acad Sci USA 114:E8382–E8391.
30. Shigenobu S, Wilson ACC (2011) Genomic revelations of a mutualism: The pea aphid and its obligate bacterial symbiont. Cell Mol Life Sci 68:1297–1309.
31. Caniard C, et al. (2013) Tyrosine pathway regulation is host-mediated in the pea aphid symbiosis during late embryonic and early larval development. BMC Genomics 14:235.
32. Wilcox JS, Dunbar HE, Wolfinger RD, Moran NA (2003) Consequences of reductive evolution for gene expression in an obligate endosymbiont. Mol Microbiol 48: 1491–1500.
33. Moran NA, Dunbar HE, Wilcox JL (2005) Regulation of transcription in a reduced bacterial genome: Nutrient-provisioning genes of the obligate symbiont Buchnera aphidicola. J Bacteriol 187:4229–4237.
34. Price DRG, et al. (2014) Aplid amino acid transporter regulates glutamine supply to intracellular bacterial symbionts. Proc Natl Acad Sci USA 111:320–325.
35. Currie RJ, et al. (2014) Dynamic recruitment of amino acid transporters to the insect symbiotic interface. Mol Ecol 23:1608–1623.
36. Standen L, et al. (2013) Genomic analysis of aphid defense in pea aphid (Acrithosiphon pisum). Genome Biol 14:132.
37. Goto A, Kadowaki T, Kitagawa Y (2003) Drosophila hemolectin gene is expressed in the brachion calf. J Insect Physiol 49:507–516.
38. Napp AJ, Christensen BM (2005) Melanogenesis and associated cytotoxic reactions: Applications to insect innate immunity. Insect Biochem Mol Biol 35:443–459.
39. Kramer JK, et al. (2001) Oxidative conjugation of catechols with proteins in insect skeletal systems. Tetrahedron 57:385–392.
40. Suderman RJ, Dittmer NT, Kanost MR, Kramer JK (2006) Model reactions for insect cuticle sclerotization: Cross-linking of recombiant cuticular proteins upon their laccase-catalyzed oxidative conjugate with cuticles. Insect Biochem Mol Biol 36: 353–365.
41. Asada N, Fukumitsu T, Fujimoto K, Masuda K (1993) Activation of prophenoloxidase with 2-propanol and other organic compounds in Drosophila melanogaster. Insect Biochem Mol Biol 23:515–520.
42. Wynn GG, Boudreaux HB (1972) Structure and function of aphid coriciles. Annu Entomol Soc Am 65:157–166.
43. Miyazaki M (1987) “Metabolism of Apherids.” Apherids: Their Biology, Natural Enemies and Control, eds Minks AK, Harrewijn P (Elsevier, Amsterdam), Vol 2A, pp 1–25.
44. Edwards JJ (1966) Defense by smear: Supercooling in the cotton wax of aphides. Nature 211:73–134.
45. Uematsu K, Kutsukake M, Fukatsu T, Shimada M, Shibao H (2010) Altruistic colony defense by menopausal female insects. Curr Biol 20:1182–1186.