plant biology

Plant Biology ISSN 1435-8603

RESEARCH ARTICLE

Comparative transcriptome reprogramming in oak galls containing asexual or sexual generations of gall wasps

E. Bellows[1], M. Heatley[2], N. Shah[2], N. Archer[3], T. Giles[2] & R. Fray[1]

1 School of Biosciences, The University of Nottingham, Nottingham, UK
2 Advanced Data Analysis Centre, The University of Nottingham, Nottingham, UK
3 School of Veterinary Medicine and Science, The University of Nottingham, Nottingham, UK

Keywords
gall wasp; *Neuroterus numismalis*; *Neuroterus quercusbaccarum*; *Quercus robur*; reprogramming; transcriptome.

Correspondence
R. Fray, School of Biosciences, The University of Nottingham, Nottingham, UK.
E-mail: rupert.fray@nottingham.ac.uk

Editor
G. Scopece

Received: 5 March 2024;
Accepted: 7 May 2024

doi:10.1111/plb.13670

ABSTRACT

- Oak gall wasps have evolved strategies to manipulate the developmental pathways of their host to induce gall formation. This provides shelter and nutrients for the developing larva. Galls are entirely host tissue; however, the initiation, development, and physical appearance are controlled by the inducer. The underlying molecular mechanisms of gall formation, by which one or a small number of cells are reprogrammed and commit to a novel developmental path, are poorly understood. In this study, we sought a deeper insight into the molecular underpinnings of this process.
- Oak gall wasps have two generations each year, one sexual, and one asexual. Galls formed by these two generations exhibit a markedly different appearance. We sequenced transcriptomes of both the asexual and sexual generations of *Neuroterus quercusbaccarum* and *Neuroterus numismalis*. We then deployed Nanopore sequencing to generate long-read sequences to test the hypothesis that gall wasps introduce DNA insertions to determine gall development.
- We detected potential genome rearrangements but did not uncover any non-host DNA insertions. Transcriptome analysis revealed that transcriptomes of the sexual generations of distinct species of wasp are more similar than inter-generational comparisons from the same species of wasp.
- Our results highlight the intricate interplay between the host leaves and gall development, suggesting that season and requirements of the gall structure play a larger role than species in controlling gall development and structure.

INTRODUCTION

Plant galls are structures induced by parasites that provide shelter and nourishment to the inducer. Oak galls are predominantly caused by gall wasps of the family Cynipidae and can be structurally complex. These small wasps lay their eggs on different tissues of the oak tree – inducing an oak gall. In the UK, more than 50 different types of oak gall can be found. The different species of oak gall wasps can be identified by the type of gall that they induce. Thus, despite being entirely made of plant-derived cells, the initiation, development, and physical appearance is determined by the insect. This phenomenon makes oak galls a classic example of an extended phenotype (Dawkins & Dennett 1999).

The asexual and sexual generations of *Neuroterus quercusbaccarum* and *Neuroterus numismalis* lay their eggs in the underside of *Q. robur* leaves; the asexual generation of *N. numismalis* can also lay its eggs in the catkins of the tree. Galls of the same generation, but different species, can be found in close proximity to each other on the same leaves. These oak gall species are rare examples of cyclical parthenogenesis (or heterogony). Each year, there is one asexual generation and one sexual generation that always follow each other (Stone *et al.* 2002). In spring the sexual generation of *N. quercusbaccarum* and *N. numismalis* emerge and lay their eggs, producing currant and blister galls, respectively, each gall contains a single larva (Fig. 1a–c). The oviposition initiates the growth of the gall, which grows to house the developing larva (Stone *et al.* 2002). The asexual generations emerge in early summer and lay their eggs, producing rough spangle galls and silk button galls, respectively, again each containing a single larva (Fig. 1d–f). These galls drop off the leaves in autumn, and the wasp hibernates inside until spring, when it emerges to start the cycle again. It is hypothesised that oviposition site and timing are very important for gall formation and that meristematic tissue is needed for gall initiation (Stone *et al.* 2002). It has been noted that the eggs are always laid on young leaves, the appearance of the different generations occurs at the same times as the first and second flushes of leaves in spring and early summer (Hough 1953).

Gall wasps do not induce uncontrolled cell division such as in *Agrobacterium*-induced crown galls (Kerpen *et al.* 2019). Instead, they trigger a host-inducer, species-specific developmental programme for a new organ. The mechanism of induction is currently unknown. One hypothesis suggests that hormones such as auxins, cytokines and indole-3-acetic acid could be produced by the inducer to stimulate growth (Stone & Schönrogge 2003). A second hypothesis is that secreted proteins could be involved in gall development. In aphids, bi-cycle proteins have been shown to change plant gene expression and gall phenotype (Korgaonkar *et al.* 2021).

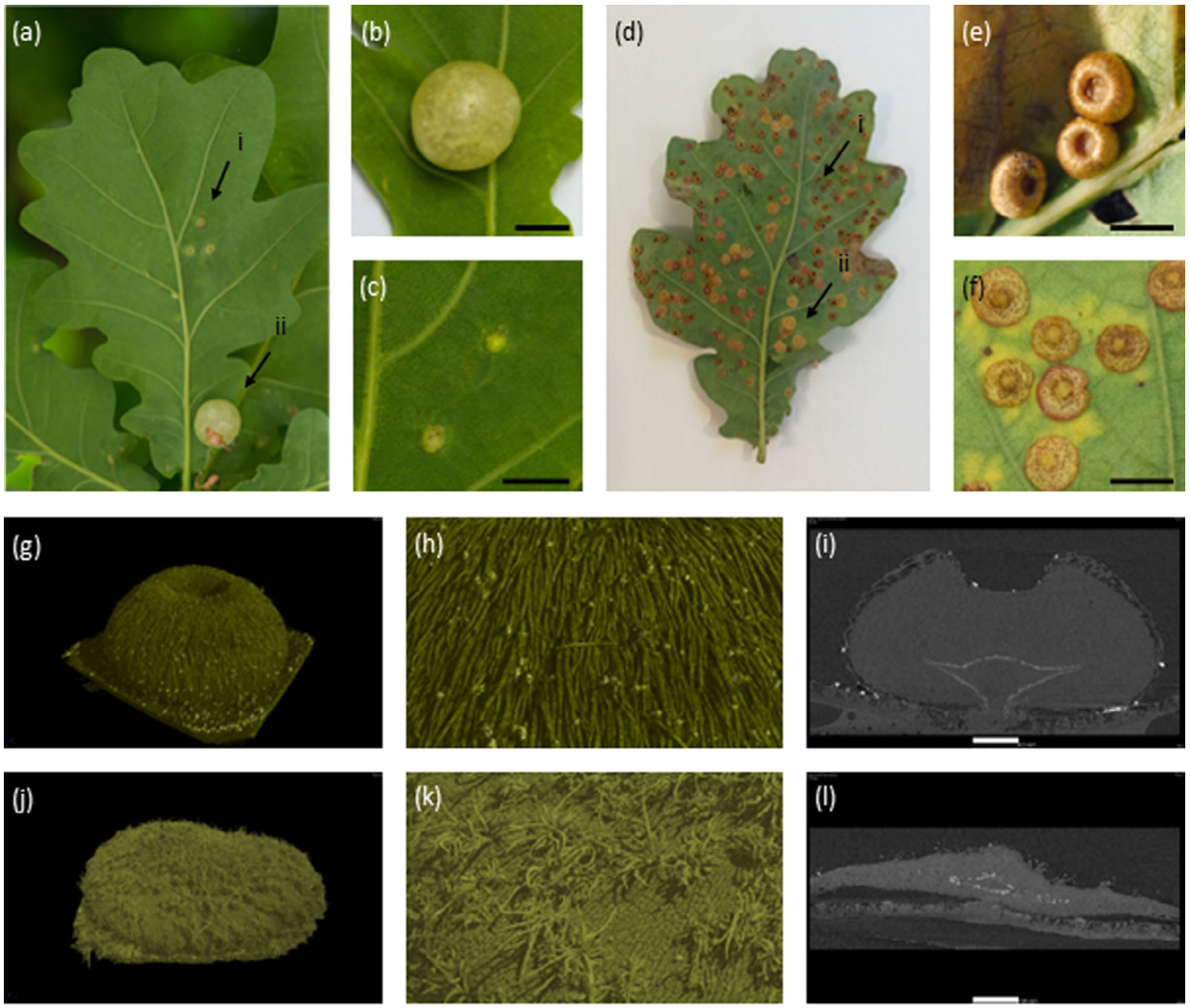

Fig. 1. Representative images and CT scans of the galls used in this study. (a) Galls containing the asexual generation of the *Neuroterus numismalis* and *Neuroterus quercusbaccarum* species, (i) blister and (ii) currant galls, respectively, appear at the same time of year on the same leaves. (b) Currant gall on an oak leaf; scale bar = 1 cm. (c) Blister gall in an oak leaf; scale bar = 1 cm. (d) Galls containing the sexual generation of the *N. numismalis* and *N. quercusbaccarum* species, (i) button and (ii) spangle galls, respectively, appear at the same time of year on the same leaves. (e) Button galls on an oak leaf; scale bar = 1 cm. (f) Spangle galls on an oak leaf; scale bar = 1 cm. (g) CT scan of a button gall. (h) Zoomed view of button gall showing trichomes. (i) CT cross-section of button gall; scale bar = 0.5 mm. (j) CT scan of a spangle gall. (k) Zoomed view of spangle gall showing the trichomes. (l) CT cross-section of spangle gall; scale bar = 0.55 mm.

A third hypothesis is that galls may be induced through novel genes introduced via exogenous DNA insertions (Cornell 1983; Jankiewicz *et al.* 2017; Gatjens-Boniche 2019). It is known that parasitic Hymenoptera wasps have symbiotic viruses or virus-like particles, which are injected with the egg into the host insect. These viruses help the developing larva to avoid the host immune system through gene transfer of virulence factors which disrupt capsule formation killing hemocytes, which are part of the insect immune system, or stopping them from adhering to foreign surfaces (Strand & Pech 1995; Moreau *et al.* 2009; Drezen *et al.* 2017).

Currently viral particles have not been found in the ovaries or venom glands of two cynipid gall wasp species, *Biorhiza pallida* and *Diplolepis rosae* (Cambier *et al.* 2019; Hearn *et al.* 2019). However, this could be a limitation of the transcriptome annotations of these species: a high proportion of transcripts from the venom glands were novel transcripts and could be of non-wasp origin. Evidence of *Wolbachia* symbiotic bacteria was also found in *de novo* assemblies of both *B. pallida* and *D. rosae* (Hearn *et al.* 2019). In some non-cynipid systems, host manipulation by bacterial symbionts has been observed. For example, in leaf mining moths, successful gall formation requires the presence of *Wolbachia* (Kaiser *et al.* 2010; Bansal *et al.* 2011; Joy 2013; Nelson *et al.* 2014). The *Wolbachia* bacteria modulate cytokine levels to create green islands in the senescent leaves where the galls are located, these islands provide a nutrient source for the developing larva (Kaiser *et al.* 2010). However, *Wolbachia* are not consistently found in gall wasps so are unlikely to play a fundamental role in gall formation (Plantard *et al.* 1999; Rokas *et al.* 2002).

The process of novel DNA insertions could be similar to that of crown galls, where a T-DNA insertion is required to hijack normal growth pathways to create a tumour-like growth. This occurs by the T-strand integrating into the host plant genome and causing malignant growth though the expression of auxin, cytokinin and opine (Tiwari *et al.* 2022). Regardless of the mechanism, the complexity of the different gall forms suggests that a controlled and regulated re-programming of plant gene expression is taking place, which would be hard to achieve by simply modulating phytohormone levels.

Oak galls have specific external and internal structures, they are comprised of three different layers: nutritional, parenchymal and epidermal (Stone *et al.* 2002). Each tissue layer provides a specific function, which differs significantly from the roles of the cells found in the original tissue. The galls attached to leaves have a broader impact on the leaf to which they are attached. Leaves with galls exhibit reduced photosynthesis, chlorophyll, and carotenoid levels (Kot *et al.* 2018b, 2020). Additionally, there are increases in free radicals and defensive responses (Kot & Rubinowska 2018; Kot *et al.* 2018a, 2019).

Previous studies have compared gene expression in morphologically different galls formed on *Glochidion obovatum*, *Eurya japonica* and *Artemisia montana* which are induced by *Caloptilia cecidophora* – a micromoth, *Borboryctis euryae* – an aphid and *Rhopalomyia yomogicola* – a gall midge, respectively (Takeda *et al.* 2019). This group found that genes associated with photosynthesis decreased in all three gall types, and genes associated with "developmental processes" were upregulated, including 38 genes which are upregulated in all species. The transcriptome of the oak gall generated by *B. pallida* on oak *Quercus robur* has been sequenced, where gene expression between galled and ungalled leaf tissue markedly differed. Genes similar to Nod factor-induced early nodulin (ENOD) genes were expressed early in development (Hearn *et al.* 2019). Proteomic analysis has been used to investigate three gall types induced by *Cynips quercusfolii*, *Cynips longiventris*, and *Neuroterus quercusbaccarum*. This analysis identified 21 proteins that showed significant change in abundance in the galls but not in the host leaf. Many of these proteins were classed as involved in "developmental regulation of plant tissue into a gall" (Pawłowski *et al.* 2017). While the levels of photosynthetic reduction and volatile production have been compared between spangle and button galls, the transcriptomes of these closely related but morphologically distinct galls have not been reported. Additionally, it is not known how the gene expression varies between the galls induced by the different asexual generations of the same species.

In this study, four galls from two wasp species were investigated. Button galls and spangle galls, containing the sexual generation and referred to here as sexual generation galls, are induced by *N. numismalis* and *N. quercusbaccarum* wasps, respectively. Similarly, blister galls and currant galls, contain the asexual generation of *N. numismalis* and *N. quercusbaccarum* wasps, respectively, and are referred to here as asexual generation galls. Previous studies have not sequenced or compared the asexual and sexual galls of the same species, or phenotypically distinct galls of closely related species that occur in the same ecological niche. Our study addresses these gaps by comparing the transcriptomes of the spangle, button, currant and blister galls. We expect that galls that look more similar will have more similar gene expression profiles. To investigate the hypothesis that DNA from the inducer organism could integrate into the host to initiate oak galls, nanopore long read sequencing was conducted on the button and spangle galls. If insertions of wasp-derived DNA into the oak genome of initiator leaf cells were required for gall induction, we would expect to be able to detect these insertions flanked by known oak sequences.

MATERIAL AND METHODS

Plant material

The four gall types were collected from the same *Q. robur* tree. The sexual generations of *Neuroterus numismalis* (button gall) and *Neuroterus quercusbaccarum* (spangle gall) were collected in August 2020. Collection of galls in August aimed to maximise the chance that the larva inside were in an active growth state. In the autumn, once the larva has developed, they enter a state of diapause before the gall drops off the leaf and the wasp inside hibernates over winter. Both button galls and spangle galls were pooled into groups of 20 to make up 50 mg needed for DNA or RNA extraction. At the same time, the leaf to which the galls were attached was collected, as were other leaves without any galls attached. Leaves were selected which only had one gall type on them. Major leaf veins were removed from the leaf samples to minimise the presence of different leaf structures, which could contribute RNA characteristics of a cell-type from which the galls were not derived. The asexual generation of *N. quercusbaccarum* (currant gall) was collected in May 2021, and the asexual generation of *N. numismalis* (blister gall) was collected in May 2022. These latter galls are larger, so 5 galls were pooled together to make up the 50 mg needed for RNA extraction. At the same time, the leaf to which the galls were attached was collected. Galls and leaves were collected in triplicate. After collection, galls were frozen in liquid nitrogen within 20 min and stored at −80 °C or processed immediately.

X-ray computed microtomography (CT) imaging and analysis

The CT scans were carried out following a modified protocol of Lundgren *et al.* (2019). Disks ca. 5 mm in diameter were excised from the leaf encompassing the oak gall and then fixed to a polystyrene block on a plastic rod. Each sample was scanned using a GE Phoenix nanotom X-ray μCT scanner (Waygate Technologies, Germany). Scan resolution was set at 2 μm, with an X-ray voltage of 72 kV and a current of 100 μA, collecting 2400 projections using a detector exposure time of 750 ms. Scan time was 30 min. 2D projections were reconstructed into 3D volumes using a filtered back-projection algorithm (Datos|X software, Waygate Technologies) prior to rendering and visualisation using VG StudioMAX software (Volume Graphics, Germany).

RNA extraction, library preparation and RNA-Seq

The RNA extraction was carried out on three biological replicates of the four galls (button, spangle, currant, blister) and the leaf to which the galls were attached. Each biological replicate consisted of 50 mg of pooled gall or 50 mg of leaf tissue ot which the galls had been attached. A modified CTAB

(Cetyltrimethylammonium bromide) method was used (Barbier *et al.* 2019; Pushkova *et al.* 2019). 0.5 ml of CTAB buffer (100 mM Tris–HCL, pH9.5, 2% CTAB, 1.4 M NaCl, 1% PEG 8000, 20 mM EDTA, 2% PVP-40, 40 mM DTT) and 2 µl proteinase K was added to the ground tissue. This was incubated for 5 min at 65 °C. 60 µl of 10% SDS were then added, the tube inverted, and 1 volume of chloroform added. The samples were vortexed for 10 s, and centrifuged (4 °C, 5500 *g*,10 min). The aqueous phase was taken and mixed with an equal volume of chloroform. After centrifugation (room temperature, 14,000 *g*, 5 min), the aqueous phase was taken and precipitated overnight with 0.5 volumes of 7.5 M ammonium acetate and 2.5 volumes of ethanol. After extraction the RNA was treated with turbo DNase (Invitrogen; AM2238) according to the manufacturer's instructions.

Total RNA was sent to Novogene for poly-A enrichment library preparation and sequencing on the Illumina NovaSeq 6000 Sequencing System.

RNA-seq data processing

This bioinformatic analysis was carried out at the Advanced Data Analysis Centre, University of Nottingham, UK. Low-quality reads and adaptor sequences were removed using the default parameters of fastp (v 0.20.1) (Chen *et al.* 2018). Reads were then processed for compatibility with the transcripts based on pseud-alignment using the default parameters of Kallisto (v 0.46.0) to quantify transcript abundances. For this quantification, the *Q. robur* transcriptome of Plomion *et al.* (2018) and (Bray *et al.* 2016) was used. Differential expression analysis was performed using Sleuth (v 0.30.1) using the default parameters (Pimentel *et al.* 2017). The differential expression compared gene expression of genes in the gall to those in the leaf to which they were attached. Genes were classified as differentially expressed if they met the criteria of a $q < 0.05$ and log2 fold-change >= 1.5.

To annotate the differentially expressed genes using Ensembl IDs, to allow GO analysis to be performed, NCBI blastn was used to map the transcripts to those of the *Quercus lobata* transcriptome (ValleyOak3.0, INSDC Assembly GCA_001633185.2). To annotate the differentially expressed genes with gene names, this was repeated using the *Q. lobata* transcriptome from NCBI (GCF_001633185.2, v 3.2).

RNA-seq analysis and visualisation

To characterise the transcription factors that are differentially expressed in the four gall types, a list of all known transcription factors in *Q. robur* was compared with our list of DEGs (Zheng *et al.* 2016). To obtain the homologues of these genes for *Q. lobata*, blastn was used to map the transcription factors to the *Q. lobata* transcriptome (ValleyOak3.0, INSDC Assembly GCA_001633185.2).

To determine enrichment, the representation factor was calculated using an online tool from the Jim Lund lab (available at http://nemates.org/MA/). Gene Ontology (GO) functional enrichment of differentially expressed genes was conducted using Gprofiler with the organism *Q. lobata* (Raudvere *et al.* 2019). To visualise the data, Venn diagrams were created using DeepVenn and the heatmap was generated using the R package heatmaply (Galili *et al.* 2017; Hulsen 2022).

DNA extraction and nanopore sequencing

The DNA was extracted from unaffected leaf, button gall and spangle gall. The button and spangle galls were pooled, with about 20 galls made up the 50 mg of tissue. DNA extraction was performed using a modified CTAB method (Barbier *et al.* 2019; Pushkova *et al.* 2019). 50 mg of galls or leaf tissue were ground and combined with 1 ml CTAB buffer (100 mM Tris–HCL, pH 9.5, 2% CTAB, 1.4 M NaCl, 1% PEG 8000, 20 mM EDTA, 2% PVP-40, 0.25% β-mercaptoethanol) and 30 µl RNase A (New England Biolabs; T3018L). This was incubated for 40 min at 65 °C before 10 µl of proteinase K (New England Biolabs; P8107S) were added and incubated for a further 20 min. This was allowed to cool to room temperature before 120 µl 10% SDS and 1 volume of chloroform weere added. The samples were mixed by vortex for 10 s and centrifuged (4 °C, 5500 *g*, 10 min). The aqueous phase was taken and mixed with an equal volume of chloroform. After centrifugation (4 °C, 5500 *g*, 10 min), the aqueous phase was taken, and the DNA was precipitated overnight with 0.8 volumes of isopropanol at −20 °C. After centrifugation (4 °C, 5500 *g*, 30 min), pellets were washed with 70% ethanol, dried, then resuspended in 50 µl distilled water.

The DNA for nanopore sequencing was further purified using a Qiagen genomic tip G/20 (Qiagen; 10223) following manufacturer's instructions. The DNA was subsequently precipitated using Short Read Eliminator (SRE) XS kit (Circulomics; SS-100-121-01) to remove any fragments <4 kb and progressively deplete fragments <10 kb.

Sequencing libraries were prepared using the Genomic DNA Ligation Kit (Oxford Nanopore Technologies; SQK-LSK110) and ca. 2 µg SRE-XS treated DNA. 250 ng of library were loaded on PromethION R9.4.1 flow cells (Oxford Nanopore Technologies; FLO-PRO002) on a PromethION Beta. During the sequencing runs, the Flow Cell Wash Kit (Oxford Nanopore Technologies; EXP-WSH004) was used halfway through to free up unavailable and blocked pores, after which the library was subsequently reloaded as described above.

Base calling was performed using the Guppy high accuracy (HAC) base calling model 2.4.

Searching for genomic insertions

This bioinformatic analysis was carried out at the Advanced Data Analysis Centre at the University of Nottingham, UK. To search for genomic insertions in the galls, a structural variant approach was taken. The nanopore sequencing from the button and spangle galls and unaffected leaf were assembled and then the gall and leaf genomes compared using structural variant analysis programs.

Raw nanopore reads were first examined via Nanoplot (v 1.38.1) before being assembled and polished via flye (v 2.9) using default settings (Lin *et al.* 2016; De Coster & Rademakers 2023). Each of these *de novo* assemblies was then assessed via BUSCO (v 5.22) against the OrthoDB Odb10 eudictot database, while contigs >10,000 bp were further examined via QUAST (v 5.02) using an upper bound assembly and settings for large, fragmented genomes (–min-contig 10000, –upper-bound-assembly, –large, –fragmented) (Mikheenko *et al.* 2018; Manni *et al.* 2021; Kuznetsov *et al.* 2022).

Structural variants (SVs) were then examined independently via Assemblytics (v 1.2.1), Minigraph (v 0.15) and Sniffles (v 1.0.12) (Nattestad & Schatz 2016; Sedlazeck *et al.* 2018; Li *et al.* 2020). For Assemblytics, leaf tissue assemblies were first individually aligned against those of each gall tissue via MUMer (v 3.23) using recommended parameters (-maxmatch, -l 100, -c 500) before calling SVs between 500 and 30,000 bp with no minimum unique sequence length (Kurtz *et al.* 2004). For Minigraph, SVs were called directly on alternative alignments performed between the same assemblies using long read mode (-x lr). For Sniffles, individual mappings were first performed for (i) leaf reads against the respective gall tissue assemblies, as well as for (ii) the respective gall tissue reads against the leaf assembly via Minimap (v 2.23) in nanopore mode and with MD tags included (-x map-ont, -MD) (Li 2018). SVs were then called on the resulting sequence alignments using default settings and examined via SURVIVOR (v 1.0.7) with the minimum read support disabled (Jeffares *et al.* 2017).

Putative indels between 500 and 30,000 bp were then extracted via custom python scripts and queried for any sequence similarities via MegaBLAST (v 2.12). For this, a custom database was used consisting of the haploid oak genome (PM1N), 32,393 microorganism genomes from the NCBI Refseq database, with an assembly level of 'chromosome' or better (31,817 bacteria, 456 archaea, 79 fungi, 41 protozoa) in addition to 315 hymenoptera genomes from the NCBI Genbank database that were categorised as either 'representative' or 'reference' genome. Among the latter, 22 assemblies belonged to members of the cynipidea family, including *N. quercusbaccarum* (accession GCA900490065.1). Because no assembly was available for *N numismalis*, the GCA900490065.1 also served as a proxy for this latter species.

In parallel, the respective nanopore reads were mapped to each of the *de novo* genome assemblies via Minimap. Those reads spanning the coordinates of putative indel sites that had matched with GCA900490065.1 were extracted via Samtools (v 1.14), together with 5000 bp flanking sequences within the respective genome assemblies (Danecek *et al.* 2021). Extracted reads were then assessed via MegaBLAST using the same custom database and putatively described as either 'oak', 'wasp' or 'chimeric' according to whether they matched either and/or both PM1N and GCA900490065.1. These categorised reads were further filtered via custom python script to extract those which had matches for the respective genome assemblies >1,000 bp and/or at least 80% identity.

RESULTS

Gall characteristics and trichome morphology in Neuroterus species galls

We undertook CT scans to better visualise the external and internal structures in live samples of the two sexual generation galls. These scans revealed the different layers of the gall and highlighted the rearrangement of trichomes on the gall surface compared to those on the leaf (Fig. 1g–l).

Trichomes on *Q. robur* leaves are typically small and have a simple, uniseriate structure with enlarged basal cell and tapered apex (Jankiewicz *et al.* 2017). However, their morphology underwent notable changes on the two gall types. On the button gall, the trichomes change to a singular fasciculate structure, characterised by long thick-walled trichomes (Fig. 1h). On the spangle gall, the trichomes have a more multiradiate structure with thick-walled trichomes which are fused at the base with the rays emanating at different levels (Fig. 1k) (Hardin 1976). These latter trichomes also significantly increase in size and number and are a distinct feature of gall morphology.

The CT imaging revealed other features, such as that the attachment between the leaf and the gall is approximately 0.2 mm in diameter. Additionally, a lighter line can be seen in the centre of both galls, which is likely to be calcium oxalate crystals, a defence mechanism of the gall (Jankiewicz *et al.* 2021).

Comparative gene expression analysis in gall formation

To investigate altered plant biological processes that are the same or differ between the galls, RNA sequencing was performed. Genes with significantly altered expression (± 1.5 log-fold 2 change, q-value <0.05) relative to the host leaf were identified (Data S1). To obtain gene names and Ensembl annotations for these genes, blastn was conducted against the *Q. lobata* transcriptome because this is better annotated than that of *Q. robur* (Table S1, Data S2 and S3).

There was significant reprogramming of gene expression in all gall types, compared to the leaf to which they were attached, with the most pronounced changes occurring in galls of the sexual generations. In spangle and button galls, 8458 and 9322 genes were upregulated, and 7063 and 6263 genes were downregulated, respectively (Table S1). Notably, there was a substantial overlap between the gene expression profiles of spangle and button galls, where 48% of the differentially expressed genes (DEG) were common to both species. The representation factor for this overlap is 1.1 (P = 9.5e-119), indicating a slightly larger overlap than expected (Fig. 2a). In contrast, fewer genes were differentially expressed in galls of the asexual generation, with 4486 and 2567 upregulated and 4115 and 1327 downregulated genes identified for currant and blister galls, respectively (Table S1). 18.2% of the DEG were shared between the two gall types, the representation factor is 1.5 (P = 4.036e-118) suggesting a larger overlap than was expected (Fig. 2b). The button and the blister galls shared 12% of the DEGs, with a representation factor of 0.9 (P = 4.2e-18), and the spangle and currant galls shared 22.4% of DEGs, with a representation factor of 0.9 (P = 3.7e-82) (Fig. 2c, d). These observations indicate a slightly smaller overlap than expected between the two gene sets. When visualising the shared genes between the four gall types using a heatmap, it became evident that the two sexual generation galls (button and spangle) were more similar to each other than to the asexual generation galls from the same species (Fig. 2e).

Gene Ontology (GO) analysis was conducted on the DEGs that were shared between the galls in the comparison to identify overrepresented terms. These genes were broken down into upregulated, downregulated or where there was a mixture of up- and downregulated genes for the different galls in the comparison. When the four gall types were compared, terms such as 'beta-fructofuranosidase' and 'fructose-bisphsphate aldolase activity' were associated with the upregulated genes. An increase in energy demand through respiration correlated with

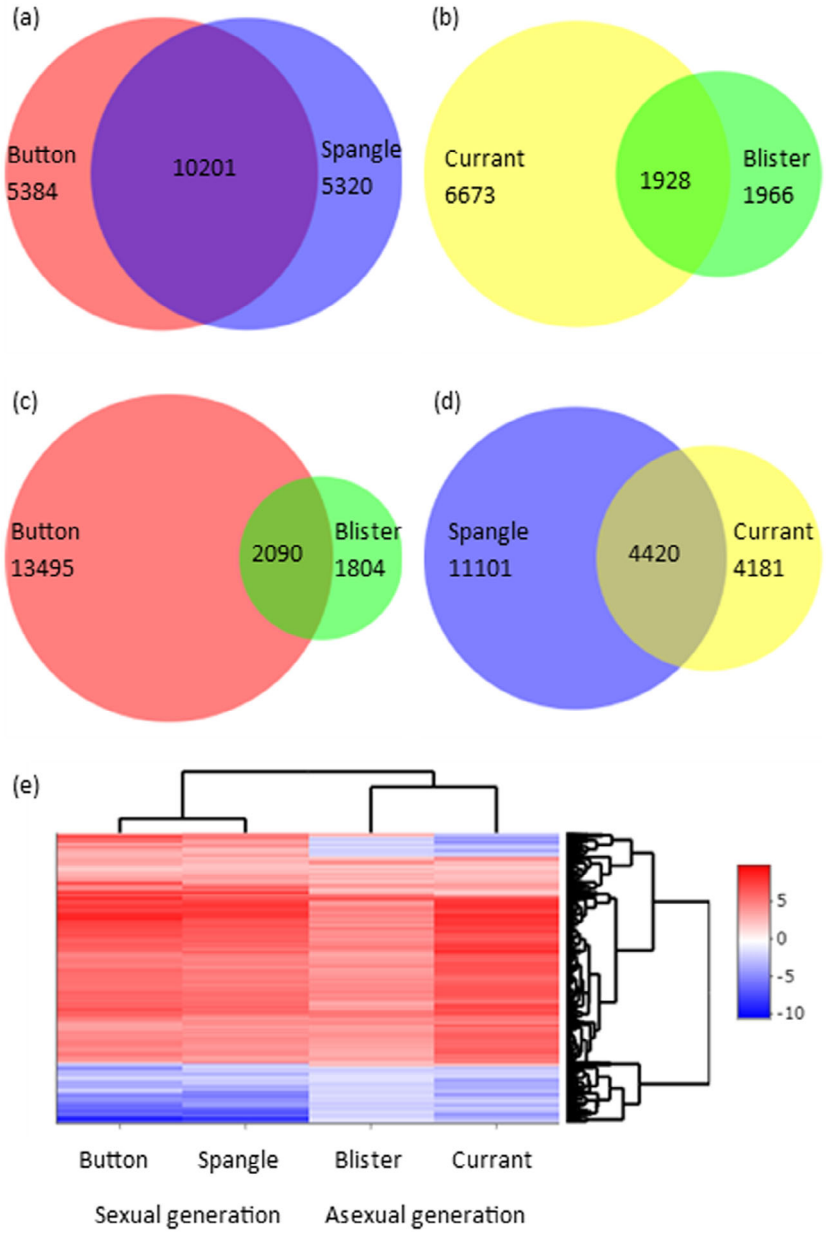

Fig. 2. Significantly up- and downregulated genes in the four gall types. (a) Venn diagram of significantly up- and downregulated differentially expressed genes in the two sexual generation galls, button and spangle (*q* value 0.05, b ± 1.5). (b) Venn diagram of significantly up- and downregulated differentially expressed genes in the two asexual generation galls, currant and blister (*q* value 0.05, b ± 1.5). (c) Venn diagram of significantly up- and downregulated differentially expressed genes in the two galls from *N. numismalis*, button and blister (*q* value 0.05, b ± 1.5). (d) Venn diagram of significantly up- and downregulated differentially expressed genes in the two galls from N. *quercusbaccarum*, spangle and currant (qvalue 0.05, b ± 1.5). (e) Heat map of the 1251 significant differentially expressed genes found in all four gall types.

the galls and their insect larva being an energy sink. In genes that changed between the gall types, the terms 'microtubule binding' and 'cytoskeleton motor activity' were present. The genes associated with these terms were upregulated in the button and spangle galls and downregulated in currant and blister galls. Interestingly, 'fructose bisphophate aldolase activity' reappeared in the downregulated GO terms. When the genes associated with the GO term were examined, fructose bisphosphate-3 and -6 were upregulated and -1 was downregulated (Table 1).

When the two asexual generation galls (currant and blister) were compared, many terms were associated with the upregulated genes, including 'response to an organic substance' and 'monocarboxylic acid binding'. The GO terms 'nucleoside triphosphate diphosphatase activity' and 'nucleoside triphosphate catabolic process' were upregulated in the blister gall but downregulated in the currant gall. 'Phloem development' was downregulated in both gall types (Table 2).

When comparing the two sexual generations (spangle and button galls), several GO terms were associated with the

Table 1. GO term analysis of the genes that are upregulated in all four gall types, that change in expression between the four types; genes that are downregulated in all four gall types.

source	term ID	term name	log qvalue
GO term for upregulated genes			
GO:MF	GO:0003824	Catalytic activity	2.529
GO:MF	GO:0004564	Beta-fructofuranosidase activity	1.915
GO:MF	GO:0008483	Transaminase activity	1.737
GO:MF	GO:0004332	Fructose-bisphophate aldolase activity	1.395
GO:BP	GO:0005975	Carbohydrate metabolic activity	1.646
GO term for genes that change between the galls			
GO:MF	GO:0003774	Cytoskeletal motor activity	2.599
GO:MF	GO:0022890	Inorganic cation transmembrane transporter activity	2.357
GO:MF	GO:0008017	Microtubule binding	1.993
GO:MF	GO:0016157	Sucrose synthase activity	1.494
GO:BP	GO:0034220	Monoatomic ion transmembrane transport	2.054
Go terms for downregulated genes			
GO:MF	GO:0008194	UDP-glycosyltransferase activity	2.384
GO:MF	GO:0004332	Fructose-bisphosphate aldolase activity	1.302

Table 2. GO term analysis of genes that are upregulated, change between galls, and are downregulated in the two asexual generation galls – blister and currant.

source	term ID	term name	log qvalue
GO terms for upregulated genes			
GO:MF	GO:0016491	Oxidoreductase activity	5.218
GO:MF	GO:0019212	Phosphatase inhibitor activity	3.838
GO:MF	GO:0033293	Monocarboxylic acid binding	3.290
GO:MF	GO:0015144	Carbohydrate transmembrane transporter activity	1.747
GO:BP	GO:1901700	Response to oxygen-containing compound	1.949
GO:BP	GO:0010033	Response to organic substance	1.811
GO:BP	GO:1901575	Organic substance catabolic process	1.509
GO:BP	GO:0005975	Carbohydrate metabolic process	1.303
GO terms for genes that change between the galls			
GO:MF	GO:0047429	Nucleoside triphosphate diphosphatase activity	1.820
GO:BP	GO:0009143	Nucleoside triphosphate catabolic process	1.565
GO terms for downregulated genes			
GO:BP	GO:0010088	Phloem development	3.898

upregulated genes, including 'microtubule', 'nuclear division' and 'chromosome'. These GO terms suggest a high level of cell division is occurring. Some genes associated with the GO term 'UDP glycosyltransferase' were upregulated in the button gall and downregulated in the spangle gall, while other genes associated with this GO term were downregulated in both types. Terms such as 'photosynthesis' and 'thylakoid' were downregulated in both gall types, indicating that the galls were becoming energy sinks, taking energy from the leaf (Table 3).

Table 3. GO term analysis of genes that are upregulated, change between galls, and are downregulated in the two sexual generation galls – button and spangle.

source	term ID	term name	log qvalue
GO terms for upregulated genes			
GO:MF	GO:0008017	Microtubule binding	22.770
GO:MF	GO:0003777	Microtubule motor activity	17.223
GO:MF	GO:0005507	Copper ion binding	9.000
GO:MF	GO:0003824	Catalytic activity	6.070
GO:MF	GO:0004185	Serine-type carboxypeptidase activity	5.022
GO:MF	GO:0005200	Structural constituent of cytoskeleton	3.681
GO:MF	GO:0030527	Structural constituent of chromatin	3.317
GO:MF	GO:0016747	Acyltransferase activity, transferring groups other than amino-acyl groups	2.633
GO:MF	GO:0004553	Hydrolase activity, hydrolyzing O-glycosyl compounds	2.251
GO:MF	GO:0016405	CoA-ligase activity	1.432
GO:MF	GO:0016878	Acid-thiol ligase activity	1.432
GO:MF	GO:0003855	3-dehydroquinate dehydratase activity	1.316
GO:BP	GO:0007017	Microtubule-based process	24.000
GO:BP	GO:0000280	Nuclear division	10.696
GO:BP	GO:0046271	Phenylpropanoid catabolic process	5.109
GO:BP	GO:0071669	Plant-type cell wall organisation or biogenesis	4.958
GO:BP	GO:0006261	DNA-templated DNA replication	4.474
GO:BP	GO:0032787	Monocarboxylic acid metabolic process	4.414
GO:BP	GO:0032392	DNA geometric change	2.934
GO:BP	GO:0032875	Regulation of DNA endoreduplication	1.604
GO:CC	GO:0005874	Microtubule	8.355
GO:CC	GO:0005576	Extracellular region	6.660
GO:CC	GO:0042555	MCM complex	3.481
GO:CC	GO:0000786	Nucleosome	3.043
GO:CC	GO:0005694	Chromosome	2.227
GO terms for genes that change between the galls			
GO:MF	GO:0008194	UDP-glycosyltransferase activity	3.684
GO terms for downregulated genes			
GO:MF	GO:0016491	Oxidoreductase activity	6.491
GO:MF	GO:0008194	UDP-glycosyltransferase activity	5.471
GO:MF	GO:0004674	Protein serine/threonine kinase activity	3.543
GO:MF	GO:0051537	2 iron, 2 sulfur cluster binding	3.460
GO:MF	GO:0022857	Transmembrane transporter activity	3.167
GO:MF	GO:0010277	Chlorophyllide a oxygenase [overall] activity	2.121
GO:MF	GO:0015276	Ligand-gated monoatomic ion channel activity	1.598
GO:MF	GO:0004888	Transmembrane signalling receptor activity	1.303
GO:BP	GO:0015979	Photosynthesis	46.000
GO:BP	GO:0048544	Recognition of pollen	3.471
GO:BP	GO:0016116	Carotenoid metabolic process	3.389
GO:BP	GO:0009657	Plastid organisation	3.320
GO:BP	GO:0042440	Pigment metabolic process	2.483
GO:BP	GO:0015994	Chlorophyll metabolic process	2.005
GO:BP	GO:0010275	NAD(P)H dehydrogenase complex assembly	1.839
GO:BP	GO:0071214	Cellular response to abiotic stimulus	1.594
GO:BP	GO:0062197	Cellular response to chemical stress	1.492
GO:BP	GO:0009416	Response to light stimulus	1.454
GO:CC	GO:0009579	Thylakoid	45.103
GO:CC	GO:0010598	NAD(P)H dehydrogenase complex (plastoquinone)	3.479

When comparing the galls of *N. quercusbaccarum*, currant and spangle, genes associated with SAM (S-adenosylmethionine) were upregulated and GO terms such as 'methionine adenosyltransferase activity' and 'S-adenosylmethionine biosynthetic process' were present. Genes associated with the terms 'DNA replication' and 'microtubule' were upregulated in spangle galls but downregulated in currant galls. Common downregulated terms in both gall types include 'photosynthesis' and 'thylakoid' (Table 4).

When the *N. numismalis* galls, blister and button, were compared there were no GO terms associated with the genes that change between these two gall types. However, GO terms such as 'Ent-kaurene oxidase activity' and 'ent-kaurene oxidation to kaurenoic acid' were associated with the upregulated genes in both gall types. This is interesting as ent-kaurene is an important step in gibberellin biosynthesis, a plant hormone that regulates many development processes, including cell elongation. Downregulated genes in both types of gall were associated with GO terms including 'photosynthesis' and 'chloroplast' (Table 5).

Table 4. GO term analysis of genes that are upregulated, change between galls, and are downregulated in the two *N. quercusbaccarum* galls – currant and spangle.

surce	term ID	term name	log qvalue
GO terms for upregulated genes			
GO:MF	GO:0003824	Catalytic activity	9.418
GO:MF	GO:0004478	Methionine adenosyltransferase activity	4.229
GO:MF	GO:0003680	Minor groove of adenine-thymine-rich DNA binding	1.951
GO:MF	GO:0008483	Transaminase activity	1.456
GO:MF	GO:0030170	Pyridoxal phosphate binding	1.449
GO:BP	GO:0044281	Small molecule metabolic process	4.917
GO:BP	GO:0006556	S-adenosylmethionine biosynthetic process	3.964
GO:BP	GO:0005975	Carbohydrate metabolic process	3.833
GO terms for genes that change between the galls			
GO:MF	GO:0008017	Microtubule binding	12.344
GO:MF	GO:0003777	Microtubule motor activity	10.610
GO:MF	GO:0030527	Structural constituent of chromatin	9.921
GO:MF	GO:0052716	Hydroquinone:oxygen oxidoreductase activity	6.288
GO:MF	GO:0046982	Protein heterodimerization activity	5.023
GO:MF	GO:0005507	Copper ion binding	4.699
GO:BP	GO:0000280	Nuclear division	13.491
GO:BP	GO:0007018	Microtubule-based movement	9.650
GO:BP	GO:0046271	Phenylpropanoid catabolic process	5.552
GO:BP	GO:0071554	Cell wall organisation or biogenesis	2.796
GO:BP	GO:0006260	DNA replication	2.669
GO:BP	GO:0051172	Negative regulation of nitrogen compound metabolic process	2.490
GO:BP	GO:0005976	Polysaccharide metabolic process	2.135
GO:CC	GO:0000786	Nucleosome	8.670
GO:CC	GO:0048046	Apoplast	4.918
GO:CC	GO:0005874	Microtubule	1.571
GO terms for downregulated genes			
GO:MF	GO:0005509	Calcium ion binding	3.029
GO:MF	GO:0016655	Oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor	2.238
GO:MF	GO:0008194	UDP-glycosyltransferase activity	1.770
GO:MF	GO:0004714	Transmembrane receptor protein tyrosine kinase activity	1.666
GO:MF	GO:0004888	Transmembrane signalling receptor activity	1.615
GO:BP	GO:0015979	Photosynthesis	22.013
GO:BP	GO:0010258	NADH dehydrogenase complex (plastoquinone) assembly	2.288
GO:CC	GO:0009579	Thylakoid	21.777
GO:CC	GO:0010598	NAD(P)H dehydrogenase complex (plastoquinone)	5.974

Transcription factor expression and trichome development in gall formation

In the majority of the transcription factor families there was a range in foldc-hange of the genes. Fewer transcription factors were differentially expressed in the blister gall, with 104 identified, compared to the other three gall types where there are 310 in the button gall, 286 in the spangle gall, and 202 in the currant gall (Data S4). This observation is consistent with the appearance of the blister gall, which is more similar to the leaf than are the other galls (Fig. 3).

There was a noticeable alteration in trichomes between button and spangle galls and oak leaves. Button galls had a singular fasciculate form at a high density, while spangle galls had a multiradiate form. The development of trichomes is controlled by a complex network of positive and negative regulators. Among these regulators, ENHANCER OF GLABRA3 (EGL3) is part of the activator complex for trichome development (Payne *et al.* 2000; Zhang *et al.* 2003). In button and spangle galls, this gene was observed to be upregulated, with a 2.4 log2 fold change (q = 6.98e-12) and 2.9 log2 fold change (q = 3.14e-8), respectively. EGL3 was not differentially expressed in the currant or blister galls, which lack trichomes.

There is no evidence of novel DNA insertions in the oak genome of sexual generation galls

It has been hypothesised that DNA insertions into the host genome could trigger the development of the gall structure, in a mechanism potentially analogous to the induction of crown

Table 5. GO term analysis of genes that are upregulated and downregulated in the two *N. numismalis* galls – button and blister.

source	term ID	term name	log qvalue
GO terms for upregulated genes			
GO:MF	GO:0052615	Ent-kaurene oxidase activity	1.739
GO:MF	GO:0003824	Catalytic activity	1.333
GO:BP	GO:0046394	Carboxylic acid biosynthetic process	2.041
GO:BP	GO:0009058	Biosynthetic process	1.776
GO:BP	GO:0010241	Ent-kaurene oxidation to kaurenoic acid	1.317
GO terms for downregulated genes			
GO:MF	GO:0008194	UDP-glycosyltransferase activity	4.659
GO:BP	GO:0009765	Photosynthesis, light harvesting	1.467
GO:CC	GO:0009507	Chloroplast	3.440

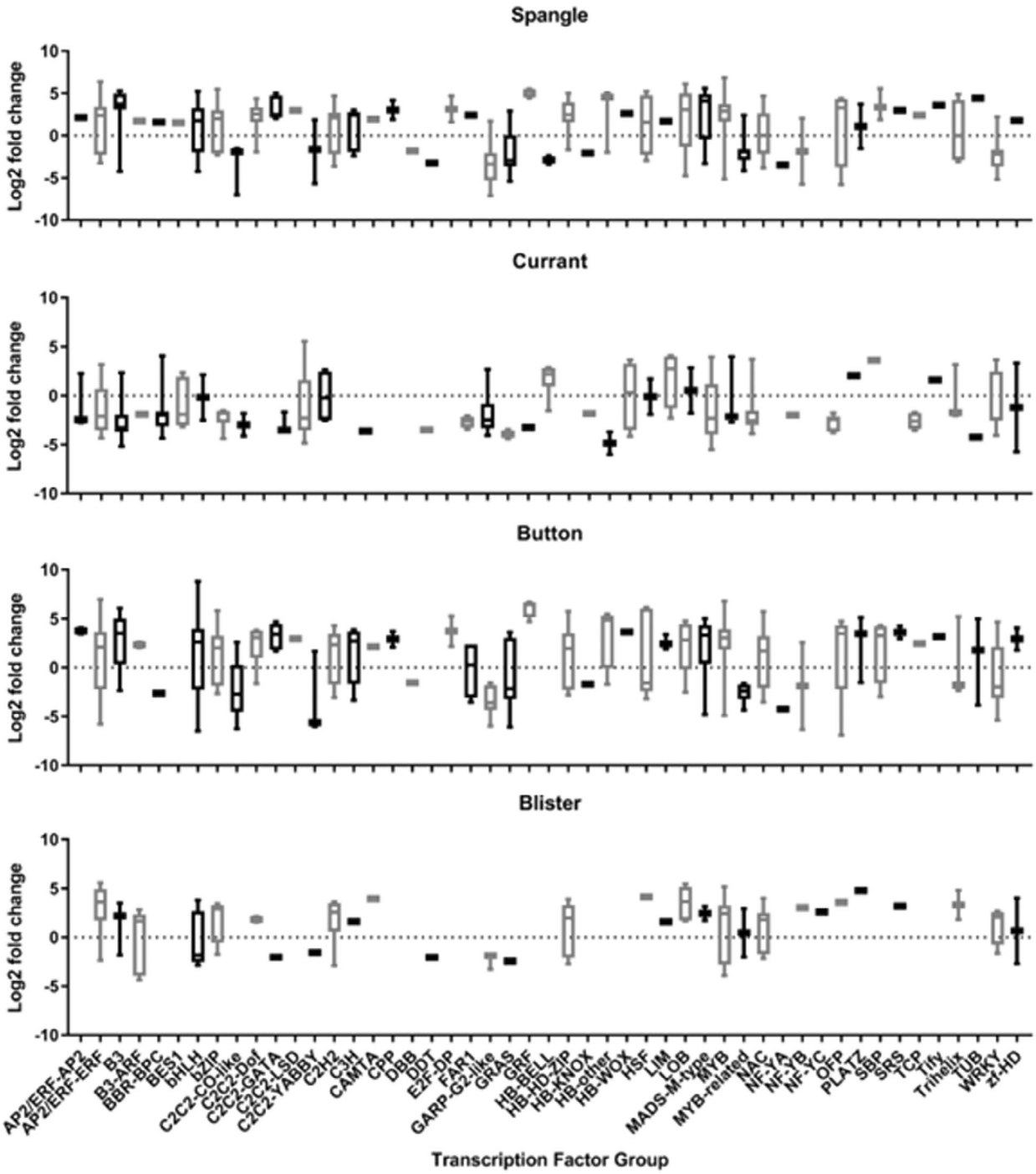

Fig. 3. Fold change in transcription factor groups in the four gall types.

gall tumours by *Agrobacterium tumefaciens* (Chilton *et al*. 1980). To explore this hypothesis, we conducted Nanopore long read sequencing of genomic DNA from button and spangle galls, as well as from a leaf of the host plant which did not have any galls attached. 86 GB of Oxford Nanopore Technologies (ONT) data for leaf (consisting of 7,198,195 reads with a mean read length of 11,954.7 and mean Q score of 13); 96 GB of ONT data for button gall (consisting of 5,284,520 reads with a mean read length of 18,277.7 and a mean Q score of 12.9) and 50 GB of ONT data for spangle gall (consisting of 5,116,190 reads with a mean read length of 9961 and a mean Q score of 12.5) were assembled using Flye. This resulted in a leaf genome with a N50 of 255784 and a length of 1,175,651,838 bp; button genome with a N50 of 317134 and a length of 1,697,337,048 bp; a spangle genome with a N50 of 1,73,653 and a length of 1,269,969,265 bp (Table 6). A structural variant approach was taken to find potential insertions in the gall genomes by comparing the spangle or button gall

Table 6. Summary of the genome assemblies.

	leaf	spangle gall	button gall
Assembly size	1,172,816,392	1,263,866,190	1,696,237,861
Number of contigs	7,215	10,929	7,727
Contigs ≥5,000 bp	7,485	11,561	7,835
Contigs ≥50,000 bp	5,843	7,465	7,041
Contigs N50 (bp)	255,784	173,653	317,134
Contigs N75 (bp)	128,757	92,514	171,541
Contigs L50	1,228	1,909	1,448
Contigs L75	2,865	4,437	3,276
Largest contig (bp)	2,489,449	6,650,687	3,430,085
GC content (%)	36.11	36.61	35.76
BUSCO completeness (%)	97.7	97.0	98

genome with the leaf genome to see if insertions or deletions could be found between them.

Through application of three independent structural variant approaches, thousands of potential insertions in the button and spangle galls were identified. The potential insertions were assessed to see if they matched the oak genome or the wasp, *N. querousbaccarum*, genome using mega blast, *N. numismalis* has not currently been sequenced. Insertions that matched the respective genome assemblies for >1000 bp and/or at least 80% identity were further analysed (Data S5). Subsequent examination of the putative insertions revealed that they did not resemble any known DNA insertion mechanism.

DISCUSSION

Comparative imaging of the sexual generations

This study aimed to investigate similarities and differences between the asexual and sexual generations of *N. quercusbaccarum* and *N. numismalis*. The CT imaging of the two sexual galls showed distinct differences in trichome organisation. Trichomes on the spangle gall have a multiradiate structure, whilst on the button gall, they have a solitary fasciculate structure. The trichomes on both gall types are much larger and at higher density than those found on the host leaf.

The CT imaging also revealed the small size of the attachment, approximately 0.2 mm, and showed the previously described calcium oxalate layer surrounding the larval growth chamber (Jankiewicz *et al.* 2021). CT imaging shows the density differences of the internal structures of the gall and enables a 3D image of the entire gall structure to be built. This approach provides a different perspective to the scanning electron images of these two galls currently published (Jankiewicz *et al.* 2017; Jankiewicz *et al.* 2021).

Novel insights into oak gall transcriptomes and transcription factor regulation

Multiple galls made up each biological replicate; however, all galls were sampled from one tree. This experimental design was chosen to ensure an isogenic background to facilitate genome assembly when searching for any novel insertions/rearrangements when comparing to the gall genome. The same tissues (three repeats each) were also used for the transcriptomics; however, this raises the potential weakness that any tree-to-tree variation or expression differences caused by local environmental influences will not have been revealed.

The gall structures have a very different appearance to that of the leaf to which they are attached, and so it is implicit that the transcriptomes of the leaf would be very different to those of the galls. The transcriptomes of several oak galls induced by various species of gall wasp have been reported (Hearn *et al.* 2019; Martinson *et al.* 2022). However, this study is the first analysis of *N. quercusbaccarum* and *N. numismalis* gall transcriptomes and, to our knowledge, the first comparison between asexual and sexual gall generations. Sequencing of the succulent oak gall induced by *Dryocosmus quercuspalustris* on *Q. rubra* revealed general trends, such as a decrease in photosynthesis and an increase in glycolysis (Martinson *et al.* 2022), similar to the gall types analysed here. In oak apples, which are induced on *Q. robur* by *B. pallida*, gene expression patterns diverge from normal bud development as the gall matures (Hearn *et al.* 2019).

We have revealed that galls found in the same niche at the same time have similar global reprogramming of plant gene expression. Button and spangle galls have a large overlap between their transcription profiles, with 48% of their DEGs being shared. This observation may be attributed to several factors: both galls are induced on the same leaf stage, and leaf stage has been documented to have a crucial role in successful gall initiation (Hough 1953). Additionally, the larva hibernates inside the gall over winter, and the similar internal structures imply similar gene expression.

The observed changes in transcription factor expression likely underpin the global transcriptional changes that result in formation of galls with wasp-specific phenotypes. To better understand the gall-forming process, it is necessary to unravel the hierarchy, regulatory networks, and mechanism(s) of initial induction of these transcription factors, as well as their target genes. There are 1163 transcription factors in the ITAK list for *Q. robur* (Zheng *et al.* 2016). In the button and spangle galls a higher number of transcription factors were differentially expressed, 310 and 286, respectively, compared to the currant and blister galls (202 and 104, respectively). One transcription factor, EGL3, a gene known to be required for trichome initiation, is highly expressed in button and spangle galls, though not differentially expressed in the galls of the trichomeless asexual generation. Transcription factor expression in oak galls of other species has not been specifically investigated. However, in wild grapevines (*Vitis riparia*), the expression of MYB33 is downregulated in late stage galls induced by the phylloxera *Daktulosphaira vitifoliae*, suggesting a role in supressing gibberellin signalling and, potentially, floral development in these structures (Schultz *et al.* 2019).

Potential mechanisms for parasitoid-induced gall formation

Our analysis detected thousands of structural variants in the genomes of the spangle and button galls. However, the sequences of the structural variants that could be found in both *N. quercusbaccarum* and the respective gall genome were ribosome sequences. These potential rearrangements or insertions are found in areas which are conserved throughout the evolutionary tree and occur at many loci, and so may simply be

due to local repeat expansions or contractions. The genome assemblies that we produced for this work have a high completeness; however, the contiguity is low but comparable to other tree genomes (Plomion *et al.* 2018). It remains unclear whether these findings are artefacts of the assembly process or expansions in these regions.

While the ribosome was long thought of as just the factory that made the proteins required for the cell, there is increasing evidence to support its role in translation regulation (Genuth & Barna 2018). Ribosome heterogeneity has been linked to key roles in development and differentiation in many organisms, from humans to plants (Martinez-Seidel *et al.* 2020; Norris *et al.* 2021). Ribosomal expansions have been demonstrated to play important functional roles as well as providing protein-binding sites (Fujii *et al.* 2018). Ribosomal proteins have also been linked to galling diseases, such as clubroot, where changes in ribosome proteins are used as markers to detect the geographic origin of the strain, as well as its virulence (Laila *et al.* 2017; Javed *et al.* 2023). Ribosome proteins in rice have also been linked to resistance against gall midges (Moin *et al.* 2021).

The results from this study provide an insight into the altered gene expression associated with asexual and sexual galls produced by two closely related wasp species, *N. quercusbaccarum* and *N. numismalis*. We did not find evidence of foreign DNA insertions into the oak genome of the gall, although there was some evidence of structural rearrangements (rRNA genes) of the genome. It remains an interesting question as to how a few cells in the leaf are induced to divide and differentiate into such structurally defined and characteristic gall forms.

AUTHOR CONTRIBUTIONS

EB: Investigation, methodology, writing – original draft preparation, formal analysis, visualization. **MH:** Formal analysis, data curation. **NS:** Formal analysis, data curation. **NA:** Supervision, formal analysis. **TG:** Supervision. **RF:** Conceptualization, funding acquisition, supervision, writing – review & editing.

ACKNOWLEDGEMENTS

This work was supported by the Leverhulme Trust, grant/award number; RPG-2020-284. We acknowledge the following people for their contribution to the study. Nanopore sequencing was conducted by Deepseq (University of Nottingham). The CT scanning was performed at the Houndsfield facility (University of Nottingham) by Dr Craig Sturrock. We also thank Laura Holt for her photography of the oak galls.

CONFLICT OF INTEREST

The authors have no conflict of interest.

DATA AVAILABILITY STATEMENT

The RNA sequencing data reported in this paper have been deposited in the NCBI GEO database, accession number GSE244168 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE244168). The DNA sequencing reported in this paper has been deposited in the SRA database Bioproject ID number PRJNA1027974 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1027974).

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1. Differentially expressed genes in the galls (button, spangle, currant, blister) compared to the leaf to which each was attached.

Data S2. Top blastn results (against *Q. lobata* transcriptome) to obtain the gene names for the DEG in the galls (button, spangle, currant, blister) compared to the leaf from which it was attached.

Data S3. Top blastn results (against *Q. lobata* transcriptome) to obtain the Ensembl IDs for the DEG in the galls (button, spangle, currant, blister) compared to the leaf from which it was attached.

Data S4. List of Transcription factors expressed in *Q. robur*, blastn results to obtain *Q. lobata* Ensembl ID, and lists of transcription factors that are differentially expressed in the galls (button, spangle, blister, currant).

Data S5. Sequences of the potential insertions/rearrangements found in the two sexual generation galls.

Table S1. The transcriptome that was used in the alignments was not annotated.

Table S2. The number of possible insertions from the structural variant analysis comparing the gall genomes to the leaf control genome.

REFERENCES

Bansal R., Hulbert S., Schemerhorn B., Reese J.C., Whitworth R.J., Stuart J.J., Chen M.S. (2011) Hessian fly-associated bacteria: transmission, essentiality, and composition. *PLoS One*, **6**, e23170.

Barbier F.F., Chabikwa T.G., Ahsan M.U., Cook S.E., Powell R., Tanurdzic M., Beveridge C.A. (2019) A phenol/chloroform-free method to extract nucleic acids from recalcitrant, woody tropical species for gene expression and sequencing. *Plant Methods*, **15**, 62.

Bray N.L., Pimentel H., Melsted P., Pachter L. (2016) Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, **34**, 525–527.

Cambier S., Ginis O., Moreau S.J.M., Gayral P., Hearn J., Stone G.N., Giron D., Huguet E., Drezen J.M. (2019) Gall wasp transcriptomes unravel potential effectors involved in molecular dialogues with oak and rose. *Frontiers in Physiology*, **10**, 926.

Chen S., Zhou Y., Chen Y., Gu J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**, i884–i890.

Chilton M.D., Saiki R.K., Yadav N., Gordon M.P., Quetier F. (1980) T-DNA from Agrobacterium Ti plasmid is in the nuclear DNA fraction of crown gall tumor cells. *Proceedings of the National Academy of Sciences of the United States of America*, **77**, 4060–4064.

Cornell H.V. (1983) The secondary chemistry and complex morphology of galls formed by the Cynipinae (Hymenoptera): why and how? *The American Midland Naturalist*, **110**, 225–234.

Danecek P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O., Whitwham A., Keane T., Mccarthy S.A., Davies R.M., Li H. (2021) Twelve years of SAMtools and BCFtools. *GigaScience*, **10**, giab008.

Dawkins R., Dennett D.C. (1999) *The extended phenotype: the long reach of the gene*. Oxford University Press, Oxford, UK.

De Coster W., Rademakers R. (2023) NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics*, **39**, btad311.

Drezen J.M., Leobold M., Bézier A., Huguet E., Volkoff A.N., Herniou E.A. (2017) Endogenous viruses of parasitic wasps: variations on a common theme. *Current Opinion in Virology*, **25**, 41–48.

Fujii K., Susanto T.T., Saurabh S., Barna M. (2018) Decoding the function of expansion segments in ribosomes. *Molecular Cell*, **72**, 1013–1020.e6.

Galili T., O'Callaghan A., Sidi J., Sievert C. (2017) heatmaply: an R package for creating interactive cluster

heatmaps for online publishing. *Bioinformatics*, **34**, 1600–1602.

Gatjens-Boniche O. (2019) The mechanism of plant gall induction by insects: revealing clues, facts, and consequences in a cross-kingdom complex interaction. *Revista de Biología Tropical*, **67**, 1359–1382.

Genuth N.R., Barna M. (2018) The discovery of ribosome heterogeneity and its implications for gene regulation and organismal life. *Molecular Cell*, **71**, 364–374.

Hardin J.W. (1976) Terminology and classification of Quercus trichomes. *Journal of the Elisha Mitchell Scientific Society*, **92**, 151–161.

Hearn J., Blaxter M., Schönrogge K., Nieves-Aldrey J.L., Pujade-Villar J., Huguet E., Drezen J.M., Shorthouse J.D., Stone G.N. (2019) Genomic dissection of an extended phenotype: oak galling by a cynipid gall wasp. *PLoS Genetics*, **15**, e1008398.

Hough J.S. (1953) Studies on the common spangle gall of oak. III. The importance of the stage in laminar extension of the host leaf. *New Phytologist*, **52**, 229–237.

Hulsen T. (2022) DeepVenn – A web application for the creation of area-proportional Venn diagrams using the deep learning framework tensorflow.js ArXiv, abs/2210.04597.

Jankiewicz L.S., Dyki B., Machlańska A., Dubert F. (2017) Oak leaf galls: *Neuroterus numismalis* and *Cynips quercusfolii*, their structure and ultrastructure. *Acta Societatis Botanicorum Poloniae*, **86**, 3537–3550.

Jankiewicz L.S., Guzicka M., Marasek-Ciolakowska A. (2021) Anatomy and Ultrastructure of Galls Induced by *Neuroterus quercusbaccarum* (Hymenoptera: Cynipidae) on Oak Leaves (*Quercus robur*). *Insects*, **12**, 850–865.

Javed M.A., Schwelm A., Zamani-Noor N., Salih R., Silvestre Vañó M., Wu J., González García M., Heick T.M., Luo C., Prakash P., Pérez-López E. (2023) The clubroot pathogen *Plasmodiophora brassicae*: a profile update. *Molecular Plant Pathology*, **24**, 89–106.

Jeffares D.C., Jolly C., Hoti M., Speed D., Shaw L., Rallis C., Balloux F., Dessimoz C., Bähler J., Sedlazeck F.J. (2017) Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nature Communications*, **8**, 14061.

Joy J.B. (2013) Symbiosis catalyses niche expansion and diversification. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **280**, 20122820.

Kaiser W., Huguet E., Casas J., Commin C., Giron D. (2010) Plant green-Island phenotype induced by leaf-miners is mediated by bacterial symbionts. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **277**, 2311–2319.

Kerpen L., Niccolini L., Licausi F., Van Dongen J.T., Weits D.A. (2019) Hypoxic conditions in crown galls induce plant anaerobic responses that support tumor proliferation. *Frontiers in Plant Science*, **10**, 56.

Korgaonkar A., Han C., Lemire A.L., Siwanowicz I., Bennouna D., Kopec R.E., Andolfatto P., Shigenobu S., Stern D.L. (2021) A novel family of secreted insect proteins linked to plant gall development. *Current Biology*, **31**, 1836–1849.e12.

Kot I., Jakubczyk A., Karaś M., Złotek U. (2018a) Biochemical responses induced in galls of three Cynipidae species in oak trees. *Bulletin of Entomological Research*, **108**, 494–500.

Kot I., Rubinowska K. (2018) Physiological response of pedunculate oak trees to gall-inducing Cynipidae. *Environmental Entomology*, **47**, 669–675.

Kot I., Rubinowska K., Michałek W. (2018b) Changes in chlorophyll *a* fluorescence and pigments composition in oak leaves with galls of two cynipid species (Hymenoptera, cynipidae). *Acta Scientiarum Polonorum, Hortorum Cultus*, **17**, 147–157.

Kot I., Sempruch C., Chrzanowski G., Czerniewicz P. (2019) The effect of leaf galls of Cynipidae on accumulation and biosynthesis of plant amines in oak trees. *Biochemical Systematics and Ecology*, **83**, 26–32.

Kot I., Sempruch C., Rubinowska K., Michałek W. (2020) Effect of *Neuroterus quercusbaccarum* (L.) galls on physiological and biochemical response of *Quercus robur* leaves. *Bulletin of Entomological Research*, **110**, 34–43.

Kurtz S., Phillippy A., Delcher A.L., Smoot M., Shumway M., Antonescu C., Salzberg S.L. (2004) Versatile and open software for comparing large genomes. *Genome Biology*, **5**, R12.

Kuznetsov D., Tegenfeldt F., Manni M., Seppey M., Berkeley M., Kriventseva E.V., Zdobnov E.M. (2022) OrthoDB v11: annotation of orthologs in the widest sampling of organismal diversity. *Nucleic Acids Research*, **51**, D445–D451.

Laila R., Robin A.H., Yang K., Choi G.J., Park J.I., Nou I.S. (2017) Detection of ribosomal DNA sequence polymorphisms in the protist *Plasmodiophora brassicae* for the identification of geographical isolates. *International Journal of Molecular Sciences*, **18**, 84.

Li H. (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 3094–3100.

Li H., Feng X., Chu C. (2020) The design and construction of reference pangenome graphs with minigraph. *Genome Biology*, **21**, 265.

Lin Y., Yuan J., Kolmogorov M., Shen M.W., Chaisson M., Pevzner P.A. (2016) Assembly of long error-prone reads using de Bruijn graphs. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E8396–E8405.

Lundgren M.R., Mathers A., Baillie A.L., Dunn J., Wilson M.J., Hunt L., Pajor R., Fradera-Soler M., Rolfe S., Osborne C.P., Sturrock C.J., Gray J.E., Mooney S.J., Fleming A.J. (2019) Mesophyll porosity is modulated by the presence of functional stomata. *Nature Communications*, **10**, 2825.

Manni M., Berkeley M.R., Seppey M., Simão F.A., Zdobnov E.M. (2021) BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, **38**, 4647–4654.

Martinez-Seidel F., Beine-Golovchuk O., Hsieh Y.-C., Kopka J. (2020) Systematic review of plant ribosome heterogeneity and specialization. *Frontiers in Plant Science*, **11**, 948.

Martinson E.O., Werren J.H., Egan S.P. (2022) Tissue-specific gene expression shows a cynipid wasp repurposes oak host gene networks to create a complex and novel parasite-specific organ. *Molecular Ecology*, **31**, 3228–3240.

Mikheenko A., Prjibelski A., Saveliev V., Antipov D., Gurevich A. (2018) Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, **34**, i142–i150.

Moin M., Saha A., Bakshi A., MS M., PB K. (2021) Study on transcriptional responses and identification of ribosomal protein genes for potential resistance against brown planthopper and gall midge pests in Rice. *Current Genomics*, **22**, 98–110.

Moreau S., Huguet E., Drezen J. (2009) *Insect infection and immunity: evolution, ecology and mechanisms*. Oxford University Press, Oxford, UK.

Nattestad M., Schatz M.C. (2016) Assemblytics: a web analytics tool for the detection of variants from an assembly. *Bioinformatics*, **32**, 3021–3023.

Nelson L.A., Davies K.A., Scheffer S.J., Taylor G.S., Purcell M.F., Giblin-Davis R.M., Thornhill A.H., Yeates D.K. (2014) An emerging example of tritrophic coevolution between flies (Diptera: Fergusoninidae) and nematodes (Nematoda: Neotylenchidae) on Myrtaceae host plants. *Biological Journal of the Linnean Society*, **111**, 699–718.

Norris K., Hopes T., Aspden J.L. (2021) Ribosome heterogeneity and specialization in development. *Wiley Interdisciplinary Reviews: RNA*, **12**, e1644.

Pawłowski T.A., Staszak A.M., Karolewski P., Giertych M.J. (2017) Plant development reprogramming by cynipid gall wasp: proteomic analysis. *Acta Physiologiae Plantarum*, **39**, 114.

Payne C.T., Zhang F., Lloyd A.M. (2000) GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics*, **156**, 1349–1362.

Pimentel H., Bray N.L., Puente S., Melsted P., Pachter L. (2017) Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods*, **14**, 687–690.

Plantard, O., Rasplus, J. Y., Mondor, G., LE Clainche, I., Solignac, M. 1999. Distribution and phylogeny of *Wolbachia* inducing thelytoky in Rhoditini and 'Aylacini' (Hymenoptera: Cynipidae). *Insect Molecular Biology*, **8**, 185–191.

Plomion C., Aury J.-M., Amselem J., Leroy T., Murat F., Duplessis S., Faye S., Francillonne N., Labadie K., Le Provost G., Lesur I., Bartholomé J., Faivre-Rampant P., Kohler A., Leplé J.-C., Chantret N., Chen J., Diévart A., Alaeitabar T., Barbe V., Belser C., Bergès H., Bodénès C., Bogeat-Triboulot M.-B., Bouffaud M.-L., Brachi B., Chancerel E., Cohen D., Couloux A., Da Silva C., Dossat C., Ehrenmann F., Gaspin C., Grima-Pettenati J., Guichoux E., Hecker A., Herrmann S., Hugueney P., Hummel I., Klopp C., Lalanne C., Lascoux M., Lasserre E., Lemainque A., Desprez-Loustau M.-L., Luyten I., Madoui M.-A., Mangenot S., Marchal C., Maumus F., Mercier J., Michotey C., Panaud O., Picault N., Rouhier N., Rué O., Rustenholz C., Salin F., Soler M., Tarkka M., Velt A., Zanne A.E., Martin F., Wincker P., Quesneville H., Kremer A., Salse J. (2018) Oak genome reveals facets of long lifespan. *Nature Plants*, **4**, 440–452.

Pushkova E.N., Beniaminov A.D., Krasnov G.S., Novakovskiy R.O., Povkhova L.V., Melnikova N.V., Dmitriev A.A. (2019) Extraction of high-molecular weight DNA from poplar plants for nanopore sequencing Plant genetic resources for breeding and producing functional nutraceutical food.

Raudvere U., Kolberg L., Kuzmin I., Arak T., Adler P., Peterson H., Vilo J. (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, **47**, W191–W198.

Rokas A., Atkinson R.J., Nieves-Aldrey J.L., West S.A., Stone G.N. (2002) The incidence and diversity of *Wolbachia* in gallwasps (Hymenoptera; Cynipidae) on oak. *Molecular Ecology*, **11**, 1815–1829.

Schultz J.C., Edger P.P., Body M.J.A., Appel H.M. (2019) A galling insect activates plant reproductive programs during gall development. *Scientific Reports*, **9**, 1833.

Sedlazeck F.J., Rescheneder P., Smolka M., Fang H., Nattestad M., Von Haeseler A., Schatz M.C. (2018) Accurate detection of complex structural variations using single-molecule sequencing. *Nature Methods*, **15**, 461–468.

Stone G.N., Schönrogge K. (2003) The adaptive significance of insect gall morphology. *Trends in Ecology & Evolution*, **18**, 512–522.

Stone G.N., Schonrogge K., Atkinson R.J., Bellido D., Pujade-Villar J. (2002) The population biology of oak gall wasps (Hymenoptera: Cynipidae). *Annual Review of Entomology*, **47**, 633–668.

Strand M.R., Pech L.L. (1995) Immunological basis for compatibility in parasitoid–host relationships. *Annual Review of Entomology*, **40**, 31–56.

Takeda S., Yoza M., Amano T., Ohshima I., Hirano T., Sato M.H., Sakamoto T., Kimura S. (2019) Comparative transcriptome analysis of galls from four different host plants suggests the molecular mechanism of gall development. *PLoS One*, **14**, e0223686.

Tiwari M., Mishra A.K., Chakrabarty D. (2022) Agrobacterium-mediated gene transfer: recent advancements and layered immunity in plants. *Planta*, **256**, 37.

Zhang F., Gonzalez A., Zhao M., Payne C.T., Lloyd A. (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development*, **130**, 4859–4869.

Zheng Y., Jiao C., Sun H., Rosli H.G., Pombo M.A., Zhang P., Banf M., Dai X., Martin G.B., Giovannoni J.J., Zhao P.X., Rhee S.Y., Fei Z. (2016) iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Molecular Plant*, **9**, 1667–1670.