The complex role of transcription factor GAGA in germline death during *Drosophila* spermatogenesis: transcriptomic and bioinformatic analyses

Svetlana Fedorova [Corresp., 1], **Natalya V Dorogova** [1], **Dmitry A Karagodin** [1], **Dmitry Yu Oshchepkov** [2], **Ilya I Brusentsov** [1], **Natalya V Klimova** [3], **Elina M Baricheva** [1]

[1] Department of Cell Biology, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russian Federation

[2] Department of System Biology, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russian Federation

[3] Department of Molecular Genetics, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russian Federation

Corresponding Author: Svetlana Fedorova
Email address: fsveta@bionet.nsc.ru

The GAGA protein (also known as GAF) is a transcription factor encoded by the *Trl* gene in *D. melanogaster.* GAGA is involved in the regulation of transcription of many genes at all stages of fly development and life. Recently, we investigated the participation of GAGA in spermatogenesis and discovered that *Trl* mutants experience massive degradation of germline cells in the testes. *Trl* underexpression induces autophagic death of spermatocytes, thereby leading to reduced testis size. Here, we aimed to determine the role of the transcription factor GAGA in the regulation of ectopic germline cell death. We investigated how *Trl* underexpression affects gene expression in the testes. We identified 15993 genes in three biological replicates of our RNA-seq analysis and compared transcript levels between hypomorphic Trl^{R85}/Trl^{362} and *Oregon* testes. A total of 2437 differentially expressed genes were found, including 1686 upregulated and 751 downregulated genes. At the transcriptional level, we detected the development of cellular stress in the *Trl*-mutant testes: downregulation of the genes normally expressed in the testes (indicating slowed or abrogated spermatocyte differentiation) and increased expression of metabolic and proteolysis-related genes, including stress response long noncoding RNAs. Nonetheless, in Flybase Gene Ontology lists of genes related to cell death, autophagy, or stress, there was no enrichment with GAGA-binding sites. Furthermore, we did not identify any specific GAGA-dependent cell death pathway that could regulate spermatocyte death. Thus, our data suggest that GAGA deficiency in male germline cells leads to an imbalance of metabolic processes, impaired mitochondrial function, and cell death due to cellular stress.

The complex role of transcription factor GAGA in germline death during *Drosophila* spermatogenesis: transcriptomic and bioinformatic analyses

Svetlana A. Fedorova[1], Natalya V. Dorogova[1], Dmitry A. Karagodin[1], Dmitry Yu. Oshchepkov[2], Ilya I. Brusentsov[1], Natalya V. Klimova[3], Elina M. Baricheva[1]

[1] Department of Cell Biology, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russian Federation
[2] Department of System Biology, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russian Federation
[3] Department of Molecular Genetics, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russian Federation

Corresponding Author:
Svetlana Fedorova[1]
Prospekt Lavrentyeva 10, Novosibirsk, 630090, Russian Federation
Email address: fsveta@bionet.nsc.ru

Abstract

The GAGA protein (also known as GAF) is a transcription factor encoded by the *Trl* gene in *D. melanogaster*. GAGA is involved in the regulation of transcription of many genes at all stages of fly development and life. Recently, we investigated the participation of GAGA in spermatogenesis and discovered that *Trl* mutants experience massive degradation of germline cells in the testes. *Trl* underexpression induces autophagic death of spermatocytes, thereby leading to reduced testis size. Here, we aimed to determine the role of the transcription factor GAGA in the regulation of ectopic germline cell death. We investigated how *Trl* underexpression affects gene expression in the testes. We identified 15993 genes in three biological replicates of our RNA-seq analysis and compared transcript levels between hypomorphic Trl^{R85}/Trl^{362} and *Oregon* testes. A total of 2437 differentially expressed genes were found, including 1686 upregulated and 751 downregulated genes. At the transcriptional level, we detected the development of cellular stress in the *Trl*-mutant testes: downregulation of the genes normally expressed in the testes (indicating slowed or abrogated spermatocyte differentiation) and increased expression of metabolic and proteolysis-related genes, including stress response long noncoding RNAs. Nonetheless, in Flybase Gene Ontology lists of genes related to cell death, autophagy, or stress, there was no enrichment with GAGA-binding sites. Furthermore, we did not identify any specific GAGA-dependent cell death pathway that could regulate spermatocyte death. Thus, our data suggest that GAGA deficiency in male germline cells leads to an imbalance of metabolic processes, impaired mitochondrial function, and cell death due to cellular stress.

Introduction

Cell death occurs in any living organism. Cell death can be accidental, i.e., caused by physical, chemical, toxic, or mechanical damage, or regulated: caused by intracellular problems or exposure to some factors in the extracellular microenvironment. The cell first attempts to cope with such perturbations and to restore cellular homeostasis, but if they are too strong or

prolonged and cannot be reversed, then the process of elimination of the potentially dangerous cell is triggered and it dies. Regulated cell death can also be involved in an organism's development or tissue renewal program; this completely physiological type of death is called programmed cell death (Galuzzi et al., 2018). In 2018, the Nomenclature Committee on Cell Death (NCCD) proposed a classification of 12 major cell death subroutines, based on molecular, morphological, and biochemical characteristics (Galuzzi et al., 2018). The best-known, "classic" types of death are apoptosis, necrosis, and autophagic death. A key feature of apoptotic cell death is the activation of caspases, apoptosome assembly, mitochondrial outer-membrane permeabilization, DNA condensation and specific fragmentation, membrane blebbing, and apoptotic-body formation (Yacobi-Sharon et al., 2013; Galuzzi et al., 2018). Necrosis is characterized by membrane damage, a release of calcium ions into the cytoplasm, bloating of organelles, and acidification of the cytoplasm. Necrosis is often accompanied by inflammation. Autophagic cell death is defined by the accumulation of autophagosomes: double-membrane vesicles that implement cell self-digestion by sequestering cellular materials to lysosomes (Yacobi-Sharon et al., 2013).

In *Drosophila* testes, ~20–30% of spermatogonial cysts undergo spontaneous cell death: germline cell death. It takes place in the apical part of a testis before meiosis and has morphological features of both apoptosis and necrosis but not autophagic cell death (Yacobi-Sharon et al., 2013). Germline cell death is regulated by mitochondrial serine protease HtrA2/Omi but is independent of effector caspases. Bcl-2 family proteins Debcl and Buffy and mitochondrial nuclease EndoG are associated with germline cell death (Yacobi-Sharon et al., 2013). It was also shown that defective spermatocytes can be eliminated by p53-mediated programmed necrosis (Napoletano et al., 2017).

Previously, we have investigated the causes of fertility loss in *Trl*-mutant males and found that spermatocytes undergo mass death (Dorogova et al., 2014; Dorogova et al., 2021). The *Trl* gene encodes transcription factor GAGA (also known as GAF), which participates in the regulation of the transcription of a large group of genes with different cellular functions in *D. melanogaster* (van Steensel et al., 2003; Omelina et al., 2011). Various studies have revealed that GAGA is required for embryogenesis and eye and wing development in *Drosophila* (Farkas et al., 1994; Bhat et al., 1996; Dos-Santos et al., 2008; Omelina et al., 2011; Bayarmagnai et al., 2012; Fedorova et al., 2019). We have researched the role of GAGA in the development and function of the *Drosophila* reproductive system and found that GAGA is required for gonadogenesis (Dorogova et al., 2014; Fedorova et al., 2019). We have demonstrated that *Trl* underexpression causes multiple disorders of oogenesis and spermatogenesis, resulting in a significant loss of fertility (Dorogova et al., 2014; Fedorova et al., 2019). Additionally, we have reported that mass autophagic death of germline cells occurs in *Trl*-mutant testes during spermatogenesis (Dorogova et al., 2014; Dorogova et al., 2021). The question has arisen as to what causes the death of spermatocytes. Because the GAGA protein is a transcription factor, can it regulate the activity of cell death genes? Or does death result from dysregulation of transcription of multiple genes and from metabolic disorders? In this study, we tried to elucidate the mechanisms triggering mass death in the testes of hypomorphic *Trl* mutants. For this purpose, we performed transcriptomic profiling of *Trl*-mutant and normal (control) *D. melanogaster* testes.

Accordingly, we report results of high-throughput RNA sequencing (RNA-seq). We found 4300 differentially expressed genes (DEGs) in hypomorphic *Trl* mutants. Nonetheless, we failed to identify a specific signaling cascade whose activation could lead to the germline cell death in *Trl* mutants. Therefore, we believe that the mass degradation of *Trl* spermatocytes represents regulated cell death caused by the cellular stress that is a consequence of imbalanced intracellular processes, primarily metabolism and mitochondrial activities.

Materials & Methods

Flies

The mutant null-allele *yw;Trl^{R85}/Sb Ser y+* strain was kindly provided by Dr. F. Karch (University of Geneva, Switzerland) and described previously (Farkas et al., 1994). *Trl^{R85}/Trl^{R85}* mutants are not viable in the homozygous state and do not survive to the imago stage; therefore, in our work, we used a combination of the *Trl^{R85}* allele with a strong hypomorphic mutation (*Trl^{362}*) that disrupts the 5′ region of the gene (Ogienko et al., 2006). *Oregon-R-modEncode* (cat. # 25211, Bloomington, USA) served as the wild-type strain. All *Drosophila* stocks were raised at 25°C on a standard cornmeal medium.

RNA-seq

For RNA-seq experiments, three biological replicates were performed. For each RNA-seq assay, total RNA was isolated from testes of 100 1–2-day-old male control flies (Oregon R) and from the same number of *yw; Trl^{R85}/Trl^{362}* testes. Young males (1–2 days old) were anesthetized and placed into Hanks' solution (BioloT). Testes were removed from the flies using dissecting needles, teased apart from seminal vesicles, and immediately transferred into a microcentrifuge tube placed in liquid nitrogen. Fifty testes were put in each tube and stored in liquid nitrogen. Total RNA was isolated using the TRIzol reagent (Invitrogen, USA). After that, genomic DNA was extracted from the organic phase and was tested by PCR for the absence of normal copies of the *Trl* gene in mutant samples. PCR was carried out in a 20 µl reaction mixture consisting of 1× PCR-buffer [16 mM $(NH_4)_2SO_4$; 67 mM Tris-HCl pH 8.9 at 25°C; 0.1% of Tween 20], 1.5 mM $MgCl_2$, 0.2 mM each dNTPs, 1 U of Taq Polymerase (Biosan, Novosibirsk, Russia), and 0.5 µM each primer. The following primers were employed: ex1a (5′-agttatccaacgttggcgag-3′) and NC70-1 (5′-cagacgttagttattagctc-3′). The PCR products were separated by electrophoresis on a 1% agarose gel in 0.5× TBE buffer with ethidium bromide staining.

The quality of the total-RNA samples was evaluated using a Bioanalyzer 2100 device (Agilent). Samples with optimal RNA integrity numbers (RINs) were chosen for further analysis. Additionally, the total RNA was quantitated on an Invitrogen Qubit™ 2.0 fluorometer (Invitrogen). Total RNA (1.2 µg) was treated with DNase (QIAGEN RNase-Free DNase Set) and purified on PureLink™ RNA Micro Kit columns (Invitrogen). RNA-seq libraries were prepared from 0.3 µg of total RNA by means of the TruSeq® Stranded mRNA LT Sample Prep Kit (Illumina) according to the manufacturer's instructions for barcoded libraries. The quality of the obtained libraries was checked on Bioanalyzer 2100 and with the DNA 1000 Kit (Agilent). After normalization, barcoded libraries were pooled and sequenced on a NextSeq550 instrument using NextSeq® 550 High Output v2.5 Kit 75 Cycles (Illumina).

The quality of the obtained raw Fastq files was tested and analyzed in FastQC. To improve the quality of the raw reads, we employed the TrimGalore software, v.0.6.7 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), Cutadapt v.1.15, and FastQC v.0.11.5 via these procedures: removal of a base from either the first or end position if the quality was low (Phred score: 20), trimming of Illumina adapters, and removal of any remaining reads that are <20 bases long. The trimmed reads were aligned with the annotated D. *melanogaster* genome retrieved from FlyBase (http://flybase.org/, dmel_r6.34). The alignment was performed in STAR aligner v.2.7.5a (Dobin et al., 2013) with determination of the number of reads per gene (options -quantMode GeneCounts) (Table 1).

Differential expression analysis was performed in DESeq2 (Anders and Huber, 2010) on the IRIS web server, which is publicly available at http://bmbl.sdstate.edu/IRIS/ (Accessed on 16 January 2020). Genes were considered differentially expressed if their average change of expression was greater than 2-fold and a Benjamini–Hochberg-adjusted P-value (P_{adj}) was less

than 0.05 to ensure statistical significance (Anders and Huber, 2010). Principal component analysis of normalized log-transformed read counts was performed by means of DESeq2 (Anders and Huber, 2010) on the IRIS web server to determine the reproducibility of the analyzed replicates (Fig. S1).

We used the scottyEstimate function of Scotty (Busby et al., 2013) and ssizeRNA 1.3.2 (Bi and Liu, 2016) to measure the statistical power of the differential expression study. We used the Scotty program with the following parameters: fc=2, pCut=0.05, minPercDetected=80, maxReps=10, minReadsPerRep=10,000,000, maxReadsPerRep=100,000,000, minPercUnbiasedGenes=50, pwrBiasCutoff=50, and alignmentRate=98. Parameters for the ssizeRNA program were as follows: nGenes = 15000, pi0 = 0.83, m = 200, mu = mu1,disp = disp1, fc = 2, fdr = 0.05,power = 0.8, maxN = 15. Both approaches showed a calculation power of more than 0.8 (Fig. S2), which means that the 3 biological repeats we used for experiment and control group are sufficiently statistically powerful to detect at least 80% of differentially expressed genes.

Validation of RNA-seq data by quantitative PCR (qPCR)

For qPCR validation, three biological replicates for controls and mutants with two technical replicates per experiment were analyzed. For qPCR, RNA was isolated as described in the transcriptome RNA-seq assay. For each biological replicate, total RNA was isolated from 100 testes of 1–2-day-old Oregon R males or from the same number of *yw; Trl^{R85}/Trl^{362}* testes.

Total RNA was isolated with the TRIzol reagent (TRIzol™ Reagent, Invitrogen, cat. # 15596026) and cleared by means of magnetic beads (RNAClean XP, Beckman, #A63987) according to the manufacturer's instructions. After that, RNA was incubated with DNase I (Thermo Scientific, #MAN0012000) and precipitated with ethanol. cDNA synthesis was performed with Maxima H Minus Reverse Transcriptase (Thermo Scientific, #EP0751) and oligo-dT_{20}. qPCR was carried out on a CFX96 Real-Time System (BioRad). The thermal cycling protocol for the amplification reaction was as follows: 5 min preincubation at 95°C; next, 40 cycles of 10 s at 95°C and 30 s at 60°C; followed by a melting curve program at 65–95°C. Primers used in the qPCRs are listed in Table S1, and the verification data are represented in Fig. S3 and Table S2.

Bioinformatics

The search for GAGA-binding sites was performed with the help of the SITECON software package (Oshchepkov et al., 2004; http://wwwmgs.bionet.nsc.ru/mgs/programs/sitecon/) via a previously developed computational approach (Omelina et al., 2011).

Gene Ontology (GO) enrichment analysis was conducted on the FlyEnrichr server (Chen et al., 2013; Kuleshov et al., 2016) and on the GOrilla server (Eden et al., 2007; Eden et al., 2009). For GOrilla, the P-value threshold was set to 10^{-3}. The sets of up- and downregulated genes ($P_{adj} \leq 0.05$, fold change [FC] >2 or <0.5) were analyzed for GO enrichment in relation to the genes expressed in testes (14 579 genes).

Results

We performed RNA-seq of poly(A)+ RNA from the testes of hypomorphic *Trl^{R85}/Trl^{362}* mutant males. The testes of the wild-type Oregon line served as a control. In three biological replicates, we identified 15993 genes out of the 17612 genes annotated in FlyBase (http://flybase.org/, dmel_r6.34). After testing for differential expression and exclusion of weakly

expressed genes (total read count <10), 14579 genes were chosen for further analysis. Vedelek et al. (2018) published a Drosophila testis transcriptome containing 15015 genes. Vedelek and coauthors cut the testes into three parts (apical, middle, and basal regions) and identified transcription profiles for genes in each testis region (Vedelek et al., 2018). Our RNA-seq results almost completely matched the testis RNA-seq data from ref. (Vedelek et al., 2018), proving the reliability of our RNA-seq results.

Then, we compared transcript levels between Trl^{R85}/Trl^{362} and *Oregon R* testes. RNA-seq analysis revealed 2437 differentially expressed genes (DEGs) ($P_{adj} \leq 0.05$ and $FC \geq 2$) in the testes of *Trl* mutants, where 1686 transcripts were upregulated and 751 were downregulated.

Gene Ontology analysis of RNA-Seq data

We used online tools FlyEnrichr and GOrilla to search for GO terms enriched in our dataset (Eden et al., 2007, 2009; Chen et al., 2013; Kuleshov et al., 2016). For the GO analysis, we subdivided the testis DEGs into upregulated and downregulated. We conducted the analysis in all three GO categories (biological processes, molecular functions, and cellular components) and compared the corresponding datasets to each other. The two tools, FlyEnrichr and GOrilla, yielded similar results (Fig. S4 and S5). As expected, analysis by biological processes showed that the set of upregulated genes was especially enriched with genes related to reproduction and development (Fig. S4 and S5). In addition, in the set of upregulated genes, we detected enrichment with metabolic and proteolysis-related genes. The 43 genes involved in proteolysis (of which four are testis-specific genes) include genes encoding enzymes with serine-type peptidase activity (17 genes), 11 other endopeptidase genes, and 15 protease genes associated mainly with the catabolism of ubiquitinated proteins.

The downregulated DEG set showed the expected enrichment with proliferation and differentiation genes, consistently with the reduced size of the testes in the *Trl* mutants. These include seven dynein complex–related genes (*Dhc36C*, *Dhc98D*, *Dhc62B*, *Dhc64C*, *Dnah3*, *sw*, and *Sdic1*) whose expression diminished 2–6-fold. Dyneins are ATPases and explain the enrichment with the GO term “ATP-dependent microtubule motor activity, minus-end-directed.” Four (*Dhc36C*, *Dhc62B*, *Dnah3*, and *Sdic1*) of these seven genes are testis-specific, and *sw* is involved in sperm individualization.

Unexpectedly, enrichment with the GO term "lysosomes" was detected among both up- and downregulated DEGs. Genes encoding enzymes of catabolic processes, primarily endopeptidases, predominated among the upregulated DEGs associated with the GO term "lysosome." The second place—among the DEGs that decreased their expression in the *Trl* mutant testes—belongs to the GO term “mannose metabolic process,” represented by lysosomal mannosidases LManIII, LManIII, LManIV, LManV, and LManVI (whose expression declined 2.0–38.5-fold) and alpha-Man-Ic (downregulated by 3.2-fold), which is presumably associated with the endoplasmic reticulum and Golgi membrane. There are no testis-specific genes among these genes, but most of them are associated with lysosomes.

Tissue specificity of the DEGs

To identify the tissue specificity of gene expression, we used data from Li and Vedelek, who have determined the specificity of expression by means of the modENCODE tissue expression database (Li et al.,2014; Vedelek et al, 2018). In their work, tissue specificity scores higher than 4 represent testis-enriched genes and lower values represent genes expressed in multiple tissues or ubiquitously; negative values denote under-representation of transcripts in testes (Vedelek et al., 2018). Vedelek and coauthors identified 2602 testis-enriched genes in the *Drosophila* testis transcriptome, including 682 long noncoding RNAs (lncRNAs). The present RNA-Seq data contain genes with different tissue specificity: 3964 genes expressed in several or

many tissues, 2454 testis-enriched genes, and 7330 genes under-represented (i.e., normally not expressed) in testes (Fig. 1). We compared our transcriptome results with Vedelek et al.'s RNA-seq results and noticed a 94.31% overlap between the testis-enriched gene sets (2454 testis-enriched genes including 625 lncRNAs). Among our DEGs, we found 1565 genes expressed in two or more tissues; 681 of them were upregulated and 884 were downregulated. As many as 997 testis-enriched genes, including 303 lncRNAs, featured differential expression in mutant testes: 152 of testis-enriched genes were upregulated (69 twofold or more) and 845 were downregulated (235 genes: twofold or more) (Table S3). It should be noted that among the genes usually under-represented in testes [i.e., genes with the negative value of the tissue specificity index (Li et al., 2014; Vedelek et al., 2018)], the expression pattern was exactly the opposite. In RNA-seq data from *Trl* mutants, 8721 genes were found to be under-represented in the testes, of which 2636 were DEGs; 1836 of them were upregulated, whereas 800 were downregulated (Fig. 1). It is worth pointing out that the expression of late spermatogenesis genes necessary for sperm formation and maturation (Fig. 2A) was the lowest in the *Trl* mutant testes, while the expression of metabolic genes was higher than normal (Fig. 2B).

LncRNAs

Among the 2,438 testis DEGs, 27.7% (675) were lncRNAs, of which 340 lncRNAs (20.17% of all overexpressed transcripts) showed increased expression, and 219 (29.2% of all downregulated transcripts) proved to be downregulated. Because mutations in the *Trl* gene affect germline cells and lead to reduced fertility and spermatogenesis abnormalities (Dorogova et al., 2014), we were primarily interested in expression alterations of testis-specific lncRNAs. In our data, we identified 625 lncRNA genes that, as reported by Li and Vedelek, can be assigned to testis-enriched genes (Li et al. in 2014; Vedelek et al., 2018), 217 of which were differentially expressed. The functions of most of testis-enriched lncRNAs are not known, but Wen et al. screened 128 testis-specific lncRNAs and found 33 lncRNAs that perform critical functions in the regulation of late spermatogenesis and whose knockouts cause spermatogenesis defects and consequent loss of male fertility (Wen et al., 2016). According to our RNA-seq dataset, for three (CR43633, CR43634, and CR43414) of those 33 testis-specific lncRNAs, expression increased by 6-fold or more, whereas for eight (CR43753, CR44412, CR44344, CR43839, CR43862, CR44371, TS15, and CR43416) expression diminished by ≥2-fold. Thus, the knockdown of the above eight RNAs in the testes of *Trl* mutants may be responsible for at least some of the spermatogenesis abnormalities observed.

In addition, in mutant organs, we identified 76 lncRNAs with significantly altered expression that are not expressed in normal testes (Wen et al., 2016; Vedelek et al., 2018). It should be noted that CR34262, CR32010, and CR45346 were among them and featured an increase in expression by 104-, 52-, and 51-fold, respectively; it is known about these three that they are upregulated in response to environmental stressors (Brown et al., 2014; Wen et al., 2016). CR44138—which is induced most strongly by paraquat treatment (oxidative stress) followed by caffeine, Cd, Cu, and Zn—as well as stress-response related CR45346 (Brown et al., 2014; Wen et al., 2016) were also upregulated in *Trl*-mutant testes, by 45-fold and 25-fold, respectively.

Mitochondrial function and structure

Earlier, we showed that *Trl*-mutant spermatocytes acquire abnormal mitochondrion structure and morphology before dying (Dorogova et al., 2014). Here, we examined the connection of our transcriptome dataset with mitochondria (Table S3). In our dataset, we found 778 genes associated with the FlyBase GO term "mitochondrion" (GO:0005739); for 267 of them, expression changed significantly: 123 manifested higher expression, with 35 being

upregulated twofold or more, and 144 were downregulated, with 13 downregulated ≥2-fold (Table S3).

Given that our previously examined mutant mitochondrial phenotype suggested that mitochondrial permeability and function were impaired, we next analyzed the genes encoding respiratory chain components, i.e., the FlyBase GO "respiratory chain complex" term (GO:0098803). Seventy-seven genes out of 79 turned out to be expressed in testes, of which 16 genes were significantly upregulated (five, including Cyt-c1, twofold or more), and eight genes were downregulated. It is worth noting that although for most of the respiratory-chain genes, there was no significant change in expression, some of the subunits underwent expression changes in opposite directions in every component of the respiratory complex and together may have led to an imbalance in the functioning of the respiratory chain as a whole.

Cell death genes

Lack of GAGA in testes leads to death and degradation of germline cells (Dorogova et al., 2021). The question we wanted to answer in this project is whether GAGA is a transcriptional regulator of cell death in *Drosophila* testes. To this end, we compared the group of FlyBase GO lists of death-related genes (corresponding to terms "cell death" GO:0008219, "apoptotic process" GO:0006915, "programmed cell death" GO:0012501, and "autophagic cell death" GO:0048102) with the RNA-seq data we obtained (Table S3, Fig. 3).

Overall, less than half cell death genes underwent changes in expression in the *Trl*-mutant testes. There were twice as many downregulated genes as upregulated genes (Fig. 3). The “Cell death” GO dataset consists of 399 genes, and 362 were found in our RNA-seq data; only 150 of them were differentially expressed in *Trl*-mutant testes, 10 genes were downregulated twofold or more: *ninaE, Dark, CG14118, Trf2, Cdk5alpha, HUWE1, CG9593, Mcm10, Dcr-2,* and *CG30428*; 22 genes proved to be overexpressed ≥2-fold: *rst, ix, pnt, NijA, Drep2, Fhos, wrapper, Eip74EF, DNaseII, Corp, Abd-B+H383A1H4:H14565, CG5860, CG2918, CG31928, Buffy, en, Mdh2, IFT57, ft, lncRNA:Hsromega, tau,* and *Orct*. The most upregulated (FC = 146) was aspartic-type endopeptidase CG31928, which is active in lysosomes. From the 276 genes officially associated with “apoptotic processes” (according to GO FlyBase), our RNA-seq contained 238; only 98 (36.7%) were differentially expressed in *Trl*-mutant testes, and 13 of them were upregulated twofold or more: *Orct*, *wrapper*, *ft, lncRNA:Hsromega*, *DNaseII, Corp, Buffy, IFT57, Abd-B, CG2918, en, Drep2*, and *pnt*; seven genes were downregulated ≥2-fold: *Cdk5alpha, Dark, Mcm10, Dcr-2, CG14118, CG30428,* and *CG9593*. It must be mentioned that *lncRNA:Hsromega* is a stress-inducible lncRNA that takes part in the metabolism regulation as well (according to FlyBase). RNA-seq data from *Trl*-mutant testes contained 28 of the 29 genes from the “autophagic cell death” GO dataset, and 12 genes were differentially expressed; *Dark* and *Trf2* were downregulated and *Eip74EF* and *Mdh2* were upregulated twofold or more.

In our previous work, we demonstrated that autophagosomes and lysosomes are abundant in the cytoplasm of dying spermatocytes. Nevertheless, the death of these cells does not involve apoptosis (Dorogova et al., 2021). Here, we analyzed the genes belonging to the “autophagy” (GO:0006914) and “lysosome” (GO:0005764) FlyBase terms (Table S3). Only 71 out of 200 autophagy genes proved to be differentially expressed in *Trl*-mutant testes, seven of them were upregulated and three downregulated ≥2-fold. The most upregulated gene, *stj* (31-fold change) encodes a voltage-gated calcium channel subunit involved in the regulation of lysosomal fusion with endosomes and autophagosomes (Tian et al., 2015). The most interesting is the Bcl-2 family member Buffy, which participates in stress-induced cell death (Sevrioukov et al., 2007), including a lysosomal alternative germ cell death pathway in *Drosophila* (Yacobi-Sharon et al., 2013). Among 106 lysosomal genes, approximately a half (46 genes) are differentially expressed

in *Trl*-mutant testes, and three of the most overexpressed genes (CG31928: FC = 146, CG4847: FC = 13, CG5860: FC = 3.4) are endopeptidases involved in cell death.

Finally, we analyzed the dataset of the "cellular response to stress" (GO:0033554) term (Table S3). Approximately 30% of genes (240 out of 619, $P_{adj} < 0.05$) underwent changes in their expression; 30 genes were upregulated and 11 were downregulated in *Trl*-mutant testes. The most upregulated gene was *Hsp70Ba* (FC = 96); it encodes a protein involved in the response to heat shock and hypoxia and in the unfolded protein response (Moutaoufik et al., 1993). Besides, upregulated stress-responsive genes included cochaperone Hsc70-3 (FC = 2.5), the hypoxia-induced *CG2918* gene encoding a protein with unfolded-protein–binding activity, and the starvation-upregulated *Sirup* gene, which codes for a critical assembly factor for Complex II in the electron transport chain of mitochondria. It is noteworthy that the official "cellular response to stress" GO dataset do not include lncRNAs, whereas we mentioned above that at least five of the most upregulated (FC = 25 to 104) lncRNAs in our dataset are stress-inducible (Brown et al., 2014; Wen et al., 2016).

GAGA target genes (analysis of GAGA-binding sites)

As mentioned above, our GO term analysis did not reveal enrichment of the DEG set with GO terms corresponding to cell death or stress. We decided to test 5′ regulatory regions of the genes (from FlyBase GO lists of genes) for the presence of GAGA-binding sites. For this purpose, we applied the SITECON software package (Oshchepkov et al., 2004) previously confirmed to recognize GAGA-binding sites (Omelina et al., 2011). Using SITECON software, we analyzed the presence of potential GAGA-binding sites of type 1 (GAGnGAG) and type 2 (GAGnnnGAG) in the -500...+1 region relative to the transcription start site (TSS). We also compared the obtained results with our findings about Flybase GO lists of genes related to development, spermatogenesis, and cellular stress (Fig. 4, Table S4) as well as all DEGs from *Trl*-mutant testes. Developmental genes tend to contain abundant GAGA-binding sites (van Steensel et al., 2003; Omelina et al., 2011); therefore, we included the GO list of developmental genes in our analysis. As expected, the highest density of the GAGA-binding sites proved to be characteristic of genes of developmental processes, with GAGA-binding sites of the second type (GAGnnnGAG) occurring more frequently in these groups of genes (Fig. 4). In addition, the sets of DEGs identified by RNA-seq in our work also have higher density of GAGA-binding sites compared to the genome-wide average. In this work, we analyzed the testis transcriptome in hypomorphic *Trl* mutants, i.e., having a lower amount of the GAGA transcription factor. Hence, DEG sets are expected to be enriched with genes regulated by this transcription factor, and the density of GAGA-binding sites in them should be higher than the genome-wide average. In 5′ (upstream) regions of the genes required for spermatogenesis or male gamete formation, the density of both types of GAGA-binding sites was found to be lower than the genome average.

27.7% of the DEGs in our transcriptome dataset were lncRNAs, we analyzed the lncRNA genes for potential GAGA-binding sites; the results are now added into Table S4 and Figure 4. We found no enrichment with putative GAGA-binding sites in the set of the 675 ncRNAs expressed in hypomorphic *Trl* testes (Figure 4). Out of the 217 testis-specific lncRNAs (Vedelek et al., 2018), only 13 genes had potential GAGA binding sites, out of the 383 testis-nonspecific lncRNA genes, 23 ones had putative GAGA binding sites. It is interesting to note that all of these GAGA-targeted testis-nonspecific lncRNAs increased their expression 3.7-fold or more in *Trl* mutants, i.e. it can be assumed that their expression in the testes is normally suppressed by GAGA factor. Out of the 33 lncRNAs involved in the regulation of spermatogenesis (Wen et al., 2016), only four lncRNAs with potential GAGA-binding sites differ in expression between *Trl*-mutant and control testes: lncRNA:CR43414 is upregulated 6.36-fold in *Trl*-mutant testes,

whereas lncRNA:CR43862, asRNA:TS15, and lncRNA:CR43753 are downregulated by 0.17-, 0.30-, and 0.37-fold, respectively.

Germline cells of Trl^{R85}/Trl^{362} mutant testes die by autophagic death preserving the nuclear envelope intact, thus manifesting impaired mitochondrial morphology and an increased number of autolysosomes and lysosomes (Dorogova et al., 2021). We expected to find enrichment with GAGA-binding sites within 500 bp upstream regions of genes from these categories. Unexpectedly, the density of GAGA-binding sites in the official GO gene lists "autophagic cell death," "autophagy," "lysosome," "mitochondrion," and "respiratory chain complex" was lower than the genome-wide average, and genes related to cell death, programmed cell death, or apoptosis had slightly higher density of GAGA-binding sites (Fig. 4). Of note, type 2 GAGA-binding sites (GAGnnnGAG) are more frequent in cell death–related lists of genes in the GO database as well as among developmental genes in that database. We decided to clarify which genes from the cell death categories commonly contain GAGA-binding sites (Table S4).

The "apoptotic process" GO term corresponds to 267 genes in the Flybase, among which 63 genes (23.60%) contain type 1 GAGA-binding sites and 55 genes (20.60%) contain type 2 GAGA-binding sites (Table S4). Ninety genes out of 267 (33.71%) contain probable GAGA-binding sites in the -500...+1 region. Among these 90 genes, three functional groups of genes stood out: {1} 10 microRNA (miRNA, miR) genes (11.11%) that suppress apoptosis; {2} 11 genes (12. 22%) that regulate cell death in response to various stressors, and {3} seven genes related to DNA fragmentation, including five (*EndoG*, Testis EndoG-Like 3, *Drep1*, *Drep3*, and *CG14118*) encoding endonucleases and two coding for apoptotic executors (nbs and Corp) that recognize DNA breaks and inhibit apoptosis, thereby enabling repair. We identified the largest number of probable GAGA-binding sites in genes *Drep3* (14 sites), *scute* (13), *Egfr* (eight), *CG14118* (12), *brinker* (nine), *Scylla* (seven), and *sickle* (10 sites). Moreover, all the genes encoding endonucleases had more than one GAGA-binding site. Two other features of the probable GAGA target genes from the GO "apoptotic process" list should be highlighted: {1} more than half of them (35 of the 63 genes containing type 1 GAGA-binding sites and 30 of the 55 genes containing type 2 GAGA-binding sites) repress apoptosis; {2} at least one-third of the GAGA target genes in this dataset are developmental genes (30.16% of type 1 and 36.36% of type 2 GAGA target genes), for example, *abdA* (three GAGA-binding sites), *abdB* (four sites), *en* (two sites), and *sc* (14 sites). Most likely, it is these developmental genes that contribute to the observed higher density of GAGA-binding sites in the -500…+1 region of death-related lists of genes in the Flybase GO database.

In the Trl^{R85}/Trl^{362} mutant testes, only 10 probable GAGA target genes from the Flybase GO list for the "apoptotic process" term were differentially expressed: endonuclease *CG14118* was 3-fold downregulated, whereas nine genes (*Orct, ft, lncRNA:Hsromega, Corp, fkh, Buffy, Abd-B*, *en*, and *pnt*) were upregulated. It is important to point out that it is impossible to obtain information about miRNA expression levels from our RNA-seq data; therefore, we cannot say anything about the expression levels of the 10 genes of apoptosis-suppressing miRNAs that contain probable GAGA-binding sites. MiRNAs bantam, miR-14, and miR-278 and the miR-2 family, which includes miR-2, -6, -11, -13, and -308, are known to independently suppress apoptosis at the post-transcriptional level (Jovanovic and Hengartner, 2006). In addition, among the target mRNAs of the listed miRNAs, there are transcripts encoding regulators of other types of cell death, such as Fhos, a component of programmed autophagic death, and NijA, a contributor to necrosis. Downregulation of these miRNAs can activate translation of hundreds of their target mRNAs and launch cell death.

Similarly, of the 333 genes in the "programmed cell death" GO list, less than half of the genes (41.14%, i.e., 137 genes, including 11 miRNA genes) contain potential GAGA-binding sites. Seventy-six genes (22.82%) contain GAGA type 1 sites, and 69 (20.72%) genes contain

type 2 sites. The highest number of GAGA-binding sites is found in the -500…+1 region of genes *Drep3* (14 sites), *scute* (13), *CG14118* (12), and *sickle* (10 sites). Furthermore, as in the case of the "apoptotic process" term, the set of probable GAGA target genes contains developmental genes.

The GO list "autophagy genes" includes 200 genes, of which only 59 (29.5%) carry potential GAGA-binding sites. Forty-two genes contain GAGA-binding sites of type 1, and 32 of type 2. The largest number of the sites was detected in genes *AMPKalpha* (eight sites), *daw* (seven), and *hid* (six).

Thus, our results on the density distribution of GAGA-binding sites in 500 bp upstream regions of genes indicate that it is unlikely that GAGA as a transcription factor directly controls regulated germline cell death in *Trl*-mutant testes. Although we observed slightly higher density of GAGA-binding sites in the Flybase GO lists of genes corresponding to "programmed cell death," "apoptotic process," and "cell death" (Fig. 4), this result may be explained by the presence of developmental genes in these lists.

Discussion

The GAGA protein is a global regulator of the expression of thousands of genes and is involved in the modulation of transcription at several levels: chromatin remodeling, transcription, and Pol II pausing (van Steensel et al., 2003; Tsai et al., 2016). In our earlier work, we showed that a GAGA deficit in *Drosophila* testes leads to autophagic death of germline cells and testis diminution (Dorogova et al., 2014; Dorogova et al., 2021). In the present report, we addressed the participation of GAGA in transcriptional regulation of male germline cell death. We conducted an RNA-seq analysis of control and mutant testes and compared the corresponding gene sets to each other. Our analysis revealed some characteristics of the regulation of cell death in the *Drosophila* testes and, in particular, answered the question what role GAGA plays in the control of cell death in *Drosophila* testes.

DEGs in *Trl*-mutant testes

An unexpected funding in the analysis of the transcriptomes is the predominance of upregulated genes over downregulated ones among the DEGs. As many as 69.18% of the DEGs proved to be overexpressed in the *Trl*-mutant testes whereas only 30.82% were underexpressed. It is known that transcription factor GAGA is a positive regulator of global gene expression (Tsai et al., 2016); consequently, we expected that its depletion in the testes should raise the expression of most genes. Indeed, in the larval imaginal discs carrying a similar combination of the *Trl*-null allele with a hypomorphic *Trl* mutation, 82% of genes turned out to be underexpressed, confirming a positive role of GAGA in global gene regulation (Tsai et al., 2016). It is possible that the expression skew toward gene activation that we registered in *Trl*-mutant testes is not due to the GAGA activity as a transcription factor but rather to other aspects of its function. For example, GAGA is involved in chromatin remodeling and can modulate gene expression by changing nucleosome density in a promoter region (Judd et al., 2021; Fuda et al., 2015). This notion is supported by our finding that the set of upregulated DEGs, according to FlyEnrichr, is enriched with the target genes of transcription factors Pc and Su(Hw) (Fig. S6). In *Trl*-mutant testes, more than half of the target genes of Trl, Pc, and Su(Hw) show upregulation (Fig. S6). Pc and Su(Hw) modulate gene expression at the chromatin level. Direct protein–protein interactions of GAGA with Pc as well as with some other components of the Pc complex are reported to enhance the activity of the Pc complex (Poux et al., 2001). No direct binding of GAGA with Su(Hw) has been reported but they can interact via Mod(mdg4) (Melnikova et al., 2004; Ghosh et al., 2001). For instance, the interaction between GAGA and Mod(mdg4) is a possible mechanism

governing gypsy insulator activity (Melnikova et al., 2004). Normally, the Pc complex and the Su(Hw) insulator repress large groups of genes at the beginning of spermatogenesis, thereby preventing premature differentiation of germline cells into spermatozoa (Zhang et al., 2017; Feng et al, 2017; Glenn and Geyer, 2019). GAGA may directly or indirectly take part in these processes; therefore, its depletion drives alterations in the composition of repressor and/or insulator complexes and induction of some repressed genes.

Another reason for the prevalence of gene activation in *Trl*-mutant testes may be the cancellation of Pol II polymerase pausing. Paused RNA polymerase II (RNA-Pol) is located ~30–50 bp downstream of the TSS of genes associated with developmental control, cell proliferation, and intercellular signaling (Tsai et al., 2016). GAGA is enriched on promoters with paused Pol II (Tsai et al., 2016; Fuda et al., 2015). The groups of genes that must be rapidly and synchronously expressed in response to developmental or stress signals, such as heat shock protein genes, are normally inactive (Vihervaara et al., 2018). For example, the *Hsp70* gene encoding a chaperone is highly repressed, but 1 min of heating is sufficient to activate it (Tsai et al., 2016; Duarte et al., 2016). Some researchers demonstrated a decline of the amount of paused RNA-Pol as well as productive *Hsp70* transcription in the absence of GAGA (Tsai et al., 2016). Thus, the increased expression of approximately $^{2}/_{3}$ of genes in the *Trl*-mutant testes may be due to the downregulation/inactivation of GAGA as a chromatin remodeler or RNA polymerase pausing enforcer.

Testis-specific genes

Normally, the GAGA protein is localized to the apical end of the testes in germline cell nuclei at early premeiotic stages (Dorogova et al., 2014). This period of spermatogenesis is characterized by a high level of gene transcription and synthetic activity. The volume of spermatocytes increases 25-fold, which requires considerable consumption of energy and resources (Fuller M.T., 1993). Germline cells at premeiotic stages have transcriptome composition that is dominated by genes responsible for mitotic division and preparation for differentiation. In this regard, our data match the results of Vedelek et al. who divided the testis into three parts—apical, middle, and basal—and analyzed the transcriptomes of each part (Vedelek et al., 2016). According to their data, the apical part of the testis is enriched with genes that are maximally expressed in other tissues, not testes. Testis-specific genes were detected mainly in the basal part and were found to play a part in sperm differentiation and morphogenesis (Vedelek et al., 2016).

In GAGA-deficient testes, germline cells are present only in the early stages of spermatogenesis; starting from the stage of spermatocytes, they begin to degrade and get eliminated (Dorogova et al., 2021). This means that the testes of *Trl* mutants actually consist of the apical part as a consequence of the early death of the germline cells. As we expected in this case, testis-specific genes are under-represented in the RNA-seq data on the *Trl* mutant (Fig. 1), in agreement with the data from ref. (Vedelek et al., 2016). Among 2602 testis-enriched genes from Vedelek et al.'s dataset, 148 genes were not detectable in our experiment, whereas the other 1855 showed underexpression. Thus, the results of RNA-seq analysis of *Trl*-mutant testes confirmed the observed mutant phenotype, i.e., the death of spermatocytes is induced before their differentiation.

A characteristic feature of our RNA-seq dataset is its enrichment with testis-under-represented genes (Fig. 1). FlyEnrichr GO analysis indicated that in the mutant testes, the upregulated genes were mainly metabolic genes (Fig. 2). We believe that this evidence may reflect increased cellular stress: when GAGA is in short supply, an imbalance develops in the expression levels of genes from different processes, for example, in the mitochondrial respiratory chain. The unbalanced amounts of products of such genes, on the one hand, drive additional

induction of genes coding for the missing components. On the other hand, an excessive amount of proteins may not be recycled in time by the proteolytic system and can create insoluble protein aggregates that trigger the unfolded protein response, which alters the expression of many genes involved in endoplasmic-reticulum quality control. At the same time, mitochondrial dysfunction leads to a lack of energy in the cell and again upregulates metabolic genes. By contrast, in the absence of the GAGA protein, transcriptional regulation is disturbed, promoting cellular stress, activation of autophagy (in an attempt to utilize some proteins and remove destroyed organelles), and in the end, cell death. This notion is supported by the observed dramatic overexpression of stress-inducible genes and lncRNA genes: e.g., the *Hsp70Ba* gene which encodes a chaperone participating in the response to heat shock and hypoxia and stress response–related lncRNA genes *CR34262*, *CR32010* and *CR45346*. Taken together, these data suggest that the observed cell death during spermatogenesis is not related to cell differentiation or seminal functions but rather is a consequence of metabolic aberrations.

LncRNAs and miRNAs

Approximately 1/3 of our DEG set consists of lncRNAs, a class of noncoding RNAs longer than 200 nucleotides. More and more evidence has been accumulating about the functions of lncRNAs in numerous biological processes and in diseases (Choudhary et al., 2021; Xu et al. 2017; Li et al., 2019; Deniz et al., 2017). LncRNAs are predominantly localized to the nucleus and implement gene expression regulation at epigenetic, transcriptional, and post-transcriptional levels (reviewed by Li et al., 2019). Underexpression of some lncRNAs may result in abnormal embryogenesis, loss or decline of *Drosophila* fertility (Wen et al., 2016; Li et al., 2019). Wen et al. revealed 33 testis-specific lncRNAs involved in the regulation of spermatogenesis (Wen et al., 2016), four of them with potential GAGA-binding sites altered their expression in *Trl*-mutant testes. Their knockdown certainly contributed to the spermatogenesis abnormalities we observed, but it should be noted that these lncRNAs act at later (postmeiotic) stages of spermatogenesis as compared to the death of *Trl*-mutant spermatocytes. The knockdowns of CR43414 and TS15 via RNA interference impaired the polarization of spermatids and caused a lag or poor alignment of individualization complexes; the CR43753 knockdown resulted in an abnormal early phase of spermatid elongation; and the CR43862 knockdown yielded scattered and curled sperm nuclei in late spermatogenesis. By contrast, in the case of hypomorphic *Trl* testes, death occurred at the spermatocyte stage, and most cells did not reach meiosis

Recent studies have revealed important functions of lncRNAs in the modulation of autophagy, the regulation of metabolism and the stress resistance in Drosophila (Li et al., 2019; Lakhotia et al., 2012). Moreover, any dysregulation of lncRNA expression is enough to reduce stress tolerance. For example, lncRNA hsrω overexpression as well as nullisomy or its RNA interference are lethal for most *Drosophila* embryos and first- or third-instar larvae under heat stress (Lakhotia et al., 2012). LncRNA hsrω contributes to omega speckle formation (which is a spatial repository of key regulatory factors bound to their pre-stress nuclear targets in cells recovering from stress) and to regulation of the protein metabolic process (Lakhotia et al., 2012; Lo Piccolo et al., 2019). In *Trl*-mutant testes, we detected 3.35-fold upregulation of lncRNA hsrω (Table S3) and huge overexpression of other stress response lncRNAs, e.g., CR34262, CR32010, and CR45346. We suppose they may contribute to the aforementioned upregulation of metabolic genes. On the other hand, expression dysregulation of some lncRNAs may induce the autophagic cell death in mutants. It is known that the highest expression level of lncRNA lncov1 coincides with the autophagic cell death in the larval ovary of the worker bee *Apis mellifera* (Choudhary et al., 2021). In humans, overexpression of BRAF-activated lncRNA (BANCR) raises the LC3-II/LC3-I ratio, a marker of autophagy (Xu et al., 2017).

The regulation of cell death via the lncRNA–miRNA axis deserves special attention and requires further research. MiRNAs are small noncoding RNAs 18–24 nucleotides long that control gene expression post-transcriptionally through mRNA degradation or translation inhibition (Leaman et al., 2005). Many miRNAs are implicated in the regulation of cell death: the largest miRNA family (miR-2, -6, -11, -13, and -308), miR-14, and bantam are inhibitors of apoptosis in *Drosophila* (Leaman et al., 2005; Xu et al., 2003); the mammalian miRNA-30a family attenuates Beclin-mediated autophagy stimulation (Xu et al., 2017). LncRNAs and miRNAs interact with each other at different levels: direct transcriptional regulation (some lncRNA sequences contain miRNA recognition elements [MREs]); lncRNA stability can be weakened by miRNAs; lncRNAs serve as miRNA decoys or sponges and can compete for target mRNAs; finally, lncRNAs can be a source of miRNAs (Xu et al., 2017). Thus, in *Trl*-mutant testes, the dysregulation of lncRNA expression may disrupt the balance of lncRNA–miRNA interactions contributing to germline cell death.

Cell death genes

The question that we addressed in this paper is whether GAGA transcriptionally regulates cell death in *Drosophila* testes. GO analysis of the testis DEGs did not reveal any enrichment with gene groups associated with cell death. Moreover, the majority of cell death–related genes did not manifest changes in expression in the mutant testes (Fig. 3), and downregulated genes were prevalent among the differentially expressed cell death–related ones.

Because germline cell death in the mutant testes is autophagic, we previously suggested that *Trl* mutations negatively affect cell metabolism by preventing the necessary magnitude of macromolecule synthesis and cell growth (Dorogova et al., 2021). The lack of energy and macromolecules as a rule activates the TOR signaling cascade, which regulates autophagy. GAGA protein deficiency can affect the expression of either genes encoding components of the TOR pathway and/or factors regulating autophagy, which can also cause ectopic death (Levine and Klionsky, 2004; Das et al., 2012). Overexpression of *Atg2*, *Atg4*, *Atg5*, *Atg7*, *Atg8*, *Atg9*, *Atg16*, *Atg17*, and *Atg18* and inactivation of TOR kinase are necessary for autophagy activation by this signaling cascade (Levine and Klionsky, 2004; Das et al., 2012). We failed to detect transcriptional activation of TOR pathway genes in the present RNA-seq study, although some individual genes from this pathway proved to be down- and upregulated. For example, no *Atg* gene underwent a more than 2-fold expression change, but genes *stj* and *lft* involved in autophagosome maturation were upregulated 31.34- and 2.25-fold, respectively (Table S3). It is possible that TOR pathway activation in the mutants proceeds in a noncanonical manner and requires further study.

Next, we looked at whether potential GAGA-binding sites were present in the regulatory regions of cell death–related and stress-related genes (Fig. 4). We chose the region of 500 bp upstream of the TSS because it was shown previously that peaks of GAGA binding are situated within this region in the majority of GAGA target genes (Tsai et al., 2016). Moreover, stress-activated genes whose induction is dependent on GAGA have a strong tendency to contain GAGA-binding sites immediately upstream of the TSS, between positions −100 and −50 (Duarte et al., 2016). Our analysis confirmed the reports that the list of developmental genes is enriched with GAGA-binding sites (van Steensel et al., 2003; Omelina et al., 2011). We detected no enrichment with GAGA-binding sites in the gene lists corresponding to terms "autophagic cell death," "autophagy," "lysosome," "mitochondrion," and "respiratory chain complex" in the Flybase GO database, but we did detect slight enrichment with GAGA-binding sites in GO gene lists corresponding to "programmed cell death," "apoptotic process," and "cell death." On detailed examination, it appeared that this increase in the density of GAGA-binding sites is attributable to the developmental genes present in the GO gene lists (Table S3). It should also be noted that

approximately one-third of the genes carrying probable GAGA-binding sites from the "apoptotic process" gene list either have nonapoptotic functions in development or are implicated in programmed cell death during development or morphogenesis. For example, the main role of *scute* is the determination of sex and the development of the nervous system; *engrailed* is essential for posterior compartment identity and for compartment boundary formation and maintenance; and *Ecdysone receptor (EcR*) launches both molting and metamorphosis (Flybase.org). Some of the GAGA targets belong to the group of genes regulating programmed cell death during development. For instance, the *Abd B* gene is involved in promotion of the apoptotic process associated with morphogenesis; *argos* is necessary for facet differentiation and programmed cell death during eye morphogenesis; *Scylla* acts as cell death activator during head development (Flybase.org); and *hid* activation causes the programmed apoptotic cell death during facet formation (Hsu et al., 2002). Additionally, some genes have been added into the list of cell death–related genes according to a prediction of their function, which is not always accurate. For example, the *peanut* gene is classified as apoptotic because its mammalian homolog, *ARTS*, has a truncated isoform that is localized to mitochondria and actually contributes to the modulation of apoptosis (Mandel-Gutfreund and Larisch, 2011). Our studies suggest that in *Drosophila*, the *peanut* gene does not have a truncated isoform, the PNUT protein is localized subcortically, and its overexpression or deficiency has no effect on cell death (Akhmetova et al., 2017; Akhmetova et al., 2015). Thus, we believe that the modest enrichment of official GO lists of cell death–related genes with GAGA-binding sites is a false positive result because it can be explained by the presence of developmental genes and genes with unproven contributions to the regulation of death in the gene lists. Only the "programmed cell death" list is truly enriched with GAGA-binding sites, not surprisingly, because it is known that the transcription factor GAGA helps to control many developmental processes in *Drosophila* (Bhat et al., 1996; Dos-Santos et al., 2008; Omelina et al., 2011; Bayarmagnai et al., 2012).

Taken together, these data indicate that it is unlikely that GAGA regulates cell death of germline cells in *Drosophila* testes at the transcriptional level. We think that the germline cell death observed in *Trl* mutants may be a consequence of imbalanced intracellular processes (primarily metabolism and mitochondrial functioning) that lead to cellular stress.

Conclusions

Previously, we have shown that in *Trl* mutants, mass death of germline cells occurs during spermatogenesis. At the earliest stages, pathological changes are detectable in the mitochondrial apparatus of cells: hypertrophy and swelling of mitochondria, matrix decondensation, and crista degradation. *Trl*-mutant spermatocytes are subject to excessive autophagy and lysis at the premeiotic growth stage (Dorogova et al., 2014; Dorogova et al., 2021). Given that the *Trl* gene encodes the transcription factor GAGA, which governs the expression of many genes, we decided in the present study to test whether it regulates cell death at the transcriptional level. We performed RNA-seq analysis of testes carrying a combination of null allele Trl^{R85} and hypomorphic mutation Trl^{362}. Examination of the results at the transcriptional level confirmed disturbances of mitochondrial structure and function and developing cellular stress in spermatocytes. Characteristic features of the RNA-seq dataset of Trl-mutant testes turned out to be 1) diminished expression of the testis-enriched genes that are essential for sperm morphogenesis at later stages, indicating slowed or abrogated spermatocyte differentiation; 2) greater expression of ubiquitous or multitissue genes, among which metabolic genes dominated, indicating cellular stress; 3) opposite gene expression changes in many biological processes, such as the respiratory chain and programmed cell death. Nonetheless, we did not identify any specific signaling cascade whose activation could lead to the death of

germline cells deficient in GAGA. Furthermore, we failed to detect enrichment with GAGA-binding sites within -500...+1 regions of genes corresponding to autophagy, cell death, or stress-related GO terms.

Altogether, these findings suggest that the mass degradation of *Trl*-mutant spermatocytes represents regulated cell death caused by the cellular stress that is probably a consequence of imbalanced intracellular processes, primarily metabolism and mitochondrial functioning.

Acknowledgements

This study was supported by The RFBR Project № 20-04-00496-a and The Ministry of Science and Higher Education of the Russian Federation Project № FWNR-2022-0015. We thank ICIG SB RAS Genomics Core Facility and Gennady Vasilyev for RNA sequencing. The English language was corrected and certified by shevchuk-editing.com.

References

Akhmetova KA, Dorogova CN, Chesnokov IN, Fedorova SA. 2015. Analysis of Phenotypic Manifestation of *peanut* Gene Expression Suppression by RNAi in *Drosophila* Oogenesis. *Genetika* 51(9):991-9.

Akhmetova KA, Dorogova NV, Bolobolova EU, Chesnokov IN, Fedorova SA. 2017. The Role of Pnut and its Functional Domains in *Drosophila* Spermatogenesis. *Russian Journal of Genetics: Applied Research* 7:29-35. doi: 10.1134/s2079059717010026.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* 11, R106 doi:10.1186/gb-2010-11-10-r106.

Arsham AM, Neufeld TP. 2009. A Genetic Screen in Drosophila Reveals Novel Cytoprotective Functions of the Autophagy-Lysosome Pathway. *PLoS ONE* 4(6): e6068. doi:10.1371/journal.pone.0006068

Bayarmagnai B, Nicolay BN, Islam AB, Lopez-Bigas N, Frolov MV. 2012. Drosophila GAGA factor is required for full activation of the dE2f1-Yki/Sd transcriptional program. *Cell Cycle* 11(22): 4191-4202. doi: 10.4161/cc.22486.

Bhat KM, Farkas G, Karch F, Gyurkovics H, Gausz J, Schedl P. 1996. The GAGA factor is required in the early *Drosophila* embryo not only for transcriptional regulation but also for nuclear division. *Development* 122;(4):1113-1124.

Bi R, Liu P. 2016. Sample size calculation while controlling false discovery rate for differential expression analysis with RNA-sequencing experiments. *BMC Bioinformatics*. 17:146. doi: 10.1186/s12859-016-0994-9

Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW,Wen J, Park S, Suzuki AM. 2014. Diversity and dynamics of the *Drosophila* transcriptome. *Nature* 512: 393–399 doi: 10.1038/nature12962.

Busby MA, Stewart C, Miller CA, Grzeda KR, Marth GT.2013. Scotty: a web tool for designing RNA-Seq experiments to measure differential gene expression. *Bioinformatics*. 29(5):656-7. doi: 10.1093/bioinformatics/btt015

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 128(14) doi: 10.1186/1471-2105-14-128.

Choudhary C, Sharma S, Meghwanshi KK, Patel S, Mehta P, Shukla N, Do DN, Rajpurohit S, Suravajhala P, Shukla JN. 2021. Long Non-Coding RNAs in Insects. *Animals* 11:1118 doi: 10.3390/ani11041118.

Das G, Shravage BV, Baehrecke EH. 2012. Regulation and function of autophagy during cell survival and cell death. *Cold Spring Harbor Perspectives in Biology* 4(6):a008813 doi: 10.1101/cshperspect.a00 8813.
Deniz E, Erman B. 2017. Long noncoding RNA (lincRNA), a new paradigm in gene expression control. *Functional and Integrative Genomics* 17:135–43 doi: 10.1007/s10142-016-0524-x.
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15-21. doi: 10.1093/ bioinformatics/bts635.
Dorogova NV, Fedorova EV, Bolobolova EU, Ogienko AA, Baricheva EM. 2014. GAGA protein is essential for male germ cell development in *Drosophila*. *Genesis* 52(8):738-751 doi: 10.1002/ dvg.22789.
Dorogova NV, Zubkova AE, Fedorova EV, Bolobolova EU, Baricheva EM. 2021. Lack of GAGA protein in *Trl* mutants causes massive cell death in *Drosophila* spermatogenesis and oogenesis. *Vavilov Journal of Genetics and Breeding* doi:10.18699/VJ21.52-o (in Russian).
Dos-Santos N, Rubin T, Chalvet F, Gandille P, Cremazy F, Leroy J, Boissonneau E, Theodore L. 2008. *Drosophila* retinal pigment cell death is regulated in a position-dependent manner by a cell memory gene. *International Journal of Developmental Biology* 52(1):21-31 doi: 10.1387/ijdb.072406nd.
Duarte FM, Fuda NJ, Mahat DB, Core LJ, Guertin MJ, Lis JT. 2016. Transcription factors GAF and HSF act at distinct regulatory steps to modulate stress-induced gene activation. *Genes and Development* 30(15):1731-46. doi: 10.1101/gad.284430.116.
Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48 doi: 10.1186/1471-2105-10-48.
Eden E, Lipson D, Yogev S, Yakhini Z. 2007. Discovering motifs in ranked lists of DNA sequences. *PLoS Computational Biology* 3:0508–22 doi: 10.1371/journal.pcbi.0030039.
Farkas G, Gausz J, Galloni M, Reuter G, Gyurkovics H, Karch F. 1994. The *Trithorax-like* gene encodes the *Drosophila* GAGA factor. *Nature* 371:806-808 doi: 10.1038/371806a0.
Fedorova EV, Dorogova NV, Bolobolova EU, Fedorova SA, Karagodin DA, Ogienko AA, Khruscheva AS, Baricheva EM. 2019. GAGA protein is required for multiple aspects of *Drosophila* oogenesis and female fertility. *Genesis* 57(2):e23269 doi:10.1002/dvg.23269.
Feng L, Shi Z, Chen X. 2017. Enhancer of polycomb coordinates multiple signaling pathways to promote both cyst and germline stem cell differentiation in the *Drosophila* adult testis. *PLoS Genetics* 13(2):e1006571. doi: 10.1371/journal.pgen.1006571.
Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, Siepel A, Lis JT. 2015. GAGA factor maintains nucleosome-free regions and has a role in RNA polymerase II recruitment to promoters. *PLoS Genetics* 11(3):e1005108 doi: 10.1371/journal.pgen.1005108.
Fuller MT. 1993. Spermatogenesis In: The development of *Drosophila melanogaster*. *NY Cold Spring Harbor. Press,* 71-147.
Galluzzi L, Vitale I, Aaronson SA. et al. 2018. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death and Differentiation* 25(3):486-541 doi: 10.1038/s41418-017-0012-4.
Ghosh D, Gerasimova TI., Corces VG. 2001. Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. *EMBO Journal* 20(10): 2518-2527 doi:10.1093/emboj/20.10.2518.
Glenn SE, Geyer PK. 2019. Investigation of the Developmental Requirements of *Drosophila* HP1 and Insulator Protein Partner, HIPP1. *Genes Genomics Genetics (Bethesda)* 9(2):345-357 doi: 10.1534/g3.118.200705.
Hsu CD, Adams SM, O'Tousa JE. 2002. Rpr- and hid-driven cell death in

Drosophila photoreceptors. *Vision Research* 42(4):507-16 doi: 10.1016/s0042-6989(01)00231-0.

Judd J, Duarte FM, Lis JT. 2021. Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription. *Genes and Development* 35(1-2):147-156 doi: 10.1101/gad.341768.120.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research* gkw377 doi: 10.1093/nar/gkw377.

Lakhotia SC, Mallik M, Singh AK, Ray M. 2012. The large noncoding hsrω-n transcripts are essential for thermotolerance and remobilization of hnRNPs, HP1 and RNA polymerase II during recovery from heat shock in *Drosophila*. *Chromosoma* 121:49–70 doi: 10.1007/s00412-011-0341-x.

Leaman D, Chen PY, Fak J, Yalcin A, Pearce M, Unnerstall U, Marks DS, Sander C, Tuschl T, Gaul U. 2005. Antisense-mediated depletion reveals essential and specific functions of microRNAs in *Drosophila* development. *Cell* 121(7):1097-108 doi: 10.1016/j.cell.2005.04.016.

Levine B, Klionsky DJ. 2004. Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Developmental Cell* 6:463-477 doi: 10.1016/s1534-5807(04)00099-1.

Li JJ, Huang H, Bickel PJ, Brenner SE. 2014. Comparison of *D. melanogaster* and *C. elegans* developmental stages, tissues, and cells by modENCODE RNA-seq data. *Genome Research* 24(7):1086-1101 doi:10.1101/gr.170100.113

Li K, Tian Y, Yuan Y, et al. 2019. Insights into the Functions of LncRNAs in *Drosophila*. *International Journal of Molecular Science* 20(18):4646 doi:10.3390/ijms20184646.

Lo Piccolo L, Mochizuki H, Nagai Y. 2019. The lncRNA hsrω regulates arginine dimethylation of human FUS to cause its proteasomal degradation in *Drosophila*. *Journal of Cell Science* 132(20): jcs236836 doi: 10.1242/jcs.236836.

Mandel-Gutfreund Y, Kosti I, Larisch S. 2011. ARTS, the unusual septin: structural and functional aspects. *Biological Chemistry* 392(8-9):783-90 doi: 10.1515/BC.2011.089.

Maruzs T, Simon-Vecsei Z, Kiss V, Csizmadia T and Juhász G. 2019. On the Fly: Recent Progress on Autophagy and Aging in *Drosophila*. *Frontiers in Cell and Developmental Biology* 7:140 doi: 10.3389/fcell.2019.00140.

Melnikova L, Juge F, Gruzdeva N, Mazur A, Cavalli G, Georgiev P. 2004. Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14806-14811 doi: 10.1073/pnas.0403959101.

Moutaoufik MT, Tanguay RM. 2021. Analysis of insect nuclear small heat shock proteins and interacting proteins. *Cell Stress and Chaperones* 26(1): 265-274 doi: 10.1007/s12192-020-01156-3.

Napoletano F, Gibert B, Yacobi-Sharon K, et al. 2017. p53-dependent programmed necrosis controls germ cell homeostasis during spermatogenesis. *PLoS Genetics* 13(9): e1007024 doi: 10.1371/journal.pgen.1007024.

Ogienko AA, Karagodin DA, Fedorova SA, Fedorova EV, Lashina VV, Baricheva EM. 2006. Effect of hypomorphic mutation in *Trithorax-like* gene on *Drosophila melanogaster* oogenesis. *Ontogenez* 37(3):211-220.

Omelina ES, Baricheva EM, Oshchepkov DY, Merkulova TI. 2011. Analysis and recognition of the GAGA transcription factor binding sites in *Drosophila* genes. *Computational Biology and Chemistry* 35(6):363-370 doi: 10.1016/j.compbiolchem. 2011.10.008.

Oshchepkov DY, Vityaev EE, Grigorovich DA, Ignatieva EV, Khlebodarova TM. 2004. SITECON: a tool for detecting conservative conformational and physicochemical properties in transcription factor binding site alignments and for site recognition. *Nucleic Acids Research* 32(Web Server issue):W208-12 doi: 10.1093/nar/gkh474.

Parsell DA., Lindquist S. 1993. The function of heat-shock proteins in stress tolerance: Degradation and reactivation of damaged proteins. *Annual Review of Genetics* 27: 437-496 doi: 10.1146/annurev.ge.27.120193.002253.

Poux S, Melfi R, Pirrotta V. 2001. Establishment of Polycomb silencing requires a transient interaction between PC and ESC. *Genes and Development* 15(19): 2509-2514 doi: 10.1101/gad.208901.

Sevrioukov EA, Burr J, Huang EW, Assi HH, Monserrate JP, Purves DC, Wu JN, Song EJ, Brachmann CB. 2007. *Drosophila* Bcl-2 proteins participate in stress-induced apoptosis, but are not required for normal development. *Genesis*. 45(4): 184-193 doi: 10.1002/dvg.20279.

Tian X, Gala U, Zhang Y, Shang W, Nagarkar-Jaiswal S, di Ronza A, Jaiswal M, Yamamoto S, Sandoval H, Duraine L, Sardiello M, Sillitoe RV, Venkatachalam K Fan H, Bellen HJ, Tong C. 2015. A Voltage-Gated Calcium Channel Regulates Lysosomal Fusion with Endosomes and Autophagosomes and Is Required for Neuronal Homeostasis. *PLoS Biology* 13(3): e1002103 doi: 10.1371/journal.pbio.1002103.

Tsai SY, Chang YL, Swamy KB, Chiang RL, Huang DH. 2016. GAGA factor, a positive regulator of global gene expression, modulates transcriptional pausing and organization of upstream nucleosomes. *Epigenetics and Chromatin* 9:32 doi: 10.1186/s13072-016-0082-4.

van Steensel B, Delrow J, Bussemaker HJ. 2003. Genomewide analysis of *Drosophila* GAGA factor target genes reveals context-dependent DNA binding. Proceedings of the National Academy of Sciences of the United States of America 100(5):2580-2580 doi: 10.1073/pnas.0438000100.

Vedelek V, Bodai L, Grézal G. et al. 2018. Analysis of *Drosophila melanogaster* testis transcriptome. *BMC Genomics* 19:697 doi: 10.1186/s12864-018-5085-z.

Vihervaara A, Duarte FM, Lis JT. 2018. Molecular mechanisms driving transcriptional stress responses. *Nature Reviews Genetics* 19(6):385-397 doi: 10.1038/s41576-018-0001-6.

Wen K, Yang L, Xiong T, Di C, Ma D, Wu M, et al. 2016. Critical roles of long noncoding RNAs in *Drosophila* spermatogenesis. *Genome Research* 26:1233-44 doi: 10.1101/gr.199547.115.

Xu P, Vernooy SY, Guo M, Hay BA. 2003. The *Drosophila* microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. *Current Biology* 13(9):790-5 doi: 10.1016/s0960-9822(03)00250-1.

Xu Z, Yan Y, Qian L, & Gong Z. 2017. Long non-coding RNAs act as regulators of cell autophagy in diseases (Review). *Oncology Reports* 37, 1359-1366 doi: 10.3892/or.2017.5416.

Yacobi-Sharon K, Namdar Y, Arama E. 2013. Alternative germ cell death pathway in *Drosophila* involves HtrA2/Omi, lysosomes, and a caspase-9 counterpart. *Developmental Cell* 25:29-42 doi: 10.1016/j.devcel. 2013.02.002.

Zhang S, Pan C, Lv X, Wu W, Chen H, Wu W, Wu H, Zhang L. 2017. Repression of Abd-B by Polycomb is critical for cell identity maintenance in adult *Drosophila* testis. *Scientific Reports* 7:5101 doi: 10.1038/s41598-017-05359-0.

Figure 1

Comparison of the expression of testis-specific and nonspecific genes

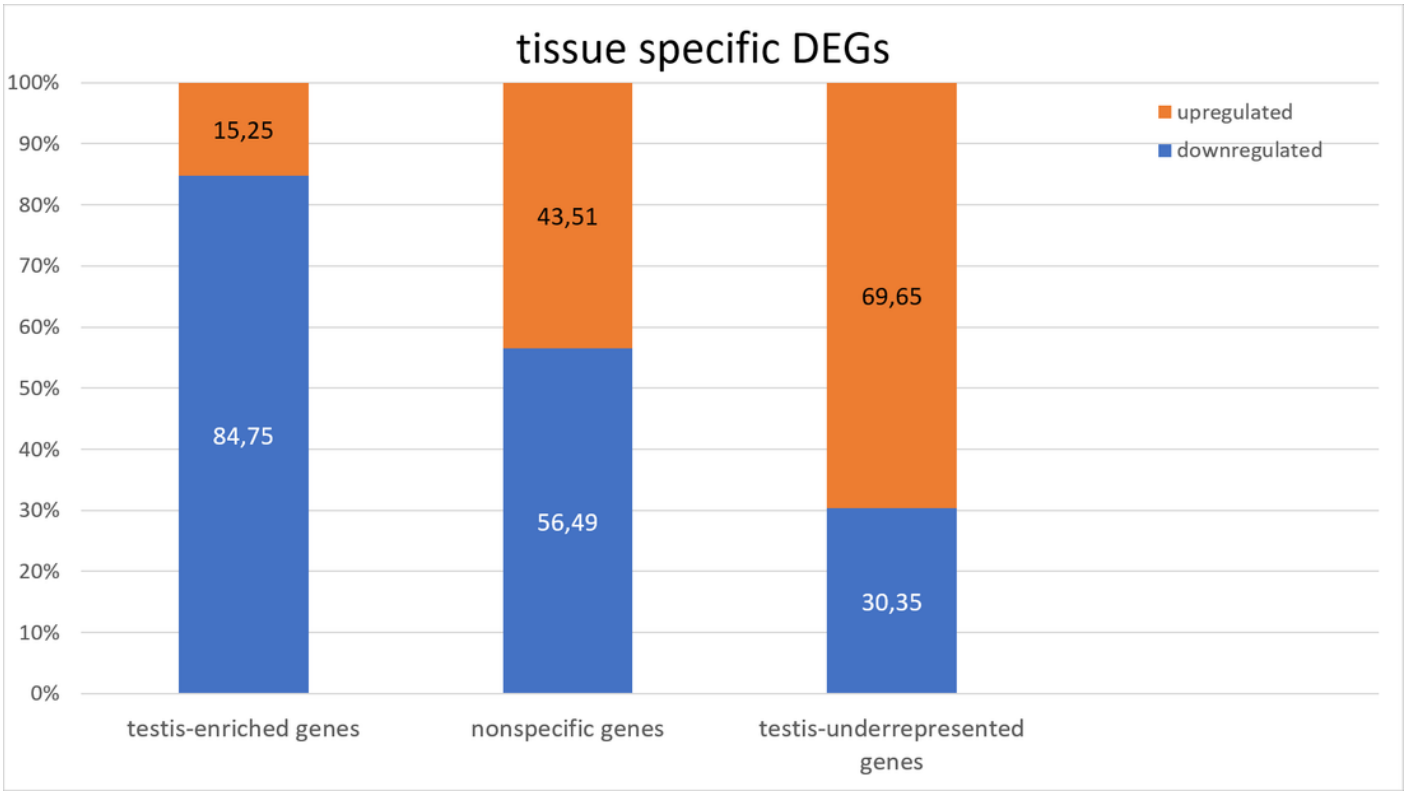

Figure 2

FlyEnrichr GO analysis of testis-specific and nonspecific genes

(A) Top GO terms enriched among testis-enriched downregulated genes. FlyEnrichr GO analysis, sorted by combined score. **(B)** GO terms most enriched among upregulated genes under-represented in testes. FlyEnrichr GO analysis, sorted by combined score.

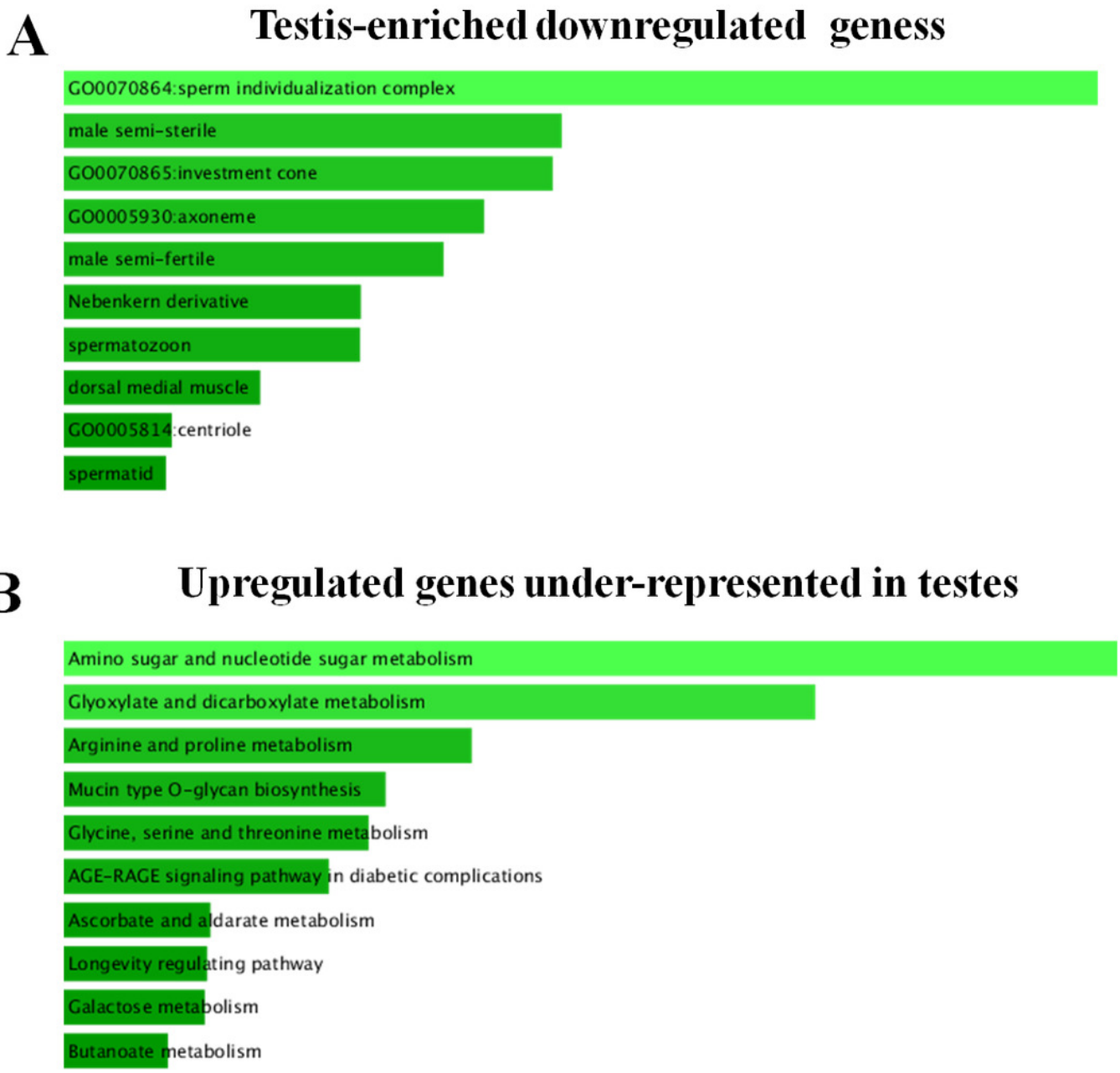
A
Testis-enriched downregulated geness
GO0070864:sperm individualization complex
male semi-sterile
GO0070865:investment cone
GO0005930:axoneme
male semi-fertile
Nebenkern derivative
spermatozoon
dorsal medial muscle
GO0005814:centriole
spermatid
B
Upregulated genes under-represented in testes
Amino sugar and nucleotide sugar metabolism
Glyoxylate and dicarboxylate metabolism
Arginine and proline metabolism
Mucin type O-glycan biosynthesis
Glycine, serine and threonine metabolism
AGE-RAGE signaling pathway in diabetic complications
Ascorbate and aldarate metabolism
Longevity regulating pathway
Galactose metabolism
Butanoate metabolism

Figure 3

Distribution of genes associated with GO cell death related terms with increased, decreased, or unchanged expression levels

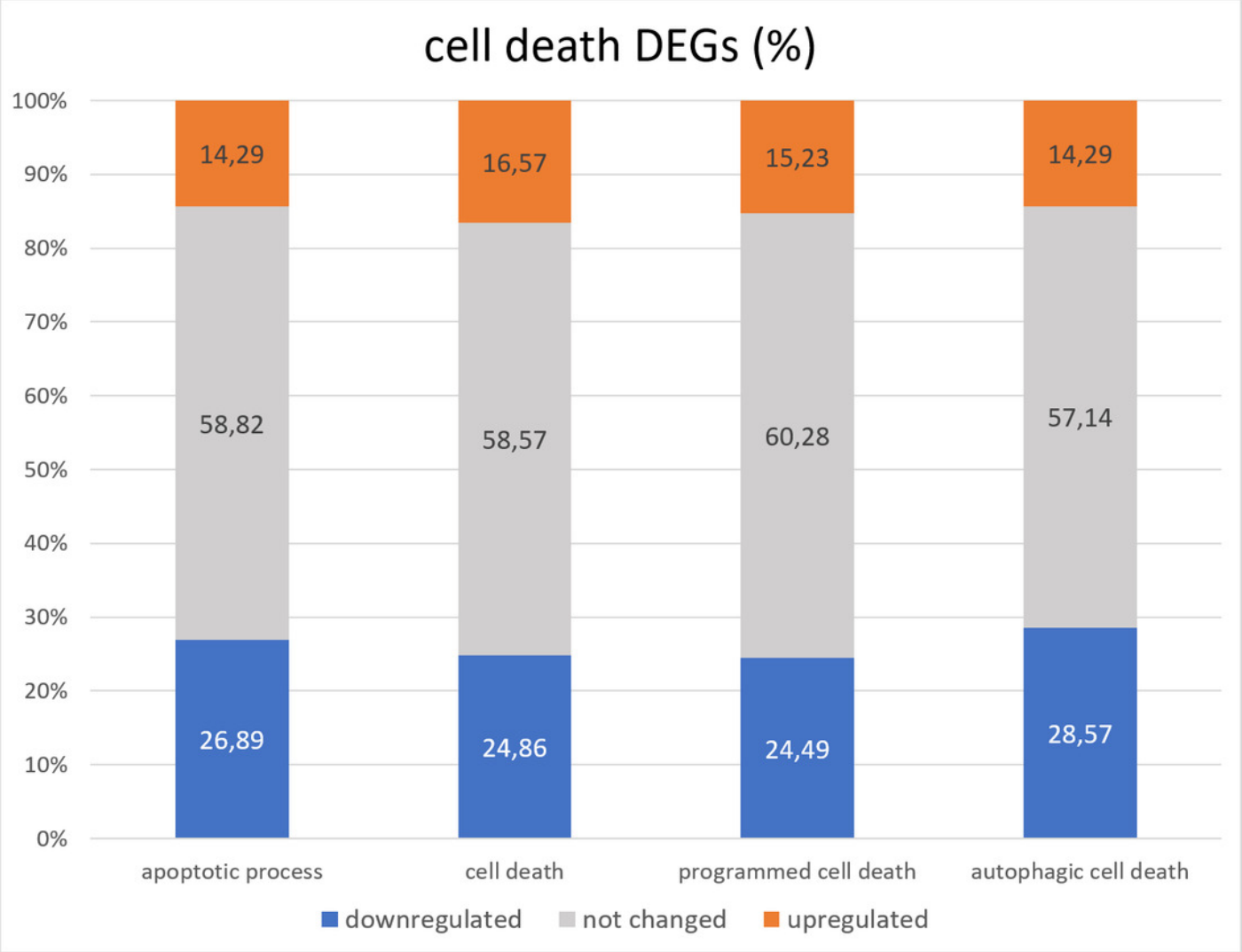

Figure 4

Relative density of probable GAGA-binding sites in various gene selections.

In the Flybase GO lists of genes, the figure presents the proportion of genes containing a GAGA-binding site(s) in the –500…+1 region relative to the TSS, with normalization to the proportion of such genes among all *Drosophila* genes, according to SITECON results on type 1 GAGA-binding sites (“GAGA1 up500” in the figure) and type 2 GAGA-binding sites (“GAGA2 up500” in the figure).

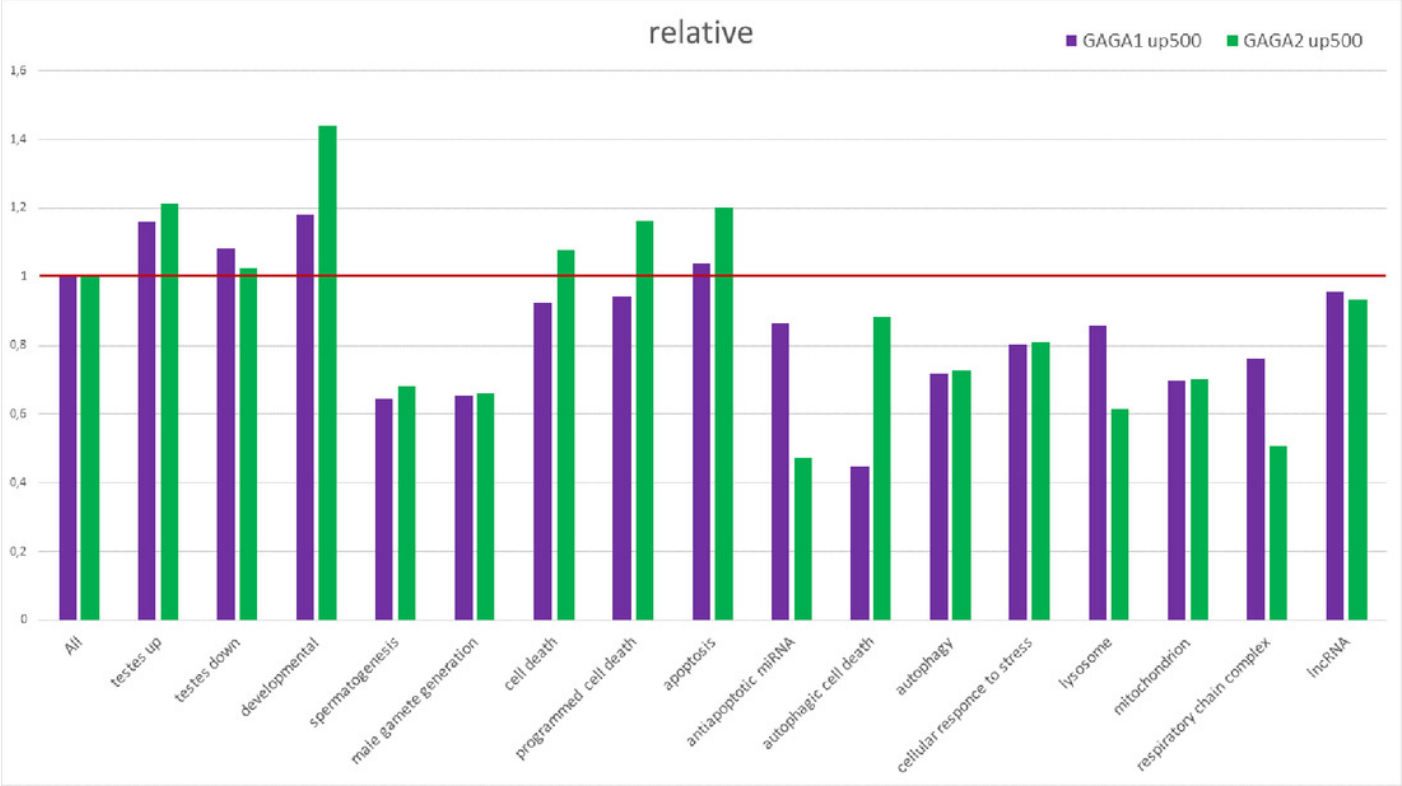

Table 1(on next page)

RNA-seq data alignment statistics

1 Table 1. RNA-seq data alignment statistics

Library	Number of reads	Number of uniquely mapped reads	Number of reads mapped to multiple loci	Number of reads mapped to too many loci
362-testes1.tr	41 397 803	39 832 848 (96.22%)	1 209 655 (2.92%)	86 802 (0.20%)
362-testes2.tr	42 823 074	40 781 860 (95.23%)	1 627 713 (3.80%)	117 174 (0.25%)
362-testes3.tr	36 436 374	34 488 834 (94.65%)	1 564 303 (4.29%)	104 373 (0.27%)
Or-testes10.tr	37 418 763	36 165 776 (96.65%)	971 386 (2.60%)	94 938 (0.23%)
Or-testes11.tr	37 927 024	36 548 954 (96.37%)	1 120 371 (2.95%)	144 254 (0.34%)
Or-testes12.tr	38 382 924	36 977 384 (96.34%)	1 110 115 (2.89%)	144 196 (0.40%)

2