Bakkali *et al. BMC Genomics* (2024) 25:1222
https://doi.org/10.1186/s12864-024-11020-8

BMC Genomics

RESEARCH Open Access

Comparative transcriptomics suggests a highly species-specific nature of the phenotypic plasticity associated with the outbreaks of the two main pest locusts

N. Bakkali[1†], S. Saadi[1†], A. Badih[2†] and M. Bakkali[1*]

Abstract

Background Locust outbreaks cause devastation and provide material for fundamental research. They associate with a case of phenotypic plasticity whereby the shift between the two extremes of the polyphenism (i.e., gregarious phase *versus* solitarious phase) affects behaviour as well as most aspects of the locusts' biology. The phenotypic changes imply changes in gene expression, the changes in behaviour characterize the locusts' phase change, and the changes in the Central Nervous System (CNS) control the changes in behaviour. Thus, understanding and tackling the phenomenon requires studying the gene expression changes that the locusts' CNS undergoes between phases. The genes that change expression the same way in different locusts would be ancestrally relevant for the phenomenon in general and some of those that change expression in a species-specific way would be relevant for the phenomenon in species-specific way.

Methods Here, we use available raw sequencing reads to build transcriptomes and to compare the gene expression changes that the CNS of the two main pest locusts (*Schistocerca gregaria* and *Locusta migratoria*) undergo when they turn gregarious. The differentially expressed genes resulting from this comparative study were compared with the content of the *L. migratoria* core transcriptional phase signature genes database. Our aim is to find out about the species-specificity of the phenomenon, and to highlight the genes that respond in the same way in both species.

Results The locust phase change phenomenon seems highly species-specific, very likely due to the inter-specific differences in the material used, and in the biology and life conditions of the different locust species. Research on locust outbreaks, gregariousness and swarming would therefore benefit from considering each locust species apart, and caution is needed when extrapolating results between species—as no species seems representative of all locust species. Still, the 109 genes and 39 non-annotated sequences that we found to change expression level the same way in the two main pest locusts, especially those previously reported as core transcriptional phase signature genes in *L. migratoria*'s CNS-related tissues (10 and 1, respectively), provide material for functional testing in search for important

†N. Bakkali, S. Saadi and A. Badih contributed equally to this work.

*Correspondence:
M. Bakkali
mbakkali@ugr.es

Full list of author information is available at the end of the article

genes to better understand, or to fight against locust outbreaks in a non-species-specific way. The large set of genes that respond in a species-specific way provide material for comparing, understanding and tackling the locust's phase change phenomenon in a species-specific way. The still uncharacterized transcripts that change expression either in a species-specific or the same way between the two species studied here provide material for gene discovery. Functional testing and confirmation are needed in all cases.

Keywords Locust, *Schistocerca gregaria*, *Locusta migratoria*, Outbreak, Solitarious, Gregarious, Transcriptome, Comparison

Introduction

The word locust is tightly linked to outbreaks and devastation. In fact, the difference between grasshoppers and locusts resides in the ability of the populations of the latter to outbreak, become gregarious and swarm. Locusts are therefore considered to be pests (e.g [1]), as several species are known to cause damage and their geographical distribution covers most parts of the world, including for instance *Dociostaurus maroccanus* in the Mediterranean area (e.g [2–4]). *Melanoplus* species in North America (e.g [5–7]). *S. cancellata* in South America [8, 9], *Chortoicetes terminifera* in Australia [10–13], *Locusta migratoria* in Asia and Africa (e.g [14–19]), and *Schistocerca gregaria* in Africa and Asia (see https://www.fao.org/locusts/en/). The last two species being considered the main pest locusts due to the wide world area that their populations' outbreaks affect, to the devastation they cause, and because they affect poor regions of the globe. Accordingly, most of the research is carried out on these two species (e.g [20–28]).

L. migratoria and *S. gregaria* populations' outbreaks are associated with a wide range of changes that affect almost every aspect of their biology (for instance see [21–23, 25, 29–35]). The two extremes of the resulting polyphenism consist of a gregarious phase—during outbreaks—and a solitarious phase—when the locust population is not dense and the insects carry a "normal" life. The differences between locusts of the two phases are so pronounced in some species that solitarious and gregarious individuals/populations of the same species were once described as two different species (for a short history of the research on phase polyphenism see [36]). Locusts' phase change is not due to genetic changes *senso-stricto* (i.e., differences, thus mutations, in the DNA sequence of the genome). It is product of the genic (see epigenetic) response to the environmental changes. Environmental changes, mainly life in low population density or in very crowded conditions, result in differences in gene expression between solitarious and gregarious locusts, respectively (for instance see [37, 38]).

Interestingly, the phase change of the locusts involves aspects that are common to all known locust species, such as the association of gregariousness with life in crowded conditions, increased behavioural activity, smaller body size and increased reproduction. Yet, there are also subtle and not so subtle inter-species differences in the characteristics of the phase-dependent changes that affect the biology of the locusts. For instance, the models that estimate locusts probability of being gregarious are not applicable to different species due, as we demonstrated in Blazquez and Bakkali [39], to species-specific differences in the changes of the morphological and behavioural traits between the two phases of the different locust species—for more details on some of the species-specific characteristics of the locusts' phase change see for instance [36, 40, 41].

A logical way of understanding the locusts' phase change phenomenon could therefore be that it shows both common and species-specific characteristics between the different locust species. A reasonable expectation would therefore be that the differences in gene expression between locust phases should include common as well as species-specific changes—the common changes could be of use to a better understanding of, or even fighting against, the phase change (see outbreaks) in locusts in general, and the species-specific changes could be of use for understanding, or even fighting against, the phase changes and its implications in particular species. For that, a between-species comparison of the molecular differences between soliarious and gregarious individuals should be done.

We therefore carry out here the first inter-species comparison of the transcriptomic changes in the Central Nervous System (CNS) that differentiate the solitarious from the gregarious phases of the two main and best characterized pest locusts, *L. migratoria* and *S. gregaria*. We highlight the common changes, we report the species-specific changes. We also report new uncharacterized sequences and we discuss the findings and suggest a set of genes that could be worth considering for further functional studies.

Materials and methods

Transcriptome assembly and annotation

The CNS transcriptomes of the solitarious and gregarious *S. gregaria* as well as their annotation, expression levels and comparative data were from [37].

The raw paired-end reads of the Illumina sequencing of RNAs from the solitarious and gregarious *L. migratoria* CNS are from [42]. They can be found in the NCBI

Bioproject PRJNA399820 (accessions: SRR5967009, SRR5967010, SRR5967011 and SRR5966534, SRR5966981 and SRR5966984). To make the results comparable, the raw sequencing reads of *L. migratoria* were treated just as we treated the *S. gregaria* sequencing reads in [37]. Briefly, they were assembled using *ABySS* [43] with the odd-numbered *Kmers* 19 to 95. The resulting transcriptomes were merged using *Linux cat* command before redundancy removal using *vsearch* [44] at a 95% identity cutoff. *CAP3* [45] was then used (with o=16, k=0 and p=95 as options) to further elongate the contigs. *BWA* [46] separately aligned the reads from solitarious and gregarious libraries against the assembled reference trancriptome. After processing the alignments using *Samtools* [47] and *xa2multi* (https://github.com/lh3/bwa/blob/master/xa2multi.pl), *htseq* [48] was used to count the reads that aligned against each contig of the reference transcriptome.

Annotation of the resulting *L. migratoria* transcriptome was also carried out the same way as the annotation of the *S. gregaria* transcriptome in [37]. That is, we used *BLASTx* against our local database of *Drosophila melanogaster* proteins, then against our local database of the protein sequences of *Acyrthosiphon pisum*, *Anopheles gambiae*, *Apis mellifera*, *Bombyx mori*, *Nasonia vitripennis*, *Pediulus humanus* and *Tribolium castaneum* (henceforth called the *Species* database—see [37]). The contigs that gave no significant hits were *BLASTx* searched against the NCBI *nr* database in our local server and the contigs that remained hitless were used for a local *BLASTn* search against the NCBI *nt* database. The uncharacterized contigs—those that gave no significant blast result—were considered "anonymous" and the contigs that gave the same blast result were considered as belonging to the same unigene.

Comparative transcriptomics and differential gene expression analysis

Identification of the transcripts that are shared between (common to) both locust species was carried out in a stepwise manner: (i) The contigs that had significant *BLAST* hit against the same sequence of the *BLAST* databases (those that share the same *BLAST* result) were identified using *Linux fgrep* command and spreadsheets. (ii) In the case of the sequences that had no significant *BLAST* result (anonymous sequences), a local *BLAST* database was built using the *S. gregaria* transcriptome before *BLASTn* searching the anonymous contigs of the *L. migratoria* transcriptome against it. The same was done in the inverse sense (i.e., *BLASTn* of the anonymous *S. gregaria* contigs against a database of the *L. migratoria* transcriptome). We merged the positive hits of both *BLASTn* searches (anonymous *S. gregaria* transcripts against *L. migratoria* transcriptome database and anonymous *L. migratoria* transcripts against *S. gregaria* transcriptome database). We then removed redundancies within the transcripts of each species, and we retained the anonymous transcripts of a transcriptome that had significant *BLAST* hit in the other transcriptome.

The *S. gregaria* and *L. migratoria* contigs that had a best *BLAST* hit against the same sequence of our local *Drosophila*, *Species*, NCBI *nr* or NCBI *nt* databases (see [37]), and those that have significant *BLAST* hits against each other, were considered as shared between the two locust species analysed herein. The remaining *S. gregaria* and *L. migratoria* contigs were considered as specific to the CNS transciptome of the corresponding species.

Just as we did for the *S. gregaria* transcriptome in [37], the sequencing reads that aligned to contigs of the same *L. migratoria* unigene were separately summed for the solitarious and the gregarious libraries. Only contigs and unigenes that were at least 75 bp long and with a significant *BLAST* hit or have at least a sequencing read aligned to them in at least two libraries were retained. The summed reads of each unigene and those of the individual contigs were then used for statistical comparison of the solitarious *versus* gregarious expression levels using the *Generalized Linear Model* (*GLM*) method in *EdgeR* [49] (as described in [50, 51]). Statistical significance was considered at a 0.05 level after *False Discovery Rate* (*FDR*) correction. The data of the differentially expressed genes in *L. migratoria* were compared to those reported for *S. gregeria* in [37].

A reviewer of a previous version of this manuscript rightly suggested that we look at the *L. migratoria* literature and compare our results with what was published for that species in the different works—unfortunately the only available RNAseq work for *S. gregaria* so far is the one in [37]. We therefore took the *L. migratoria* gene identifiers of the core transcriptional *L. migratoria* phase signature genes available in the http://www.locustmine.org:8080/locustmine database [52]—we used both the whole set of those genes as well as only the genes from the available CNS-related tissues (i.e., the brain, ganglia and antennae). The sequences of those genes were extracted from the *L. migratoria* genome assembly v2.4.1 [53], and a local *BLAST* database was built in our server computer using those sequences. *BLASTn* searches were then used in order to identify the general and CNS core transcriptional phase signature *L. migratoria* genes that are also differentially expressed between phases in the *S. gregaria* CNS. Comparison with our results revealed those genes that are also differentially expressed between phases both in the *L. migratoria* and in the *S. gregaria* CNS transcriptomes used for this work.

Gene Onthology (GO) analyses were carried out using *Panther Classification System* (pantherdb.org) and gene

Table 1 Sequencing statistics

Locust	Data accession	Locust phase	Total reads	Read length	Q30	%Ns	%GC
Lm	ξ	Solitarious	204,438,252	101	~90	< 0.005	40.5
		Gregarious	201,037,488				40.37
Sg	ο	Solitarious	95,259,912				42.98
		Gregarious	76,248,284				41.92

Lm: *Locusta migratoria*. *Sg*: *Schistocerca Gregaria*. ξ: PRJNA399820: Institute of Zoology, Chinese Academy of Sciences. ο: PRJNA381887: Mohammed Bakkali's Laboratory, Universidad De Granada, Spain

Table 2 Transcriptome assembly statistics

Locust	Contigs	%GC	N50	Max length
Lm	221,511 (481442)	41.14	1696 (1343)	26,662
Sg	110,764	41.72	1296	35,064

Lm: *Locusta migratoria*. *Sg*: *Schistocerca Gregaria*

networks were built using *STRING* database (https://string-db.org/) and *Drosophila melanogaster* as reference.

Results

Assembly and annotation

Quality check of the *fastq* files (Table 1) shows that sequencings of the CNS RNAs of solitarious and gregarious *L. migratoria* and *S. gregaria*, that were carried out using the same technology (Illumina) and at the same read-length level (100 bases), gave results of similar quality. They show insignificant number of undetermined bases (Ns), most of the bases (around 90%) have less than 1 in a 1000 chance of being product of error (Q30), and the %GC is similar, especially within species. The numbers of reads are similar within species but *L. migratoria* counts on a little over twice the number of *S. gregaria* sequencing reads.

While the reference transcriptome of *S. gregaria*'s CNS was from [37], the *L. migratoria* transcriptome had to be assembled *de novo* (in order to obtain both transcriptomes using the same method and make them comparable, see Material and Methods). The latter transcriptome produced around twice the number of contigs compared to *S. gregaria*'s reference transcriptome. But, the overall characteristics of both transcriptomes were similar; so that they show similar N50, largest contig length and %GC (Table 2). Supplemental files *Lm_CNS.fasta* and *Sg_CNS.fasta* provide the assembled sequences of the *L. migratoria* and *S. gregaria* CNS reference transcriptomes, respectively.

Separate stepwise *BLAST* annotations of the two transcriptomes (see Material and Methods) produced results that seem in accordance with the larger reference transcriptome assembled for *L. migratoria* compared to the one assembled for *S. gregaria* (Tables 3 and 4). Indeed, the reference transcriptome of the former species shows lower contig to unigene ratios in each *BLAST* result. However, the proportion of contigs that gave significant *BLAST* results (transcripts of known genes) compared to that of the contigs that gave no significant *BLAST* result (anonymous transcripts) are similar between the two transcriptomes (being the number of the anonymous sequences in *L. migratoria* transcriptome twice that number in the smaller *S. gregaria* transcriptome). Table S1 summarizes the *BLAST* and GO annotation results.

Table 3 Annotation statistics

Locust	Database								No Blast
	Local		Nr.		Nt.		Total		
	Con.	Uni.	Con.	Uni.	Con.	Uni.	Con.	Uni.	
Lm	30,517	15,088	28,301	19,849	12,429	2746	71,247	37,683	64,675
Sg	59,513	14,700	6243	1794	10,313	1126	76,069	17,620	34,696

Lm: *Locusta migratoria*. *Sg*: *Schistocerca Gregaria*. Local: local database of insect proteins (see [37]). Con.: contigs. Uni. Unigenes

Table 4 BLAST annotation statistics

Locust	With BLAST hit Contigs (unigenes)		With no BLAST hit Contigs (unigenes)		
	Same	Different	With BLAST hit in the other species	With no BLAST hit in neither species	Absent in the other species
Sg	23,596 (6740)	5747 (4208)	1207 (980)	4443	9673
Lm	17,846 (6740)	33,663 (15586)	12,451 (3784)	9043	59,829

Lm: *Locusta migratoria*. *Sg*: *Schistocerca Gregaria*. Same: number of sequences with the same best BLAST hit. Different: number of BLAST positive sequences that have different best BLAST hits in the *L. migratoria* and *S. gregaria* reference transcriptomes compared here

Table 5 Comparison between the transcripts that have a significant BLAST result and identification of those that show significant changes in expression level between the solitarious and gregarious states both in *Schistocerca Gregaria* and *Locusta migratoria*

Species	Common	Sig.	Sig.-Non.	Sig.-Sig.	Over	Under	Over-Over	Under-Under	Over- Under	Under-Over
Sg	6740 (23596)	4560 (5720)	4325 (5396)	235 (324)	221 (310)	14 (14)	99 (186)	10 (10)	123 (124)	4 (4)
Lm	6740 (17847)	370 (370)	135 (135)	235 (235)	101 (101)	134 (134)	99 (99)	10 (10)	4 (4)	123 (124)

Common: common to both species. Sig.: significant change in expression level between the solitarious and the gregarious states. Non.: non-significant change in expression level between the solitarious and the gregarious states. Over: over-expressed in the gregarious state. Under: under-expressed in the gregarious state. Between parentheses are the contigs. Sig.-Non.: significant change in expression level between the solitarious and the gregarious states in the species but not in the other. Sig.-Sig.: significant change in expression level between the solitarious and the gregarious states in both species. Over-over: over expressed in the gregarious state in both species. Over-Under: over-expressed in the gregarious state in the species and under-expressed in the gregarious state in the other species. Under-Under: under expressed in the gregarious state in both species

Table 6 Comparison between the transcripts that have no significant BLAST result and identification of those that show significant changes in expression level between the solitarious and gregarious states both in *Schistocerca Gregaria* and *Locusta migratoria*

Species	Common	Sig.	Sig.-Non.	Sig.-Sig.	Over	Under	Over-Over	Under-Under	Over-Under	Under-Over
Sg	3489	3489	3444	42 (45)	39 (42)	3	39 (42)	0	0	3
Lm	2693	187	145	42	42	0	42	0	3	0

Common: common to both species (i.e., they gave significant BLAST against each other). Sig.: significant change in expression level between the solitarious and the gregarious states. Non.: non-significant change in expression level between the solitarious and the gregarious states. Over: over-expressed in the gregarious state. Under: under-expressed in the gregarious state. Between parentheses are the contigs. Sig.-Non.: significant change in expression level between the solitarious and the gregarious states in the species but not in the other. Sig.-Sig.: significant change in expression level between the solitarious and the gregarious states in both species. Over-over: over expressed in the gregarious state in both species. Over-Under: over-expressed in the gregarious state in the species and under-expressed in the gregarious state in the other species. Under-Under: under expressed in the gregarious state in both species

Gene expression

Of the 6740 genes that are common to the two reference transcriptomes compared in this work, 4560 show significant difference in gene expression levels between solitarious and gregarious *S. gregaria*, wherase only 370 genes show such difference between solitarious and gregarious *L. migratoria*. Of these latter, the expression levels of 235 genes are significantly different between the solitarious and gregarious phases of both species. 109 of those 235 genes show change of expression between phases in the same direction in both species (Table 5 shows the detailed comparative transcriptomics). 99 of those 109 genes increase expression in the geragrious phase of both species, and the remaining 10 genes decrease it. Among the genes that show consistent change of expression between phases of both species, we can highlight a choline transporter-like, an odorant binding protein, a G protein subunit and a defence protein precursor.

When it comes to the anonymous sequences, after *BLAST* searching the anonymous *L. migratoria* transcripts against the *S. gregaria* transcriptome and the *S. gregaria* anonymous sequences against the *L. migratoria* transcriptome, merging the results, and removing redundancies (see Material and Methods), 3445 *S. gregaria* transcripts corresponded to 2693 *L. migratoria* transcripts. Of these, 1344 were *BLASTn* matches between 1075 *S. gregaria* anonymous sequences and 1110 *L. migratoria* anonymous sequences (i.e., these are still uncharachterized sequences that do not correspond to any known gene and that are expressed in the CNS of both species analyzed here). Of such common anonymous sequences, all the *S. gregaria* transcripts show significant difference in expression levels between the solitarious and gregarious phases of that species, wherease only 145 *L. migratoria* transcripts show such difference. 42 *S. gregaria* and 45 *L. migratoria* anonymous transcripts show significant differences in expression levels between the solitarious and gregarious states in both species. Of these all but 3 transcripts are over-expressed in the gregarious phase of both species. The latter 3 transcripts are over-expressed in the gregarious phase in *S. gregaria* but are under-expressed in the same phase of *L. migratoria* (Table 6). Table S2 shows the gene expression comparison results.

Figure 1 shows how threre is a clear overall clustering by species, so that the general distribution of gene expression levels of a phase of a species is more similar to the other phase of the same species than to the same phase of the other species. There is also a notorious species-specificity both in the levels of gene expression as well as in the differences in gene expression levels between phases. Such species-specificity applies both to all the expressed genes and transcripts (Fig. 1A), to the *BLAST* positive (i.e., known) differentially expressed transcripts (Fig. 1B), and to the anonymous differentially expressed genes (Fig. 1C). The *BLAST* positive transcripts having visibly more conserved inter-phase diferential expression in the two species.

Functional annotation

GO analysis shows that the *L. migratoria* reference transcriptome assembled here contains sequences pertaining to 4504 biological processes, whereas the sequences of the *S. gregaria* reference transcriptome used here belong to 4046 biological processes. Most of these processes (3921) appear in both transcriptomes and only 34

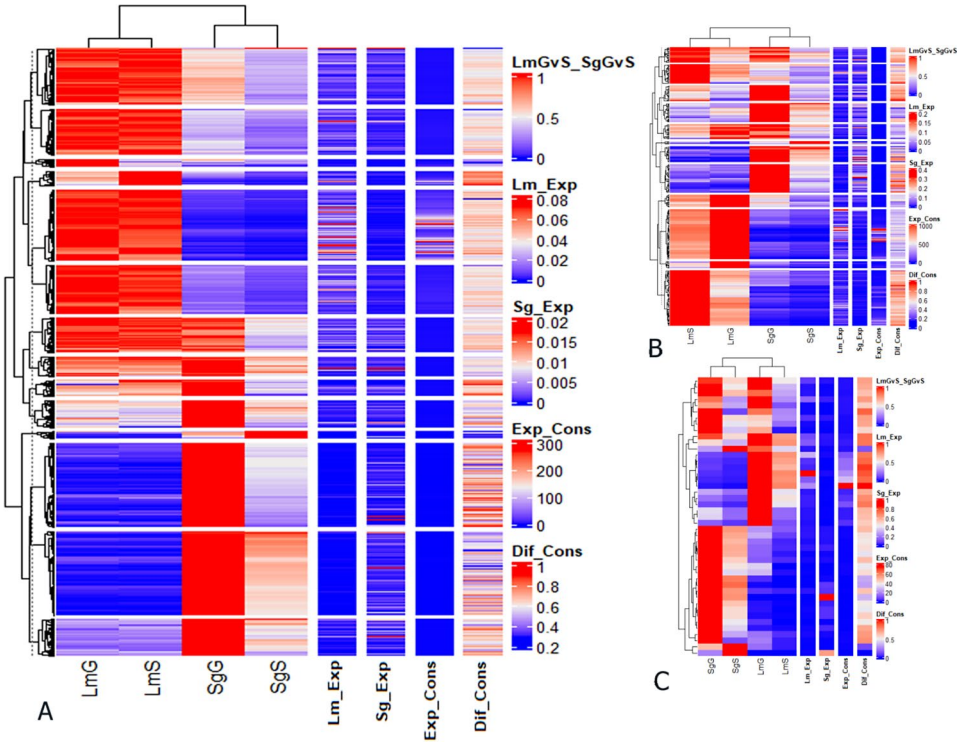

Fig. 1 Heatmaps of the overall expression levels of the common expressed (**A**) common blasted differentially expressed (**B**) and common non-annotated differentially expressed genes (**C**) of the solitarious and gregarious *Locusta migratoria* and *Schistocerca gregaria*. LmS: Solitarious *L. migratoria*, LmG: Gregarious *L. migratoria*, SgS: Solitarious *S. gregaria*, SgG: Gregarious *S. gregaria*, Lm_Exp: Average solitarious and gregarious gene expression in *L. migratoria*, Sg_Exp: Average solitarious and gregarious gene expression in *S. gregaria*, Exp_Cons: Degree of conservation of the average solitarious and gregarious gene expression level between *L. migratoria* and *S. gregaria*, Dif_Cons: Degree of conservation of the solitarious *versus* gregarious differential gene expression level between *L. migratoria* and *S. gregaria*

processes show significantly different number of genes between both transcriptomes (Fig. 2). Except for a neuropeptide signalling process, a G-protein coupled receptor signalling process and a lifespan related process, that appear increased in *S. gregaria*, the rest of those 34 significantly different biological processes are related to transcription and appear increased either in *S. gregaria* or in *L. migratoria*. The functional annotation data are in Table S1.

The sequences that have the same *BLAST* result in both species correspond to 3773 biological processes (Table S1), where the constitutive (such as transcription-related) and tissue-specific (i.e., neural) processes expectedly predominate (Fig. 3 lists the Biological Processes that contain 30 or more of the common sequences of both transcriptomes).

Among the known genes that appear in both reference transcriptomes (i.e., the sequences that have the same *BLAST* result in both transcriptomes), those that are significantly differentially expressed between solitarious and gregarious locusts in both species belong to 375 biological processes. They show a notorious presence of stress- and immunity-related processes—together with the always predominant metabolism- and transcription-related processes (Table S1, Fig. 4). The genes that increase expression in the gregarious phase of both species belong to 185 biological processes where transcription- and immunity-related processes are notorious (Table S1 and Fig. 5). The genes that decrease their level of expression in the gregarious phase of both species, however, belong to just ten biological processes; being these: positive regulation of BMP signalling pathway (GO:0030513), lipid transport (GO:0006869), lipid metabolic process (GO:0006629), inositol catabolic process (GO:0019310), imaginal disc-derived wing vein specification (GO:0007474), glyoxylate catabolic process (GO:0009436), glycolytic process (GO:0006096), glycine biosynthetic process, by transamination of glyoxylate (GO:0019265), glucose metabolic process (GO:0006006) and fatty acid metabolic process (GO:0006631).

Gene network analysis

No experimentally demonstrated direct interaction at confidence levels of 0.9 to 0.5 (highest to medium) is detected between the genes that significantly change expression level in the gregarious phases of both *L. migratoria* and *S. gregaria*—full *STRING* network. Some direct interactions appear between those genes when

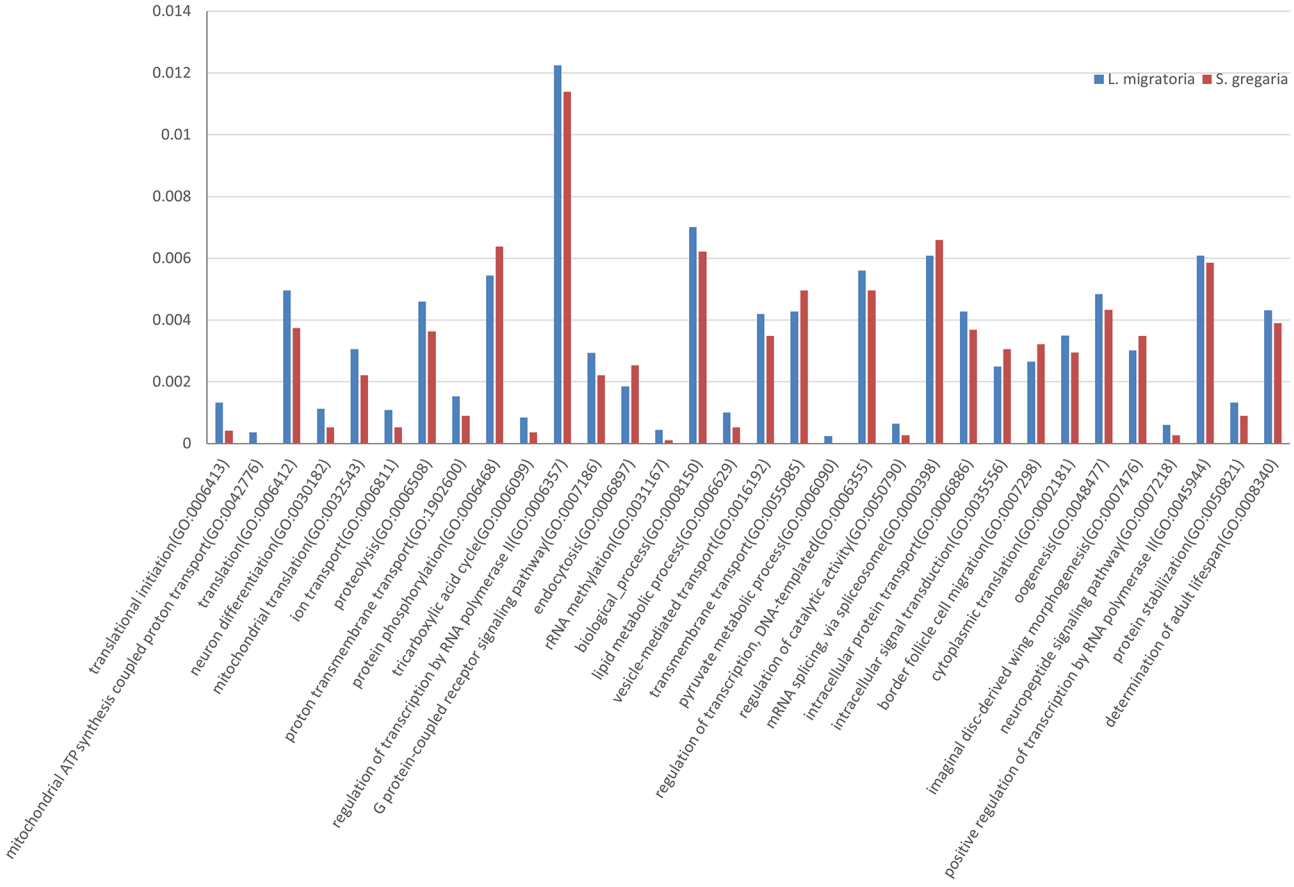

Fig. 2 The biological processes that contain significantly different numbers of sequences in the *Locusta migratoria* and *Schistocerca gregaria* reference transcriptomes compared in this work—as by Fisher's exact test. Y axis shows the proportion that the biological process represents among all the processes (number of sequences of a transcriptome that belong to a specific biological process/total number of sequences of the same transcriptome that belong to any biological process)

we add more sources of information in *STRING* (all the databases in this case, i.e., Text mining, Experiments, Databases, Co-expression, Neighbourhood, Gene Fusion, Co-occurrence). Specifically, an expected interaction emerges between the S-adenosylmethionine decarboxylase proenzyme (SamDC) and S-adenosylmethionine synthetase, isoform c (Sam-S) at the highest (0.9) confidence level (Fig. 6a). The also expected interaction between the Ankyrin repeat domain-containing protein 49; Lethal 2 35Be (l(2)35Be) and the Lethal (2) 34fc isoform b (l(2)34Fc) genes is added at high (0.7) confidence level (Fig. 6b). The network becomes somewhat more populated when we add up to 10 of the first shell of genes interacting with the members of the list of genes that significantly change expression level in the gregarious phases of both *L. migratoria* and *S. gregaria*. However, all the interacting additions are ribosomal proteins that interact with the Small subunit ribosomal protein s21e (40 S ribosomal protein S21, RpS21) present in the submitted list of genes that significantly change expression level in the gregarious phase in both species (Fig. 6c). All in all, no clear interaction network is observed between the genes that change expression between phases both in the *S. gregaria* and in the *L. migratoria* CNS transcriptomes analyzed here, and no clear hub gene is found among those genes.

Comparison with the L. Migratoria core transcriptional phase signature genes database

Of all the genes that are differentially expressed between phases in the *S. gregaria* CNS, 673 transcripts correspond to 440 *L. migratoria* core transcriptional phase signature genes. Of these, 228 *L. migratoria* core transcriptional phase signature genes were from *L. migratoria* CNS-related tissues (they correspond to 373 differentially expressed transcripts between phases in *S. gregaria* CNS), see Table S2.

Only 33 of those 228 genes gave results in a *Panther* GO search, being the biological processes signal transduction (GO:0007165), carbohydrate metabolic process (GO:0005975), metabolic process (GO:0008152), transmembrane transport (GO:0055085) and G protein-coupled receptor signalling pathway (GO:0007186) the only

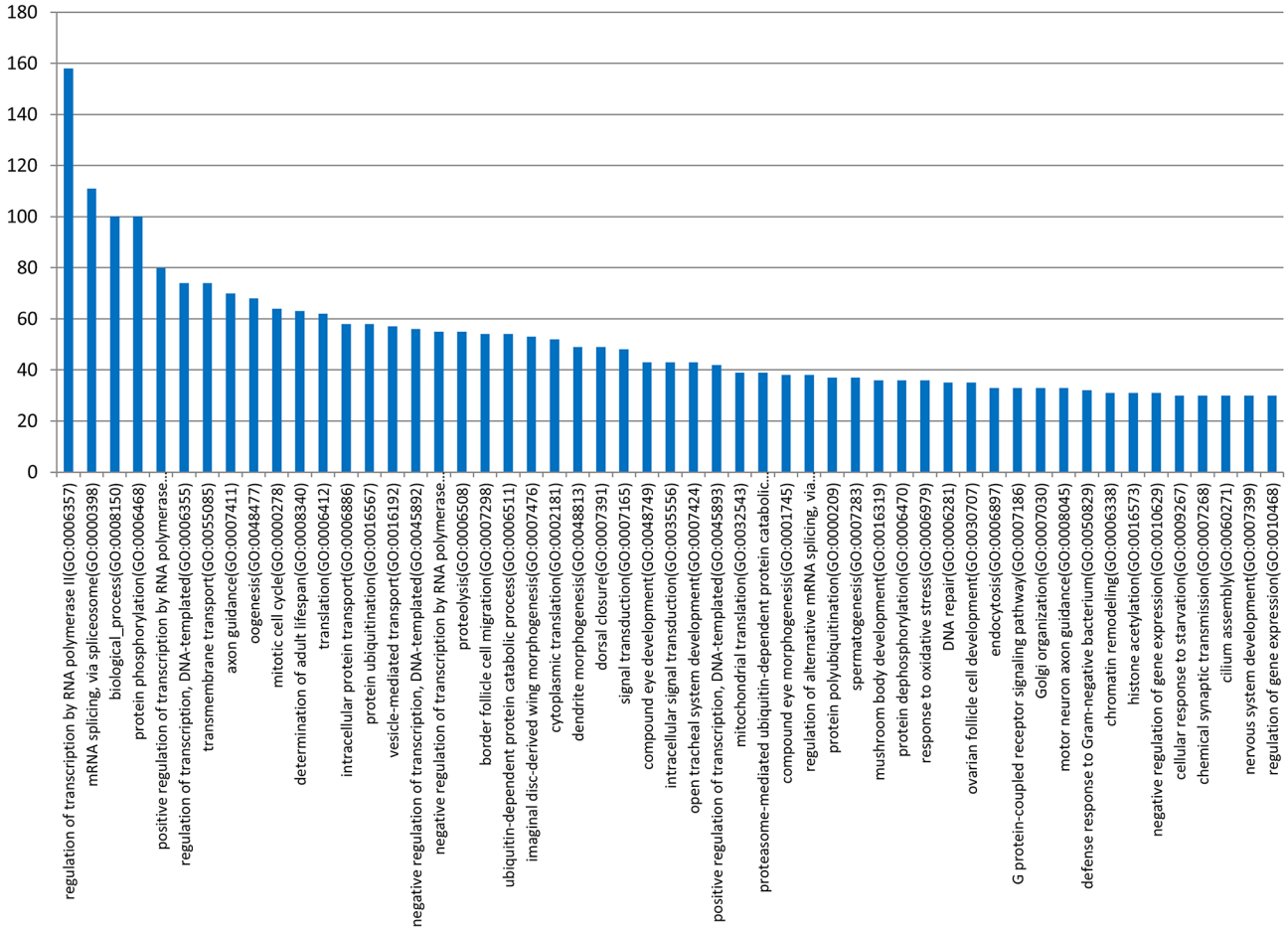

Fig. 3 The biological processes with 30 or over genes to which belong the sequences of the part of the *Locusta migratoria* and *Schistocerca gregaria* reference transcriptomes used here that have the same best BLAST hit in both transcriptomes

processes with more than one gene (see Figure S1 and the functional annotation in Table S1).

The *STRING* network of those 228 genes shows only three genes (Lysosomal alpha-mannosidase II, Ankyrin 2 and Rad17) with no direct interaction between them at confidence levels high to medium (Figure S2a). The network becomes more populated when we add up to 10 first shell interacting genes, but still there is no direct link between the sub-networks of the three query genes (Figure S2b).

Of the 235 transcripts that we found to be differentially expressed between phases both in the *S. gregaria* and in the *L. migratoria* CNS transcriptomes compared in this work, only 35 genes show change of expression level in the *L. migratoria* core transcriptional phase signature genes too, 24 of these were from the CNS-related tissues. Of these 24 genes, only 15 changed expression level between phases in the same direction (in a consistent way) in both species (11 of these in CNS-related tissues of the *L. migratoria* core transcriptional phase signature genes). 13 of the 15 genes that change expression level in the same direction increase it in both two species (10 in the CNS-related tissues of the *L. migratoria* core transcriptional phase signature genes) and two decrease it (1 in CNS-related tissues of the *L. migratoria* core transcriptional phase signature genes). Table S2 lists the aforementioned genes.

Table 7 shows the 11 *L. migratoria* CNS-related tissues core transcriptional phase signature genes that change expression level between phases in the same direction in both the *S. gregaria* and *L. migratoria* CNS transcriptomes compared in this work. The String network of those 11 genes shows only two genes (Serine/threonine-protein kinase grp, and Glucose dehydrogenase [FAD, quinone] short protein) with no direct interaction between them at confidence levels medium to high (Figure S2c) and no connection between the networks of both two genes appears when up to 10 genes of the first shell of interacting are added (Figure S2d).

3 of the 42 anonymous sequences that we found to be differentially expressed in *S. gregaria* and *L. migratoria* transcriptomes analysed in this work were also in the *L. migratoria* core transcriptional phase signature gene database. None of them appeared in CNS-related tissues

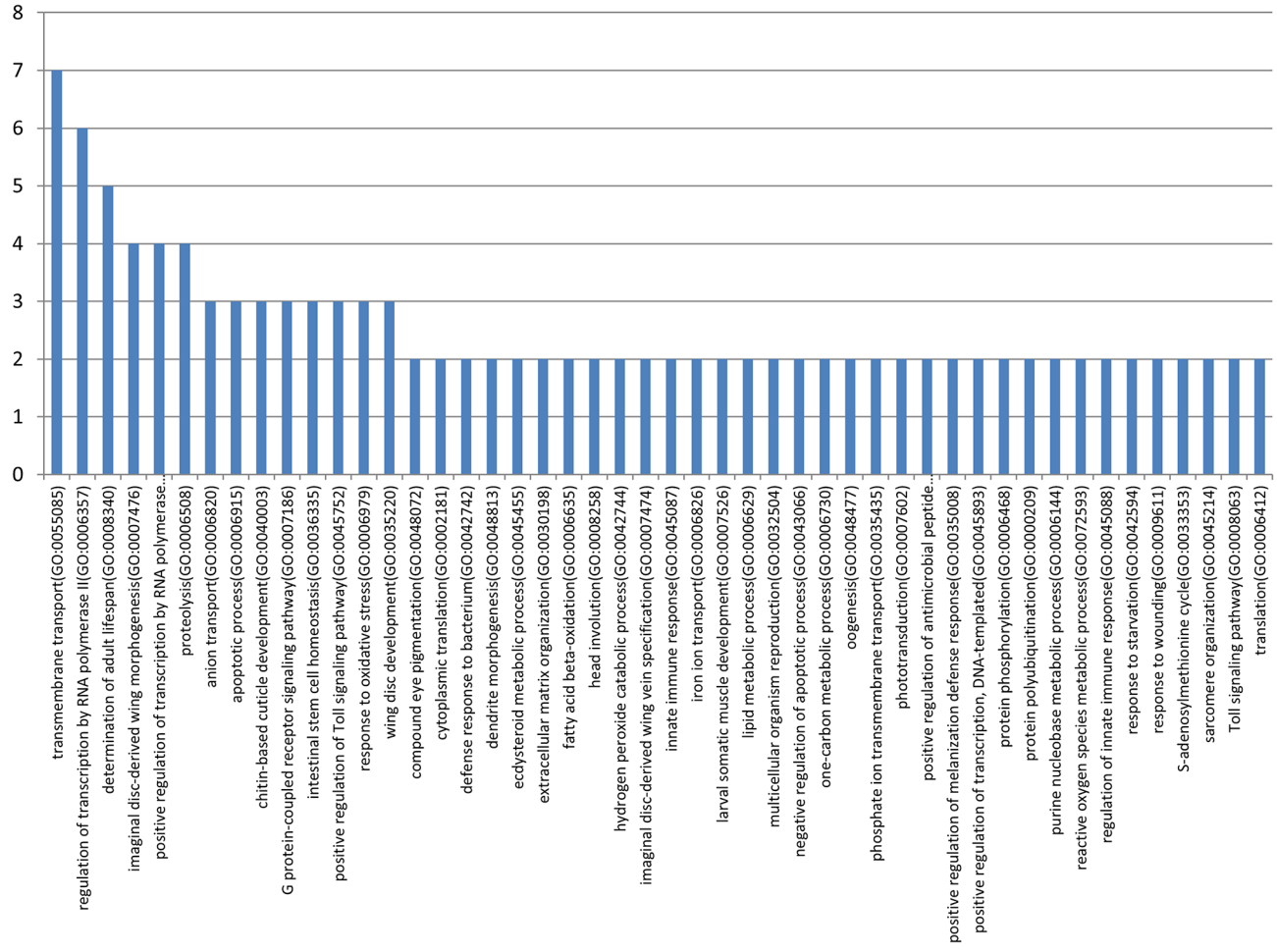

Fig. 4 Biological processes to which belong the genes that are significantly different between the solitarious and gregarious states both in *Locusta migratoria* and *Schistocerca gregaria* and that contain over one gene

in that database and only one showed consistent change of expression level (same direction of change) between phases of both species (Table S2).

Discussion

Several works researched the genetics and molecular biology of locusts' phase change. Some of these had significant impact (e.g [20, 54], for a review see [38]). Yet, until very recently, no work approached the issue from a comparative inter-specific perspective. While we were writing this manuscript, a couple of works explored this phenotypic plasticity issue using two different inter-species comparative approaches [55]. compared locust and aphid transcriptomes and produced a list of few hundred genes that show similar inter-species differences between crowded and isolated animals. For their part [56], compared locust and grasshopper species and found no single gene to react to crowding the same way between the studied species. However, the results of the first work might have been affected by the phylogenetic disparity of the studied species—large phylogenetic distances would reduce homologies and shared traits and reactions. On the other hand, the second work compared a locust (pest that outbreak) with grasshopper (non-pests that do not outbreak), which might have reduced, or left aside, the molecular aspects of the locust phase change per se. The current work is thus the first to compare two locust species. We use the two main pest locusts, *L. migratoria* and *S. gregaria*, that are widely used for studies on locusts, well characterized, and phylogenetically relatively closely related to each other.

Being the phenomenon of locust phase change so tightly linked to perception and behaviour, an obvious system of study is the Central Nervous System—that is also the target of many pesticides (e.g., pyrethroids and neonicotinoids). In addition, being the locust phase change always due to gene expression changes in response to changes in the living conditions, we aims at identifying the genes that change expression level between locust phases in a general manner (in both two locusts), and those that change expression between locust phases in a species-specific manner. We provide

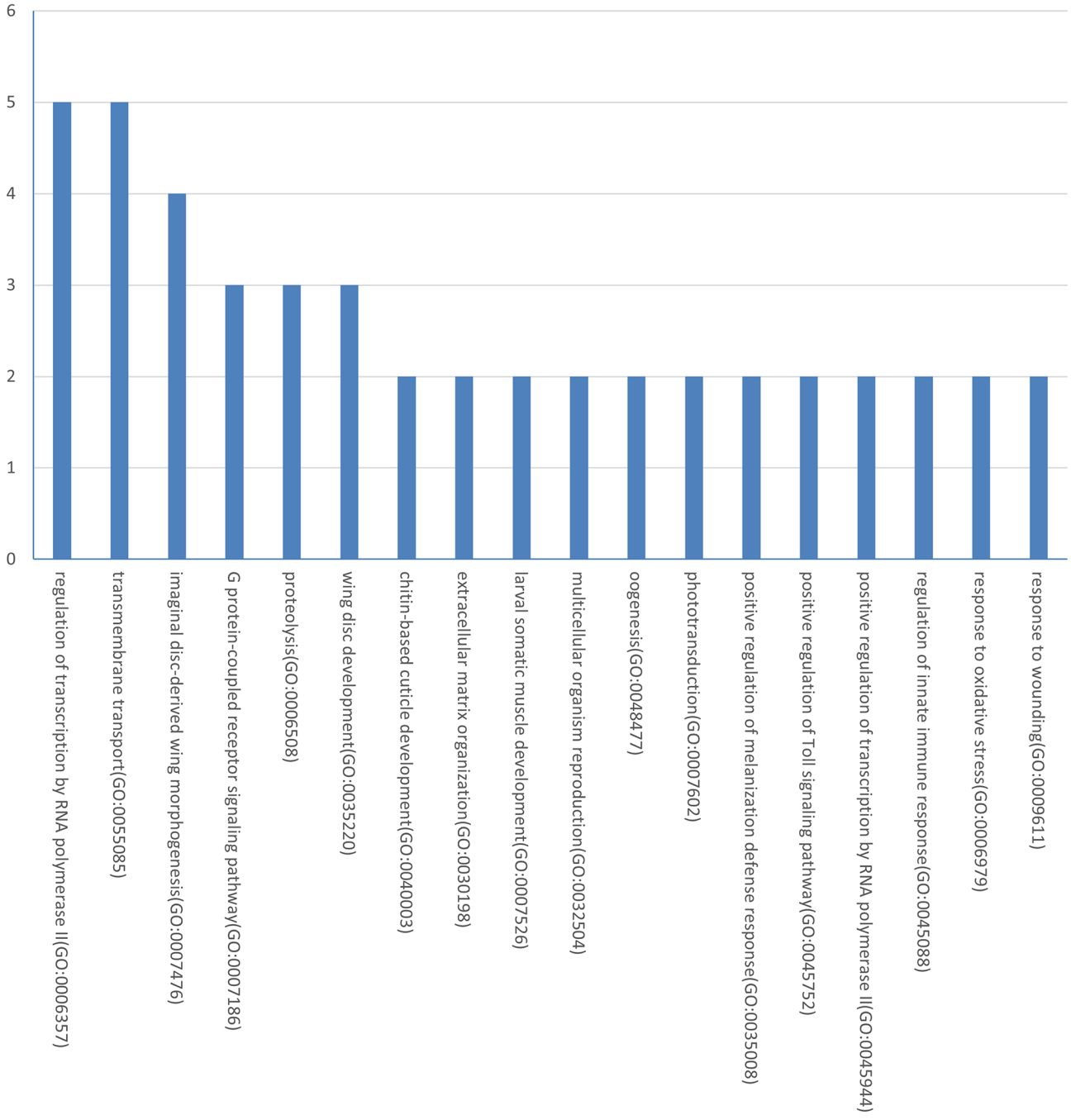

Fig. 5 Biological processes that contain more than one gene and to which belong the genes that increase their levels of expression in the gregarious phase of both *Locusta migratoria* and *Schistocerca gregaria*

lists of genes for future functional genomics studies, and we infer on the species-specificity of the locusts' phase change phenomenon. We also infer on whether findings on a single species could be extrapolated to the rest of locust species, so the efforts can focus on a (or few) representative species or, instead, studying each locust species apart is needed—bearing in mind that studying more species is always recommended, just as recommended is considering the findings on other species. We also highlight genes that consistently change their levels of expression significantly in the CNS when different locust species change phase.

The raw RNA sequencing data used here were obtained using the same sequencing technology (Illumina) and are of similarly good quality (as by number of Ns and Q30 score) as to confidently use and compare them. Their similar %GC, and the higher similarity of that percentage within than between species, suggest that the data

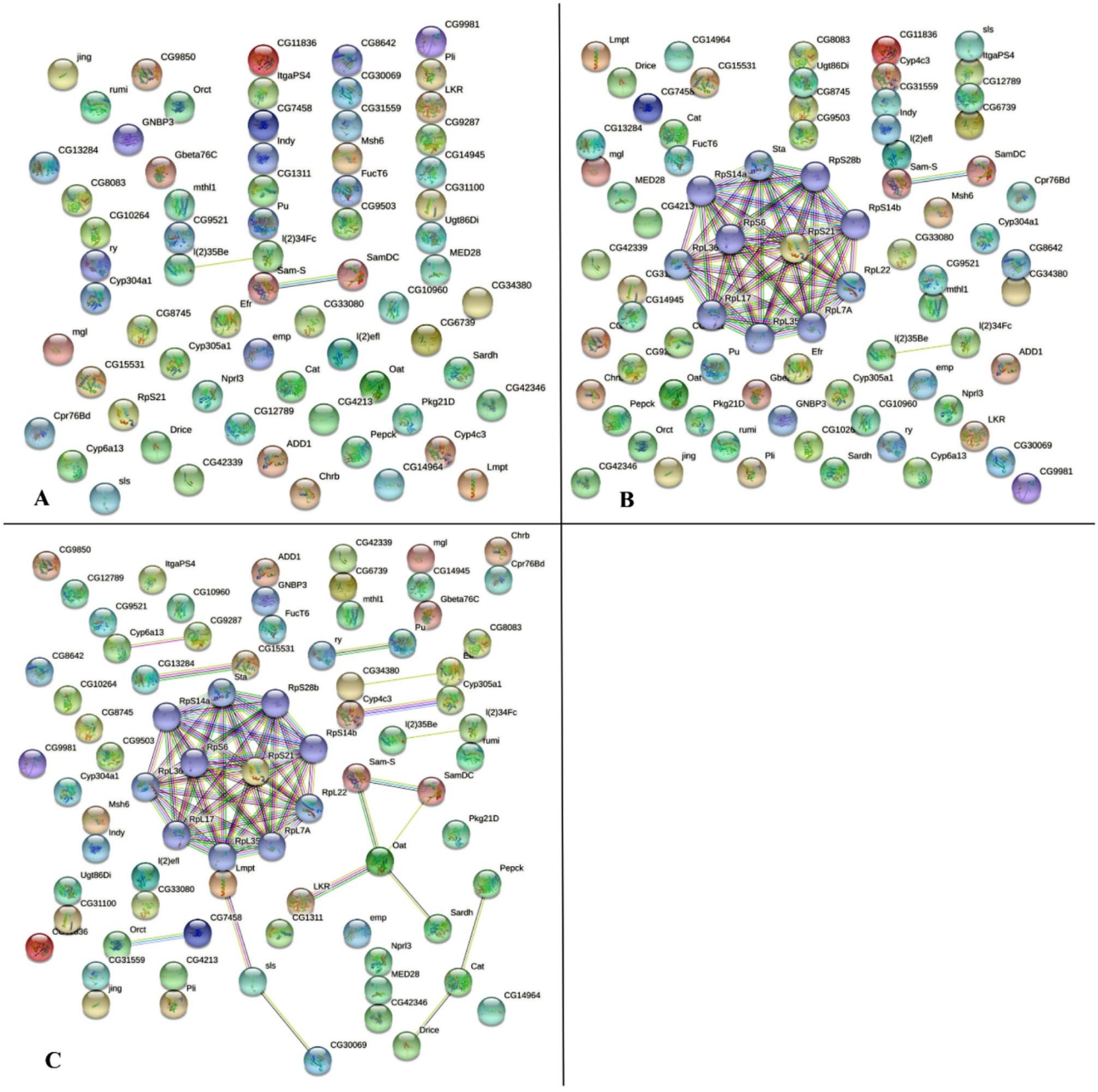

Fig. 6 Gene network analysis of the genes that show significant change in expression levels between the solitarious and the gregarious phases both in *Locusta migratoria* and in *Schistocerca gregaria*. **A**: Our gene list against all the String databases at the highest confidence level (0.9). **B**: Our gene list against all the String databases the high confidence level (0.7). **C**: Our gene list plus the first 10 genes shell in all the String databases at high (0.7) confidence level

do not contain significant contaminants. In addition, the within-species clustering and significant differences between phases confirm that the locusts used were indeed at different physiological states (i.e., solitarious and gregarious). Furthermore, the annotation results are in accordance with the data being from locusts' CNS tissues. Moreover, the number of sequencing reads for each phase is similar within each species and is quite over the minimum recommended for a good transcriptomics work (see [57]). Indeed, the raw sequencing reads used here were produced, used and published by expert laboratories as representatives of the tissue, phase and species we compare [37, 42]. In addition, we choose to process the *L. migratoria* data just as we did for the *S. gregaria* data rather than using two reference genomes or the processed data in [42]. We did that for homogeneity and to avoid any differences in the results that might be due to differences in raw data processing. Comparing *de novo* assembled transcriptomes, also allowed focussing only on the CNS-expressed part of the genome

Table 7 The 11 genes that are consistently differentially expressed between the solitarious and gregarious phases in the CNS transcriptomes of both *S. Gregaria* and *L. migratoria* compared in this work and that are reported in [52] as core transcriptional phase signature genes in *L. Migratoria* CNS-related tissues

Gene symbol	Gene name	Function	Molecular function	Biological process
TTPAL	PREDICTED: clavesin-1-like [Apis mellifera]	Alpha-tocopherol transfer protein-like	GO:0005215; transporter activity	GO:0006810; transport
Tlr13	Toll-like receptor 4 [Zootermopsis nevadensis]	Toll-like receptor 13	GO:0005515; protein binding	GO:0007165; signal transduction
TGM3	Hemocyte protein-glutamine gamma-glutamyltransferase [Camponotus floridanus]	Hemocyte protein-glutamine gamma-glutamyltransferase	GO:0003810; protein-glutamine gamma-glutamyltransferase activity	GO:0018149; peptide cross-linking
Miox	PREDICTED: inositol oxygenase-like [Apis mellifera]	Inositol oxygenase	GO:0005506; iron ion binding; GO:0050113; inositol oxygenase activity	GO:0019310; inositol catabolic process; GO:0055114; oxidation-reduction process
LOCMI14612	conserved hypothetical protein [Pediculus humanus corporis] gb|EEB18644.1| conserved hypothetical protein [Pediculus humanus corporis]	Putative uncharacterized protein	NA	NA
LOCMI10214	hypothetical protein L798_09821 [Zootermopsis nevadensis]	Putative chitin binding peritrophin-a domain protein	GO:0008061; chitin binding	GO:0006030; chitin metabolic process
LOCMI06870	NA	NA	NA	NA
l(2)efl	heat shock protein 20.7 [Schistocerca gregaria]	Protein lethal(2)essential for life	NA	NA
GRP	GNBP3 [Locusta migratoria]	Beta-1,3-glucan-binding protein	GO:0004553; hydrolase activity, hydrolyzing O-glycosyl compounds	GO:0005975; carbohydrate metabolic process
Gld	Glucose dehydrogenase [acceptor] [Zootermopsis nevadensis]	Glucose dehydrogenase [FAD, quinone]	GO:0008812; choline dehydrogenase activity; GO:0016614; oxidoreductase activity, acting on CH-OH group of donors; GO:0050660; flavin adenine dinucleotide binding	GO:0006066; alcohol metabolic process; GO:0055114; oxidation-reduction process
DDB_G0269228	PI-PLC X domain-containing protein 1 [Zootermopsis nevadensis]	PI-PLC X domain-containing protein DDB_G0269228	GO:0008081; phosphoric diester hydrolase activity	GO:0006629; lipid metabolic process

and an easier identification of the still uncharacterized sequences (anonymous) that are expressed in the two species.

The fact that the assembled CNS reference transcriptome of *L. migratoria* contained about twice the number of sequences than that of *S. gregaria* can be explained by the deeper sequencing in the former species, together with the use of the same assembly procedure to obtain both transcriptomes. However, the similar overall characteristics of both transcriptomes (N50, largest contig length and %GC) (Table 2), and the over 6000 known genes and over 4000 anonymous sequences that are common to both transcriptomes, mean that those transcriptomes could be considered as representatives of their respective species and tissue, and that they provide enough overlap as to be useful for a comparative work. In agreement with that, the GO analysis shows that both transcriptomes contain genes pertaining to similar sets of biological processes, so that over 87% of the biological processes appear in both species and only less than 1% appear significantly differentially represented in both transcriptomes. In addition, the biological processes represented in each of the two transcriptomes were as one would expect from the studied tissues (i.e., with both constitutive and neural-related processes).

The larger proportion, over ten times, of genes that show significant differential expression in *S. gregaria* compared to that proportion in *L. migatoria* is stricking. One possibility could be that the lower sequencing depth in *S. gregaria* might have caused more false positives in that species. However, that is quite unlikely, given that: (i) the sequencing depth in *S. gregaria* is over the recommended minimum threshold (see [57]), (ii) the assembled *S. gregaria* CNS transciptome in [37] showed good indicators and its annotation is the expected for a locust CNS, (iv) our laboratory PCR amplified several sequences from *S. gregaria*'s CNS transcriptome, including anonymous sequences, (v) the expression results were qPCR confirmed in vitro for several genes (see [37]), (vi) several other differentially expressed genes were confirmed by comparaison with the literature (see [37]), and (vii) the large overlap between the two locusts trascriptomes

is not likely to be obtained from insuficient sequencing or false positives,. Morever, the diference in sequencing depth would not cause such a high difference in the number of diferentially expressed genes as the one we see between both species. In addition, the over 4000 anonymous sequences present in both two transcriptomes are unlikely to happen just by chance, and indicate that, in addition to the known genes, at least the common anonymous sequences are real. Considering all these, the large difference in the number of differentially expressed genes between phases of *S. gregaria* and *L. migratoria* seems at least largely real and seems to point towards a possibly more active molecular nature of the phase polyphenism in *S. gregaria* compared to *L. migratoria*. It also suggests that the locusts phase change is complex and species-specific. Signs of such complexity and species-specificity could be infered from what we already know about the phase change in different locust species. In an earlier work [28], we found that the models for infering the phase of *S. gregaria* locusts are not useful for *L. migratoria* due to differences in the morphologic and behavioural chenges between phases in each species. In fact, some morphological, behavioural and physiological characteristics of the phase change are different between species (see [28, 36, 40, 41]). Furthermore, and in addition to laboratory induced differences, that Pener and Simpson [41] suggested as possible explaination of some stricking differences between the results of some laboratories, species-specificity might be another possible reason for some of such differences and inconsistencies. Species-specificity could explain some of the cases cited in [40, 41], as well as the case of serotonin; which was suggested to induce gregariousness in *S. gregaria* [54] and that, four years later, was associated with the solitarious phase in *L. migratoria* [58].

The fact that we found the gene expression levels to be more similar between phases of the same species than between the same phase in different species, together with the notorious differences in the levels and differential expression of genes between phases, further highlight the species-specificity of the phase polyphenism associated with the outbreak of locust populations.

The number of genes that we found to change expression levels the same way in both locusts is more similar to the number of genes that change expression levels the same way between locust and aphid species [55] than between a locust and very closely related grasshoppers [56]. While both works show more inter-specific dissimilarities than similarities in the transcriptomic answer to crowdedness, their results are incongruent with the phylogenetic distances between the compared species. Apart from unlikely but possible artifacts in one of these works, a possibility could be that locusts share more similarities in their answer to crowdedness with aphids than with grasshoppers because locusts and aphids are regularilly exposed, and are thus adapted, to crowdedness, while grasshoppers do not. However, answering this question falls beyound the scope of the current work and needs further testing.

High species-specificity seems to contrast with the cross-gregarizing effect of *S. gregaria* on *L. migratoria* and *viceversa*, reported in [59]. However, high species-specificity does not mean incompatibility, and the cross-gregarizing effect might find room within the genes that change expression level the same way in both species—among which an odorant binding protein could be a potential chemical- (see pheromone-) detecting molecule that might be involved in sensing the environment and triggering the gregarious state in more than one locust species.

Due to such species-specificity, only 235 genes show significantly different expression levels between the solitarious and gregarious phases in both *L. migratoria* and *S. gregaria* CNS transcriptomes. Furthermore, only 109 of those genes show a change in the same direction in both species; the rest (126) showing incongruent direction of change between species.

Another result that species-specificity could explain is the lack of a clear gene network and the consequent absence of hub gene(s) among the genes that change expression in both species. Indeed, the genes that change expression level in the same way between *L. migratoria* and *S. gregaria* phases do not associate to each other. The exception being the S-adenosylmethionine decarboxylase proenzyme (SamDC) and S-adenosylmethionine synthetase isoform c (Sam-S) (both related to methylation), the Ankyrin repeat domain-containing protein 49, Lethal 2 35Be (l(2)35Be) and Lethal (2) 34fc isoform b (l(2)34Fc) (both related to cell division and polarity), and the Small subunit ribosomal protein s21e (40 S ribosomal protein S21, RpS21), that is related to gene expression. Methylation genes could highlight the importance of the epigenetic regulation of gene expression during locust phase change in general. Gene expression genes could highlight the changes in gene expression as common aspect of the development of the gregarious phase in locusts—for instance, ribosomal proteins were associated to the phase change in *L. migratoria* [60]. For their part, cell polarity genes might be explained by the cell shape remodelling (especially of the neurons) required for the plastic neural response to life in crowded and stressful conditions. However, our inferences on the involvement of those genes in the locusts' phase change remain for experimental testing.

The 109 genes that change expression level in the same way when the two main pest locusts outbreak and turn gregarious is an important result. Further studies on those genes might give insights into the general

molecular basis of the locusts phase change phenomenon, and might even provide molecules of potential use for non-species-specific fight against locust outbreaks. Among those genes, the presence of less specialized (see constitutive and/or pleiotropic) genes is notorious. That might seem surprising (i.e., one would naively expect mostly specific genes; such as genes involved in neurotrasmission, response to stimulus, stress…). However, among the phase change features that are common to all locusts are the higher activity of the gregarious locusts, their increased metabolism, increased contagion, increased overall transcription, transcriptome remodeling, cuticular changes (size, shape and color)…. Such processes involve constitutive and pleotropic genes; such as genes involved in metabolism, immunity, transcription and its regulation, cuticule formation…. Several of such genes we find similarily differentially expressed between phases of the two locusts. Yet, the presence of stress-related (including starvation), G-protein coupled receptor signaling and dendrite morphogenesis processes among the genes that increase expression level in the gregarious phase of both locusts is worth highlighting. The first process could relate to the stressing life in crowded conditions, the second to perception of the stimuli resulting from such conditions, and the third to neural remodeling in response to such conditions—both three processes could thus be related to common conditions for the development of the gregarious phase in locusts. Genes such as the Choline transporter-like 1 (Ctl1), the Odorant binding protein 11, the G protein alpha subunit and the precursor of the l(2)34Fc - Defense protein l(2)34Fc precursor are to highlight.

A triply interesting result of our work are the sequences that had no significat *BLAST* result against any of the databases used here (see Material and Methods). (i) The presence of thousands of those non-annotated sequences in the CNS transcriptomes of both locusts supports our previous inference on the presence of thousands of still uncharacterized transcripts in locusts' genomes [37]. (ii) Such anonymous sequences also provide material for functional characterization and gene discovery. (iii) In addition, relevant to the present work, the over 40 anonymous transcripts that show significant differences in expression levels between phases in both locusts offer further material for researching the molecular basis of the locust's phase change, and might even offer transcripts for potential locust-specific tageting of locusts.

Some of the genes that show differential expression between phases in a species-specific way might also prove to be of interest either for further understanding the molecular basis of the locust phase change and its inter-species specificities, or might even include potential targets for species-specific fight strategies. For highlighting might be genes pertaining to the neuropeptide signalling, G-protein coupled receptor signalling and lifespan related processes—processes that are differentially represented in the CNS transcriptomes of the locust species studied here.

Being the genes that change expression level the same way in the CNS transcriptomes of both locust species very likely ancestrally related and/or important to the development and/or maintenance of the phase change in locusts, reasons for the species-specific change of expression level that most genes show between those locusts' phases are not scarce. They are expected to include not only species-specific differences in the phase change characteristics (such as those described in [28, 36, 40, 41]), but also species-specific differences in many aspects of those locusts' biology (including differences in development, morphology, behaviour, physiology, reproduction, life conditions, habitat …).

The additional comparison of our results with the content of the *L. migratoria* core transcriptional phase signature genes database further confirmed the highly species-specific nature of the locusts' phase change. In fact, the database of *L. migratoria* core transcriptional phase signature genes from CNS-related tissues contains only 10 known genes and one anonymous transcript out of the 109 known genes and 39 anonymous transcripts that we found to change expression consistently in both *L. migratoria* and *S. gregaria* CNS transcriptomes. Those 10 known genes and the anonymous sequence are thus very likely ancestrally and/or importantly related to locusts' phase change.

We therefore identify here few genes that consistently change expression level between phases in two locust species—which is unlikely to happen just by chance, or due to mistakes or differences. They could be ancestrally and/or importantly related to the development and/or maintenance of the outbreak state in locusts, and them might prove to be useful for a better understanding of, and probably even fighting against, locust outbreaks in a non-species-specific way. Not all of the many genes that we found to change expression between locust phases in an inconsistent way between the two locus species are expected to be related, specific or characteristic of the phase change. is the expression differences of some of them are expected to relate to inter-species differences in the biology, life conditions, material and its handling… rather than to differences in the phase change per se. For their part, the consistently as well as inconsistently expressed sequences that are still uncharacterized—which we called anonymous—offer material for gene discovery and might even be of potential interest to locust-specific fight strategies.

All that being said, and being this a comparative transcriptomics work, we limited ourselfs to general interpretation of the results and to highligting some genes that we

consider might be of interest. The real functional implications, importance and potential utility of each of those genes remain for functional testing—as suggested in [61]. We also highlight that this work is not void of limitations, including: (i) differences between the materials (both natural—e.g., differences in the biology and life conditions of the different locust species—and laboratory/experiment/technique-dependent—e.g., differences in physiological states of the locusts, in material handeling, and/or in technical details of RNA extraction, library construccion and sequencing). The effect of such differences would explain part of, but not all, the large inter-species differences that we report here. However, differences in the materials and hadelings would not produce the set of genes that show consistent change of expression between phases of the two locust species. Still, that latter set might also be affected by the comparison of only two species, which is another limitation of the current work. Adding more locust species to the comparison is logically expected to reduce the set of genes that show consistent change of expression between phases in different locust species—and that would be a further support to the high species-specificity of the locusts phase change phenomenon. Unfortunatly, adding more species is currently still not possible, due to the lack of similar and comparable transcriptomics data from other locust species. Finally, being this an entirely in silico work, it lacks experimental validation of the results. While qPCRs and literature searches validated the in silico results of some genes in the works that produced the raw sequencing data used here ([37] and [42]), more qPCRs would further validate the results of the genes that we report in the present work. Aditionally, and as stated before, functional testing is needed for confirmation of our interpretations as well as of the implication, importance and potential usefulness of the genes that we report and highligh here.

Conclusion

Comparing the changes in gene expression levels between the solitarious and gregarious CNS of the two main pest locust species, *L. migratoria* and *S. gregaia*, suggests that locusts' phase change, associated with locust populations outbreaks and swarming, is complex and mostly species-specific. That explains some of the reported differences and inconsistencies between works on different locust species and suggests that no one species can serve as single model for studying the phase change in all locust species.

The large set of genes that change expression differently between phases of the two locust species very likely contains genes involved in live conditions and physiological aspects that differ between species as well as genes genuinely involved in the phase change phenomenon per se. These latter might provide material for understanding and probably even dealing with the locusts' populations outbreak and swarming phenomenon in a species-specific way.

At least some of the few genes that change expression level in the CNS the same way when the two main pest locusts change phase might very likely be involved in the locust phase change in an important and/or ancestral way. They might also provide material for understanding and probably even dealing with the locusts' populations outbreak and swarming phenomenon in general.

The tens of transcriptome contigs that have no *BLAST* match in the databases (the anonymous sequences) seem genuine, for being found in two different species. Just like the known genes, those anonymous sequences that show expression level differences between the solitarious and gregarious locust CNS might provide material for either species-specific or general studies and probably even for fight strategies against locusts. They also provide material for gene discovery.

Finally, it is worth highlighting that this computer-based work is not void of limitations (e.g., effect of between-species differences in biology and life conditions, of laboratory and technical differences, and of the reduced number of species in the comparison). This work would certainly benefit from experimental validation of the results obtained in silico. The inferences on potential implications, importance and possible usefulness of the genes reported here also need experimental functional testing.

Abbreviations

CNS	Central nervous system
GO	Gene onthology
GLM	Generalized linear model
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11020-8.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Acknowledgements

M. Bakkali wants to thank the Agencia Estatal de Investigación, Ministerio de Ciencia e Innovación (Spain) for the grant Ref. PGC2018.097678.B.I00 that funded this work. M. Bakkali and S. Saadi want to thank the ERASMUS+ International Dimension Programme for the studentship at the University of Granada for the studentship to S. Saadi. We thank the anonymous reviewer who suggested comparing our findings with those published for Locusta migratoria.

Author contributions
A.B. Analyzed data, prepared some figures and tables, and participated in the interpretation of the results and in revising the manuscript. M.B. Conceived and planed the research, secured funding, data analysis, prepared some figures and tables, wrote part of the first draft of the manuscript, and participated in the interpretation of the results and in revising the manuscript. N.B. Analyzed data, prepared some figures and tables, wrote part of the first draft of the manuscript, and participated in the interpretation of the results and in revising the manuscript. S.S. Analyzed data, prepared some figures and tables, and participated in the interpretation of the results and in revising the manuscript.

Funding
Agencia Estatal de Investigación, Ministerio de Ciencia e Innovación (Spain), grant Ref. PGC2018.097678.B.I00.

Data availability
The datasets generated and/or analysed during the current study are available in the NCBI database as follows: The Locusta migratoria raw sequencing reads can be found in the Bioproject PRJNA399820.The Schistocerca gregaria raw sequencing reads can be found in the Bioproject PRJNA381887. The Locusta migratoria transcriptome shotgun assembly project has been deposited at DDBJ/EMBL/GenBank under the accession DAWVHX000000000 (Nucleotide database of the NCBI). The version described in this paper is the first version, DAWVHX010000000. The Schistocerca gregaria transcriptome shotgun assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GKTK00000000 (Nucleotide database of the NCBI). The version described in this paper is the first version, GKTK01000000. Nucleotide sequence data reported are available in the Third PartyAnnotation Section of the DDBJ/ENA/GenBank databases under the accessionnumbers TPA: BK064842-BK064843.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Departamento de Genética, Facultad de Ciencias, Universidad de Granada, Fuentenueva S/N, Granada 18071, Spain
[2]Departamento de Ciencias Naturales, Agora Granada College International School, Urbanización Llanos de Silva, S/N, Atarfe, Granada 18230, Spain

Received: 18 March 2024 / Accepted: 8 November 2024
Published online: 19 December 2024

References

1. Zhang L, Lecoq M, Latchininsky A, Hunter D. Locust and Grasshopper Management. Ann Rev Entomol. 2019;64:15–34.
2. Pasquier R, Colonna-Cesari X, Bonfils J. [The determination of the gregarigenic regions of the Moroccan cricket, Dociostaurus Maroccanus Thunbg., in Corsica]. Comptes Rendus Hebdomadaires des seances de l'Academie des Sci. 1952;235(19):1157–9.
3. Rambier A. [Gregarious and solitary populations of the Moroccan cricket (Dociostaurus Maroccanus Thunb.) In continental France]. Comptes Rendus Hebdomadaires des seances de l'Academie des Sci. 1951;233(23):1491–3.
4. Aziz Z, Nabil R, Said E, Houria N, Khadija T, Abderrahim L, Lahsen EG. Preliminary study of the intestinal microbial diversity of three Acridoidae: Oedipoda Fuscocincta, Dociostaurus Moroccanus, and Calliptamus barbarus (Acrididae: Orthoptera), in the Moroccan Middle Atlas. Indian J Microbiol. 2022;62(1):123–9.
5. Lee J, Marshall JC, Schmitz OJ, Caccone A. Genetic divergence of Connecticut Melanoplus femurrubrum populations. J Heredity. 2006;97(3):290–3.
6. Olfert O, Weiss RM, Giffen D, Vankosky MA. Modeling Ecological dynamics of a Major Agricultural Pest Insect (Melanoplus sanguinipes; Orthoptera: Acrididae): a cohort-based Approach incorporating the effects of Weather on Grasshopper Development and abundance. J Econ Entomol. 2021;114(1):122–30.
7. Dingle H, Haskell JB. Phase polymorphism in the Grasshopper Melanoplus differentialis. Science. 1967;155(3762):590–2.
8. Piou C, Zagaglia G, Medina HE, Trumper E, Rojo Brizuela X, Maeno KO. Band movement and thermoregulation in Schistocerca cancellata. J Insect Physiol. 2022;136:104328.
9. Hawkes F, Rzepka J, Gontrand G. The scent glands of the male south American Locust Schistocerca cancellata, an electron microscope study. Tissue Cell. 1987;19(5):687–703.
10. Hewitt GM, John B. The cytogenetic systems of grasshoppers and locusts. I. Chortoicetes terminifera. Chromosoma. 1971;34(3):302–23.
11. Chapuis MP, Popple JA, Simpson SJ, Estoup A, Martin JF, Steinbauer M, McCulloch L, Sword GA. Eight polymorphic microsatellite loci for the Australian plague Locust, Chortoicetes terminifera. Mol Ecol Resour. 2008;8(6):1414–6.
12. Wang B, Deveson ED, Waters C, Spessa A, Lawton D, Feng P, Liu L. Future climate change likely to reduce the Australian plague Locust (Chortoicetes terminifera) seasonal outbreaks. Sci Total Environ. 2019;668:947–57.
13. Lambert MR. Some factors affecting flight in field populations of the Australian plague Locust, Chortoicetes terminifera (walker), in New South Wales. Anim Behav. 1972;20(2):205–17.
14. Wu B, Kashiwagi T, Kuroda I, Chen XH, Tebayashi SI, Kim CS. Antifeedants against Locusta migratoria from the Japanese Cedar, Cryptomeria japonica II. Biosci Biotechnol Biochem. 2008;72(2):611–4.
15. Wang XH, Kang L. Differences in egg thermotolerance between tropical and temperate populations of the migratory Locust Locusta migratoria (Orthoptera: Acridiidae). J Insect Physiol. 2005;51(11):1277–85.
16. Chapuis MP, Lecoq M, Michalakis Y, Loiseau A, Sword GA, Piry S, Estoup A. Do outbreaks affect genetic population structure? A worldwide survey in Locusta Migratoria, a pest plagued by microsatellite null alleles. Mol Ecol. 2008;17(16):3640–53.
17. Zheng X, Zhong Y, Duan Y, Li C, Dang L, Guo Y, Ma E. Genetic variation and population structure of oriental migratory Locust, Locusta migratoria manilensis, in China by allozyme, SSRP-PCR, and AFLP markers. Biochem Genet. 2006;44(7–8):333–47.
18. Uvarov BP. A revision of the genus Locusta, L.(= Pachytylus, Fieb.), with a new theory as to the periodicity and migrations of locusts. Bull Entomol Res. 1921;12(02):135–63.
19. da Silva R, Lange AB. Tyramine as a possible neurotransmitter/neuromodulator at the spermatheca of the African migratory Locust, Locusta Migratoria. J Insect Physiol. 2008;54(8):1306–13.
20. Kang L, Chen X, Zhou Y, Liu B, Zheng W, Li R, Wang J, Yu J. The analysis of large-scale gene expression correlated to the phase changes of the migratory Locust. Proc Natl Acad Sci USA. 2004;101(51):17611–5.
21. Simpson SJ, McCaffery A, HAeGELE BF. A behavioural analysis of phase change in the desert Locust. Biol Rev. 1999;74(4):461–80.
22. Ellis PE. Changes in the social aggregation of Locust hoppers with changes in rearing conditions. Anim Behav. 1963;11(1):152–60.
23. Stower W. The colour pattern of hoppers of the Desert Locust (Schistocerca Gregaria Forskal). Anti-Locust Bull. 1959;32:75.
24. Chen S, Yang P, Jiang F, Wei Y, Ma Z, Kang L. De novo analysis of transcriptome dynamics in the migratory Locust during the development of phase traits. PLoS ONE. 2010;5(12):e15633.
25. Ayali A, Pener M. Density-dependent phase polymorphism affects response to adipokinetic hormone in Locusta. Comp Biochem Physiol Part A: Physiol. 1992;101(3):549–52.
26. Tanaka S, Nishide Y. Do desert Locust hoppers develop gregarious characteristics by watching a video? J Insect Physiol. 2012;58(8):1060–71.
27. Islam MS, Roessingh P, Simpson SJ, Mccaffery AR. Effects of population density experienced by parents during mating and oviposition on the phase of hatchling desert locusts, Schistocerca gregaria. *Proceedings of the Royal Society of London B: Biological Sciences* 1994, 257(1348):93–98.
28. Martin-Blazquez R, Chen B, Kang L, Bakkali M. Evolution, expression and association of the chemosensory protein genes with the outbreak phase of the two main pest locusts. Sci Rep. 2017;7(1):6653.

29. Roessingh P, Simpson SJ, James S. Analysis of phase-related changes in behaviour of desert Locust nymphs. Proc Royal Soc Lond B: Biol Sci. 1993;252(1333):43–9.
30. Bouaichi A, Roessingh P, Simpson SJ. An analysis of the behavioural effects of crowding and re-isolation on solitary-reared adult desert locusts (Schistocerca Gregaria) and their offspring. Physiol Entomol. 1995;20(3):199–208.
31. Simpson S, Raubenheimer D, Behmer S, Whitworth A, Wright G. A comparison of nutritional regulation in solitarious-and gregarious-phase nymphs of the desert Locust Schistocerca Gregaria. J Exp Biol. 2002;205(1):121–9.
32. Hamouda AB, Tanaka S, Hamouda MHB, Bouain A. Density-dependent phenotypic plasticity in body coloration and morphometry and its transgenerational changes in the migratory Locust, Locusta Migratoria. J Entomol Nematology. 2011;3(7):105–16.
33. Symmons P. A morphometric measure of phase in the desert Locust, Schistocerca Gregaria (Forsk). Bull Entomol Res. 1969;58(04):803–9.
34. Stower W, Davies D, Jones I. Morphometric studies of the desert Locust, Schistocerca Gregaria (Forsk). J Anim Ecol. 1960;29:309–39.
35. Davey P. Quantities of food eaten by the desert Locust, Schistocerca Gregaria (Forsk.), in relation to growth. Bull Entomol Res. 1954;45(03):539–51.
36. Pfluger HJ, Braunig P. One hundred years of phase polymorphism research in locusts. *Journal of comparative physiology A, Neuroethology, sensory, neural, and behavioral physiology* 2021, 207(3):321–326.
37. Bakkali M, Martín-Blázquez R. RNA-Seq reveals large quantitative differences between the transcriptomes of outbreak and non-outbreak locusts. Sci Rep. 2018;8(1):9207.
38. Wang X, Kang L. Molecular mechanisms of phase change in locusts. Ann Rev Entomol. 2014;59:225–44.
39. Martín-Blázquez R, Bakkali M. Standardization of multivariate regression models for estimation of the gregariousness level of the main pest Locust. Entomol Exp Appl. 2017;163(1):9–25.
40. Ayali A. The puzzle of Locust density-dependent phase polyphenism. Curr Opin Insect Sci. 2019;35:41–7.
41. Pener MP, Simpson SJ. Locust Phase Polyphenism: an update. Adv Insect Physiol. 2009;36:1–272.
42. Guo X, Ma Z, Du B, Li T, Li W, Xu L, He J, Kang L. Dop1 enhances conspecific olfactory attraction by inhibiting miR-9a maturation in locusts. Nat Commun. 2018;9(1):1193.
43. Birol I, Jackman SD, Nielsen CB, Qian JQ, Varhol R, Stazyk G, Morin RD, Zhao Y, Hirst M, Schein JE. De novo transcriptome assembly with ABySS. Bioinformatics. 2009;25(21):2872–7.
44. Rognes T, Flouri T, Nichols B, Quince C, Mahe F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584.
45. Huang X, Madan A. CAP3: a DNA sequence assembly program. Genome Res. 1999;9(9):868–77.
46. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
48. Anders S. HTSeq: Analysing high-throughput sequencing data with Python. *URL* http://www.huberemblde/users/anders/HTSeq/doc/overview.html 2010.
49. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.
50. Hidalgo-Gutierrez A, Barriocanal-Casado E, Bakkali M, Diaz-Casado ME, Sanchez-Maldonado L, Romero M, Sayed RK, Prehn C, Escames G, Duarte J et al. beta-RA reduces DMQ/CoQ ratio and rescues the encephalopathic phenotype in Coq9 (R239X) mice. *EMBO molecular medicine* 2019, 11(1).
51. Gonzalez-Garcia P, Hidalgo-Gutierrez A, Mascaraque C, Barriocanal-Casado E, Bakkali M, Ziosi M, Abdihankyzy UB, Sanchez-Hernandez S, Escames G, Prokisch H, et al. Coenzyme Q10 modulates sulfide metabolism and links the mitochondrial respiratory chain to pathways associated to one carbon metabolism. Hum Mol Genet. 2020;29(19):3296–311.
52. Yang P, Hou L, Wang X, Kang L. Core transcriptional signatures of phase change in the migratory Locust. Protein Cell. 2019;10(12):883–901.
53. Wang X, Fang X, Yang P, Jiang X, Jiang F, Zhao D, Li B, Cui F, Wei J, Ma C. The Locust genome provides insight into swarm formation and long-distance flight. Nat Commun 2014, 5.
54. Anstey ML, Rogers SM, Ott SR, Burrows M, Simpson SJ. Serotonin mediates behavioral gregarization underlying swarm formation in desert locusts. Science. 2009;323(5914):627–30.
55. Kouhei Toga HB. Meta-analysis of transcriptomes in insects showing density-dependent polyphenism. *bioRxiv* 2022, https://doi.org/10.1101/2022.05.09.490177
56. Foquet B, Castellanos AA, Song H. Comparative analysis of phenotypic plasticity sheds light on the evolution and molecular underpinnings of Locust phase polyphenism. Sci Rep. 2021;11(1):11925.
57. Wang Y, Ghaffari N, Johnson CD, Braga-Neto UM, Wang H, Chen R, Zhou H. Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. BMC Bioinformatics. 2011;12(10):S5.
58. Guo X, Ma Z, Kang L. Serotonin enhances solitariness in phase transition of the migratory Locust. Front Behav Neurosci 2013, 7.
59. Niassy A. Interactions between Schistocerca gregaria (FORSKA.L) and Locusta migratoria migratorioides (REICH & FARMAIRE) in relation to phase polymorphism 1996.
60. Li J, Wei L, Wang Y, Zhang H, Yang P, Zhao Z, Kang L. Cytosolic and mitochondrial ribosomal proteins mediate the Locust phase transition via divergence of translational profiles. Proc Natl Acad Sci USA. 2023;120(5):e2216851120.
61. Bakkali M. A bird's-eye view on the modern genetics workflow and its potential applicability to the Locust problem. C R Biol. 2013;336(8):375–83.

Publisher's note