CRISPR/Cas9 mediated genetic resource for unknown kinase and phosphatase genes in *Drosophila*

Menghua Wu1,2, Xuedi Zhang2, Wei Wei1, Li Long1, Sainan An3 & Guanjun Gao2,1 ✉

Kinases and phosphatases are crucial for cellular processes and animal development. Various sets of resources in *Drosophila* have contributed significantly to the identification of kinases, phosphatases and their regulators. However, there are still many kinases, phosphatases and associate genes with unknown functions in the *Drosophila* genome. In this study, we utilized a CRISPR/Cas9 strategy to generate stable mutants for these unknown kinases, phosphatases and associate factors in *Drosophila*. For all the 156 unknown gene loci, we totally obtained 385 mutant alleles of 105 candidates, with 18 failure due to low efficiency of selected gRNAs and other 33 failure due to few recovered F0, which indicated high probability of lethal genes. From all the 105 mutated genes, we observed 9 whose mutants were lethal and another 4 sterile, most of which with human orthologs referred in OMIM, representing their huge value for human disease research. Here, we deliver these mutants as an open resource for more interesting studies.

Phosphorylation, the most common post-translational modification (PTMs) of proteins, is involved in multiple biological processes in eukaryotic organisms. Kinases and phosphatases collaborate to regulate the levels of this modification1,2, and mutations of these always act as causal factors in human diseases3,4. Thus, a deep exploration of kinase and phosphatase genes function will aid in the study of human diseases in the clinical5. *Drosophila melanogaster* is an ideal system for the dissection of kinase and phosphatase gene function because of its high gene conservation with the human genome and low gene redundancy in its own genome6. P-element-mediated insertion and RNAi both are well-established methods for gene function analysis in *Drosophila*, and there are abundant genetic resources for both of these systems7,8. In fact, *Drosophila* genetic screening using these resources has contributed significantly to the identification of kinases, phosphatases, and their regulators9–11. However, there are still many unknown functional genes with predicted kinase and phosphatase domains in the *Drosophila* genome, and we sought to develop a high-throughput gene-targeting strategy to aid with characterizing them.

The emergence of CRISPR/Cas9 made it possible for us to manipulate high-throughput mutagenesis in *Drosophila*12–15. In our previous work, the CRISPR/Cas9 mediated mutagenesis frequency even could reach 100% in some cases15. Besides, we also applied this CRISPR/Cas9 system to a *Drosophila* testis specific lncRNA knock-out project and demonstrated its high-throughput application16. In this study, we carried out another large-scale mutagenesis for 156 unknown kinase and phosphatase genes and finally obtained 385 mutant alleles for 105 individual genes with 33 high potential lethal genes. Using these mutants, we uncovered several functional kinases and phosphatases, providing valuable genetic resource for phosphorylation clinical research.

Results

Unknown kinase, phosphatase and associate factors in *Drosophila* genome. *Drosophila* genome encodes 376 kinases, 159 phosphatases and 27 associate factors that associated with these enzymes, such as cyclins and regulatory subunits2,9,17. After thoroughly searching in FlyBase, we compiled a list of 85 kinases, 63 phosphatases and 9 associate factors with ambiguous functional annotations (Supplementary Table 1). This accounts for 23%, 40% and 33% of all kinases, phosphatases and associate factors encoded in the *Drosophila* genome, respectively. Among all these 156 unknown kinase, phosphatase and associate genes, we found 70% (119/156) were conserved with human genome, 90% (95/119) of which human orthologs were linked to human disease in Online Mendelian Inheritance in Man (OMIM). Gene group category divided these candidates into several families. Gene ontology analysis showed they might be involved in multiple developmental processes (Fig. 1). These

---

1School of Life Sciences, Tsinghua University, Beijing, 100084, P.R. China. 2School of Life Science and Technology, ShanghaiTech University, Shanghai, 201210, P.R. China. 3National Institute of Biological Sciences, Beijing, 102206, P.R. China. ✉E-mail: gaogj@shanghaitech.edu.cn
all indicated their important functions, while perhaps due to current limited resources, their functions have not been defined yet. Thus, we planned to generate stable mutants to define the function of all these 156 unknown kinases and phosphatases.

**CRISPR/Cas9 mediated mutagenesis of the unknown candidates.** To ensure the mutagenesis efficiency and completely disrupt the gene function, we designed two gRNAs per gene around the translational start codon or at least before 1/3 of the coding sequence when carrying out mutagenesis according to the standard CRISPR/Cas9 strategy. In case that these unknown genes could be lethal when mutated, we used different FRT flies for microinjection according to their genomic distribution in order to perform tissue-specific mosaic analysis. Through separate operation on the five individual FRT groups, we finally obtained 385 mutant alleles for 105 genes from all the 156 candidates and another 33 s failed due to few F0, indicating high probability of lethal genes, which we called potential lethal genes. While for the remnant 18 failure, we found they were nothing with the gRNA working efficiency, length, sequence specificity or PAM sequence specificity (Supplementary Figure 1), which perhaps mainly caused by limited detected F0/F1. Consistent with results from other organisms, we also detected insertions, deletions including frameshift and in-frame mutations among the 385 alleles (Fig. 2, Supplementary Note 1), in which case there were 11 genes of the 105 with only in-frame mutants in the resource. Of the 385 mutant alleles, we randomly chose 6 mutants to detect the mRNAs of the corresponding mutant genes using qRT-PCR. As expected, we observed reduced mRNAs to different degrees of detected genes (Supplementary Figure 2), demonstrating our resource with corresponding gene deficiency.

**Obvious phenotypes in the CRISPR/Cas9 resource.** Considering complicated functions of the kinases and phosphatases, we performed simple lethality and sterility screening for the resource only in this study. Taken together, we observed 9 genes with lethal mutants, 1 gene with male sterile mutants and 3 genes with female sterility...
mutants, these genes were predicted involved in various developmental processes (Table 1). To be recommended, of the 3 female sterile events, mutations of CG3608 were all in-frame styles, implicating the importance of the missed amino acids in the mutants, meanwhile, we could not exclude the possibility that there were off-targeting effects in the mutants. The same situation occurred in the rescue experiments for the phenotypic mutants. Of the 13 phenotypic mutants, we totally generate 4 transgenic lines including CG17028, CG14305 and CG15743 to rescue the corresponding mutants, but we found the CG17028 transgenic lines could not rescue the lethality of CG17028 mutants, which might be caused by ectopic transgenes or off-targeting effects. Besides these, we also observed other abnormal lethal events in the resource, such as CG12229 and CG17027, with different lethality between similar or same genotypic mutants, even one line of CG17028, without any mutation detected near the gRNA recognition sites (Supplementary Figure 3, Supplementary Note 2). Anyway, despite uncertainty of the total resource, we could define more certain phenotypes of unknown candidates such as lethality and sterility. We look forward to more interesting phenotypes in our resource.

Discussion
Our work in this study delivered a valuable genetic resource for unknown kinase, phosphatase and associate genes in Drosophila. Using this resource, we can carry out lots of meaningful and interesting screenings besides of lethality and fertility. Of the lethal and sterile genes identified in this study (Table 1), we re-searched for gene function and found some of them had been experimentally investigated, such as CG34380, proved to be required for normal thermal nociception in a RNAi-based screening study19; CG6767, proved to affect olfactory behavior using P-element insertional mutagenesis together with targeted RNAi20; and CG34455, prove to encode a
The female sterility phenotype of CG3608 was observed from the in-frame mutant alleles, because we did not obtain frameshift mutant alleles for CG3608 in our resource. Obvious phenotypes in the CRISPR/Cas9 resource. aThe lethality phenotype of CG33671 and male sterility phenotype of CG14305 both can be rescued by their own original genomic transcripts. bThe lethality phenotype of CG17028 cannot be rescued by its original genomic transcript. cThe female sterility phenotype of CG33671MVK cannot be rescued by its original genomic transcript. dThe female sterility phenotype of CG33671 was observed from the in-frame mutant alleles, because we did not obtain frameshift mutant alleles for CG33671 in our resource.

### Methods

#### Fly strains.

w1118, FM7a, Sco/Cyo, Sb/TM6B, Sco/Cyo; Sb/TM6B were stocks from our own lab (ShanghaiTech University, China); FRT19A was from Jose Carlos Pastor-Pareja’s lab (Tsinghua University, China); FRT40A, FRT42D, FRT79D were from Xin Cai’s lab (Tsinghua University, China); FRT82B was from Renjie Jiao’s lab (Tsinghua University, China); attP40, attP2 were from Xiaolin Bi’s lab (Dalian Medical University, China); FRT40A, FRT42D, FRT79D were from Ting Xie’s lab (Tsinghua University, China); FRT19A, FRT82B were from Shao Dong’s lab (Institute of Biophysics, CAS, China); attP40, attP2 were from Xiaolin Bi’s lab (Dalian Medical University, China).

All flies were cultured at 25°C.

#### gRNA in vitro transcription and microinjection.

gRNAs were transcribed as previously described15, then mixed with Cas9 mRNA at a final concentration of 500 ng/μl and injected into different FRT flies.

#### T7EI assay and mutation identification.

Dead larvae or single flies were squashed in 30 μL of squashing buffer containing 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 25 mM NaCl, 1 mg/mL proteinase K (Takara, Beijing, China), and incubated at 37°C for 1 h, followed by heating to 95°C for 2 min and used as PCR templates. PCR was performed in a 2 × Taq MasterMix (Aidlab Biotech, Beijing, China), and PCR products were then digested with T7 Endonuclease I (Vieswold Biotech, Beijing, China), and T7EI positive F1 PCR products were sequenced for mutants.

#### Fertility screening.

Every three males or virgin-females of our resource with viable homozygous were crossed with w1118 virgin-females or males. Parents were kept for 7 days and discarded. Then adult flies of each cross were counted for fertility of the viable resource. The ones with no offspring were defined as sterile.
Plasmid constructions and generation of transgenic flies. Full length genomic fragments were amplified using the following primers: CG33671-KpnIF, 5′-GATATCTGagttacCGCTGCTGAGCACTCTTA-3′, CG33671-NotIR, 5′-GATATCTGagttacCGCTGCTGAGCACTCTTA-3′; CG14305-KpnIR, 5′-GATATCTGagttacCGCTGCTGAGCACTCTTA-3′, and CG14305-NotIR, 5′-GATATCTGagttacCGCTGCTGAGCACTCTTA-3′. PCR products were then cut with KpnI and NotI, and then cloned into the pUAST-attB plasmid. pUAST-attB-CG33671 and pUAST-attB-CG14305 were introduced into the attP2 site by the phiC31 integration method, and the pUAST-attB-CG17028 and pUAST-attB-CG14305 were introduced into the attP40 site.

Data availability
The flies will be available by mail by contacting the corresponding author, Dr. Gao, who will maintain the stocks in the laboratory.

Received: 18 November 2019; Accepted: 13 April 2020;
Published online: 30 April 2020

References
1. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. Nature 429, 192–194, https://doi.org/10.1038/nature02671 (2004).
2. Schürrer, J., Krämer, K., Deitmar, A. & Evers, B. M. Genome-wide identification of phospho-regulators of Wnt signaling in Drosophila. Proc. Natl. Acad. Sci. USA 112, 10296–10301, https://doi.org/10.1073/pnas.1502425112 (2015).
3. Sopko, R., Cong, L., Yu, Z., Evers, B. M. & Schäfer, U. Highly efficient genome modifications mediated by CRISPR/Cas9 in Drosophila. Nucleic acids research 43, 7998–8006, https://doi.org/10.1093/nar/gkv1024 (2015).
4. Brown, E. B., Rayens, E. & Rollmann, S. M. The Gene CG6767 Affects Olfactory Behavior in Drosophila melanogaster. Genetics 199, 1139–1145, https://doi.org/10.1534/genetics.114.169646 (2015).
5. Cohen, P. Protein kinases—the major drug targets of the twenty-first century? Nature reviews Drug discovery 1, 309, https://doi.org/10.1038/nrd7773 (2002).
6. Fortini, M. E., Skupski, M. P., Boguski, M. S. & Hariharan, I. K. A survey of human disease gene counterparts in the Drosophila genome. The Journal of cell biology 150, F57–F62, https://doi.org/10.1083/jcb.150.2.157 (2000).
7. Bellini, H. J. et al. The BDGP gene disruption project. Genetics 167, 761–781, https://doi.org/10.1534/genetics.104.026427 (2004).
8. Ni, J.-Q. et al. A genome-scale shRNA resource for transgenic RNAi in Drosophila. Nature methods 8, 405, https://doi.org/10.1038/nmeth.1592 (2011).
9. Deitmar, A. Genome-wide identification of phospho-regulators of Wnt signaling in Drosophila. Development 142, 1502–1515, https://doi.org/10.1242/dev.116715 (2015).
10. Mukherjee, T., Witten, D. M., Schäfer, U. & Zeidler, M. P. Identification of Drosophila Genes Modulating Janus Kinase/Signal Transducer and Activator of Transcription Signal Transduction. Genetics 172, 1683–1697, https://doi.org/10.1534/genetics.105.046904 (2006).
11. Read, R. D. et al. A Kinome-Wide RNAi Screen in Drosophila Glia Reveals That the RIO Kinases Mediate Cell Proliferation and Survival through TORC2-Akt Signaling in Glialblasta. Plos Genetics 9, e1003253, https://doi.org/10.1371/journal.pgen.1003253 (2013).
12. Mali, P. et al. RNA-guided human genome engineering via Cas9. Science 339, 823–826, https://doi.org/10.1126/science.1232033 (2013).
13. Cong, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339, 819–823, https://doi.org/10.1126/science.1231143 (2013).
14. Gratz, S. J. et al. Genome engineering of Drosophila using the CRISPR RNA-guided Cas9 nuclease. Genetics 194, 1029–1035, https://doi.org/10.1534/genetics.113.152710 (2013).
15. Yeh, K. et al. Highly efficient genome modifications mediated by CRISPR/Cas9 in Drosophila. Genetics 195, 289–291, https://doi.org/10.1534/genetics.113.153825 (2013).
16. Wen, K. et al. Critical roles of long noncoding RNAs in Drosophila spermatogenesis. Genome Res 26, 1233–1244, https://doi.org/10.1101/gr.199547.115 (2016).
17. Soppke, R. et al. Combining genetic perturbations and proteomics to examine kinase-phosphatase networks in Drosophila embryos. Developmental cell 31, 114–127, https://doi.org/10.1016/j.devcel.2014.07.027 (2014).
18. Varshney, G. K. et al. High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. Genome Research 25, 1030, https://doi.org/10.1101/gr.186379.114 (2015).
19. Conde, J. M., Mauthner, S. E., Wang, Y., Skene, J. P. & Tracey, W. D. Jr Nociceptor-enriched genes required for normal thermal nociception. Cell reports 16, 293–303, https://doi.org/10.1016/j.celrep.2016.06.003 (2016).
20. Brown, E. B., Rayens, E. & Rollmann, S. M. The Gene CG6767 Affects Olfactory Behavior in Drosophila melanogaster. Behavior genetics 49, 317–326, https://doi.org/10.1007/s10519-019-09949-8 (2019).
21. Marzio, A., Meriglino, C., Gatti, M. & Verni, F. Sugar and chromosome stability: clastogenic effects of sugars in vitamin B6-deficient cells. PLoS genetics 10, e1004199, https://doi.org/10.1371/journal.pgen.1004199 (2014).
22. Evers, B. M. et al. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. Nature Biotechnology 34, 631, https://doi.org/10.1038/nbt.3536 (2016).
23. Moreno-Mateos, M. A. et al. CRISPRscn: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nature methods 12, 982, https://doi.org/10.1038/nmeth.3543 (2015).
24. Cullott, G. et al. CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. Nature communications 10, 1136, https://doi.org/10.1038/s41467-019-09006-2 (2019).
25. Smylaker, I. M. et al. Rationally engineered Cas9 nucleases with improved specificity. Science 351, 84–88, https://doi.org/10.1126/science.aad5227 (2016).
26. Hu, Y. et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC bioinformatics 12, 357, https://doi.org/10.1186/1471-2105-12-357 (2011).
27. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research 37, 1–13, https://doi.org/10.1093/nar/gkt923 (2008).

Acknowledgements
We thank Dr. Ting Xie (Tsinghua University, China) for valuable advices on designing the project; Dr. Xin Liang (Tsinghua University, China) for critical suggestions on manuscript preparation.
Author contributions
M.W. and G.G. designed the experiments and wrote the manuscripts; M.W., W.W., and L.L. generated the mutants; M.W. performed phenotype classification; X.Z. performed qRT-PCR; S.A. analyzed the data. All authors reviewed the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-64253-4.

Correspondence and requests for materials should be addressed to G.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020