Genome-wide CRISPR screening reveals genes essential for cell viability and resistance to abiotic and biotic stresses in *Bombyx mori*

Jiasong Chang, Ruolin Wang, Kai Yu, et al.

Genome Res. published online May 18, 2020
Access the most recent version at doi:10.1101/gr.249045.119

Genome-wide CRISPR screening reveals genes essential for cell viability and resistance to abiotic and biotic stresses in *Bombyx mori*

Jiasong Chang[1,#], Ruolin Wang[1,#], Kai Yu[1], Tong Zhang[1], Xiaoxu Chen[1], Yue Liu[1], Run Shi[1], Xiaogang Wang[1], Qingyou Xia[1,2,3,*], Sanyuan Ma[1,2,3,*].

1 Biological Science Research Center, Southwest University, Chongqing 400716, China

2 Chongqing Engineering and Technology Research Center for Novel Silk Materials, Southwest University, Chongqing 400716, China

3 Chongqing Key Laboratory of Sericulture, Southwest University, Chongqing 400716, China

# These authors contributed equally.

* Corresponding author, E-mail: xiaqy@swu.edu.cn, masy@swu.edu.cn

Running title

Genome-wide CRISPR screening in *Bombyx mori*

Keywords: CRISPR/Cas9, sgRNA library, Genome-wide screen, *Bombyx mori*, Essential genes, Environment stimuli, Host-pathogen interaction

Abstract

High-throughput genetic screens are powerful methods to interrogate gene function on a genome-wide scale and identify genes responsible to certain stresses. Here, we developed a *piggy*Bac strategy to deliver pooled sgRNA libraries stably into cell lines. We used this strategy to conduct a screen based on genome wide clustered regularly interspaced short palindromic repeat technology (CRISPR)/Cas9 in *Bombyx mori* cells. We first constructed a single guide RNA (sgRNA) library containing 94,000 sgRNAs, which targeted 16,571 protein-coding genes. We then generated knockout collections in BmE cells using the *piggy*Bac transposon. We identified 1,006 genes that are essential for cell viability under normal growth conditions. Of the identified genes 82.4% (829 genes) were homologous to essential genes in seven animal species. We also identified 838 genes whose loss facilitated cell growth. Next, we performed context-specific positive screens for resistance to biotic or non-biotic stresses using temperature and baculovirus separately which identified several key genes and pathways from each screen. Collectively, our results provide a novel and versatile platform for functional annotations of *B. mori* genomes and deciphering key genes responsible for various conditions. This study also demonstrates the effectiveness, practicality, and convenience of genome-wide CRISPR screens in non-model organisms.

Introduction

A major task in biology is to functionally annotate genomes and to identify genetic elements underlying normal cellular processes or diseases. In the past decade, genome-wide loss-of-function screens, which is a powerful hypothesis-free approach to discover genes underlie certain biological processes, have been used successfully to address many fundamental biological questions (Boutros et al. 2004; Carette et al. 2009). In diploid eukaryotic cells, high throughput loss-of-function screens have primarily been achieved by RNA interference (RNAi), either in well-by-well arrays or in barcoded pooled libraries (Mohr et al. 2014). Recently, the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein 9 (Cas9) (CRISPR) system has been explored for pooled genome-scale functional screening (Shalem et al. 2014; Wang et al. 2014). The CRISPR system largely overcomes major drawbacks of RNAi such as incomplete loss-of-function and off-target effects. More fitness genes can be identified using the CRISPR system compared to using RNAi (Hart et al. 2014; Shalem et al. 2014; Wang et al. 2014). Furthermore, the CRISPR system has been confirmed to be efficient in identifying genes for drug resistance (Koike-Yusa et al. 2014; Hou et al. 2017), tumorigenesis (Chen et al. 2015; Song et al. 2017; Xu et al. 2017), immune response (Parnas et al. 2015), host-pathogen interactions (Ma et al. 2015; Kim et al. 2018), and cancer immunotherapy (Patel et al. 2017).

Currently, CRISPR library screenings are only available in mammalian cells, bacteria (Liu et al. 2017; Wang et al. 2018), and *Drosophila* cells (Bassett et al. 2015; Viswanatha et al. 2018). To gain a comprehensive understanding of the genomic basis of universal biological processes or particular phenomena, genome wide

loss-of-function screens in more organisms is indispensable and urgently needed. *B. mori* is a lepidopteran insect with both research and agricultural importance. Genetics and genomics studies in *B. mori* have provided insight into many fundamental biological questions, such as domestication history, sex determination, metamorphosis, and silk production (Goldsmith et al. 2005; Xia et al. 2009; Omenetto and Kaplan 2010; Kiuchi et al. 2014; Xia et al. 2014). However, approximately over 80% of the genes in *B. mori* remain functionally uncharacterized.

In the current study, we developed a novel method to conduct genome wide CRISPR screens in *B. mori* cells. A plasmid library containing 94,000 single guide RNAs (sgRNAs) was generated and stably delivered into BmE cells using the *piggy*Bac transposon. Moreover, we showed that this system can be used to identify genes essential for cell viability in normal conditions, and screen resistance genes for both biotic and non-biotic stresses. This system provides a powerful platform to rapidly elucidate the functional genome of *B. mori* and to identify key genes and mechanisms associated with particular biological processes. Furthermore, since *piggy*Bac is a universal tool that can drive gene delivery in a wide range of organisms, including insects, plants and animals, this approach should be amenable to adapting CRISPR screens in other cell lines.

Results

Development of a *piggy*Bac library delivery method for *B. mori* cells

In mammalian cells, pooled DNA libraries are often delivered by lentiviral vectors. However, in *B. mori*, lentiviral vectors are extremely inefficient. As an

alternative strategy to deliver the sgRNA library, we chose *piggy*Bac, which has been shown to have high transformation activity in both *B. mori* germline and cultured cells. Previously, we established a binary transient CRISPR vector for gene knockout in the *B. mori* embryonic cell line BmE (Ma et al. 2017). To allow for the integration of both Cas9 and sgRNA expression cassettes into the genome of a single cell, we combined two cassettes into an all-in-one vector, pB-CRISPR (Supplemental Materials). A zeocin expression cassette in which a zeocin antibiotic resistance gene is driven by the *IE2* promoter, was also included in pB-CRISPR to facilitate the establishment of stable transgenic lines (Fig. 1A).

To test the knockout efficiency of pB-CRISPR, we first established a cell line that stably expressed enhanced green fluorescent protein (EGFP; BmE-EGFP) using a second general transposon, *Minos* (Supplemental Fig S1A and S1C; Supplemental Materials). Three sgRNAs targeting EGFP were designed and constructed into pB-CRISPR, and two nonspecific sgRNAs were used as controls (Supplemental Fig S2; Supplemental Table S1). Each vector was co-transfected with a *piggy*Bac transposase expression vector (A3-Helper) into BmE-EGFP. After two months of selection with zeocin, EGFP fluorescence was totally abolished in cells transfected with all three EGFP targeting sgRNAs, while cells transfected with either control sgRNA remained unchanged (Fig 1C; Supplemental Fig S1D). Flow cytometry (FCM) analysis also showed a significant decrease in EGFP signals in the targeting sgRNA transfected cells, whereas the signals were not decreased in control cells (Fig. 1B).

The targeting regions of EGFP were amplified by the polymerase chain reaction (PCR) and cloned for sequencing. Sequencing results from a total of 44 clones showed that nearly 100% of the EGFP alleles had genetic variations with > 70%

116 causing frame shift mutations (Fig. 1D; Supplemental Fig S1B). The editing
117 efficiency was much higher than that achieved by transient transfection with binary
118 vectors (Ma et al. 2017). Similar to our previous observations, small deletions
119 dominated the types of mutations (71.4–92.3%) (Ma et al. 2014). From these results,
120 we concluded that *piggy*Bac was suitable to stably deliver CRISPR constructs into the
121 *B. mori* genome and that our established pB-CRISPR vector was suitable for efficient
122 gene knockout in BmE cells.

123

124 **Design and construction of the CRISPR sgRNA library**

125 To construct the genome-wide sgRNA library, we designed a strategy as
126 illustrated in Fig. 2A. The library was constructed according to the following
127 procedure. (1) The whole genome of *B. mori* was searched for all exons which were
128 used to design a total of 1,534,227 sgRNAs (Fig. 2B). The sgRNAs were then ranked
129 using the following two criteria. First, location of an sgRNA within the first half of
130 the coding region to ensure any frame shift at the target site could disrupt the protein
131 function. Second, to avoid potential off-target effects, sgRNA sequences were unique
132 at the seed region located 12 bp downstream of the protospacer adjacent motif (PAM)
133 and were as unique as possible in the non-seed region. Finally, 94,000 sgRNAs were
134 chosen (~ 6 sgRNAs per gene) and synthesized on a microarray chip (Supplemental
135 Table S2). (2) The sgRNA oligonucleotide pool was subsequently amplified by PCR
136 and cloned into pB-CRISPR to form the CRISPR knockout plasmid library
137 (pB-CRISPR-lib; Fig. 2A). More than 10^8 clones on 240 petri dishes were selected
138 with ampicillin and > 1,000 clones per sgRNA were carried to maintain the sgRNA
139 library diversity. Then we evaluated the quality of library by deep sequencing. A total

of 51,431 sgRNAs were detected, which covered 96.3% of all *B. mori* genes. Over 87.4% of the genes had more than two sgRNAs (Fig. 2C) and 72% of sgRNAs had 11–200 reads (Fig. 2D). These results suggested that the coverage, accuracy and diversity of the pB-CRISPR-lib were sufficient for the subsequent experiments. (3) Then the pB-CRISPR-lib and the *piggy*Bac transposase expression vector (A3-Helper) were co-transfected into BmE cells. To maintain the high diversity of the original library, ~2,000 cells per sgRNA were transfected and ~2,000 cells per passage carried. To enrich sgRNA-harboring cells, the transfected cells were selected with zeocin for two months followed by culturing with complete medium without antibiotic. Cells surviving these conditions composed the BmE genome-scale CRISPR/Cas9 knockout library (BmEGCKLib).

To evaluate the coverage and diversity of the BmEGCKLib, genomic DNA was extracted from 4×10^7 cells and the sgRNA regions were amplified by PCR for deep sequencing. A total of 48,982 sgRNAs, which accounted for 95.2% of the pB-CRISPR-lib, were detected with at least one read (Fig. 2E). The compositions of the BmEGCKLib and pB-CRISPR-lib libraries were highly correlated (Fig. 2F; correlation coefficient = 0.99). The number of genes containing 1–6 sgRNAs in the BmEGCKLib was basically the same as the corresponding number in the pB-CRISPR-lib vector library (Fig. 2C). These results indicated that the genome-wide knockout cell library of BmE was successfully generated with sufficient coverage to perform further genetic screens.

Screening of essential or growth-restricting genes under normal conditions

After generating the BmEGCKLib, we next aimed to identify essential genes that affected the survival on normal growth of BmE based on depletion of their sgRNAs in the library population, as well as growth-restricting genes that provided growth advantages based on sgRNA enrichment over time. BmEGCKLib cells were grown under normal conditions and approximately 4×10^7 cells harvested at each of three time points: immediately (BmE-Lib1), one month (BmE-Lib2), and two months BmE-Lib3) after completion of zeocin selection (Supplemental Fig S3A). Deep sequencing of the libraries at different time points revealed a reduction in the correlation coefficients over time (Supplemental Fig S3B); the number of sgRNAs also decreased in the BmE-Lib2 and BmE-Lib3 groups (Fig. 3A). Although the overall abundance of sgRNAs showed a gradual depletion, some sgRNAs were significantly enriched (Fig. 3B). These observations suggested that BmEGCKLib could be used for screening essential and growth-restricting genes.

We used the MAGeCK program (Li et al. 2014) to compute the negative (sgRNA depleted) or positive (sgRNA enriched) scores for each gene. Both the $\log_{10}$ fold-changes of individual sgRNAs and $\log_{10}$(negative scores) of genes showed good correlations between two biological replicates (Supplemental Fig S4; Supplemental Table S3 ; r = 0.87 and r = 0.76, respectively), indicating a considerable high reproducibility of our fitness screens. All of the *B. mori* genes were ranked by negative or positive scores and a p-value less than 0.05 was used as a threshold value. We identified 1,006 genes as essential for normal growth (Supplemental Fig S3C; Supplemental Table S4) and 838 genes as growth-restricting genes (Supplemental Fig S3D; Supplemental Table S4). The relative ratios of essential or growth-restricting genes to the total number of coding genes (~6.1% and ~5.1%, respectively) were

187 equal to those reported in other species (Fig. 3C). Furthermore, we compared the
188 1,006 essential genes to those identified in *Homo sapiens*, *Mus musculus*, *Drosophila*
189 *melanogaster*, *Saccharomyces cerevisiae*, *Danio rerio*, *Caenorhabditis elegans*,
190 *Arabidopsis thaliana*, and *Bacillus thuringiensis*. Considerable overlap was found
191 between *B. mori* essential genes and available eukaryotic species; however, essential
192 genes of BmE only partially overlapped with the essential genes reported for the
193 prokaryote *B. thuringiensis* (Fig. 3E). Since true essential genes must be
194 transcriptionally active, we measured the expression level of essential genes using
195 RNA-seq data to validate the CRISPR results (Fig. 3D; Supplemental Table S5). As
196 expected, the essential genes exhibited higher transcriptional levels compared to the
197 average expression level of all genes, whereas the transcriptional levels for the
198 growth-restricting genes were lower. Besides, the same tendency was also found
199 among three gene groups in the silk gland, suggesting that the essential genes defined
200 in BmE cells may also be essential in specialized tissues, or even in whole organisms
201 (Fig 3D).

202

203 **Analysis of essential genes for *B. mori* cells**

204 We next investigated the essential and growth-restricting genes of *B. mori* cells
205 from different aspects, including genome-wide distribution, functional categorization,
206 and cellular localization. Comparison of all genes density and selected genes revealed
207 that overall both essential and growth-restricting genes were evenly distributed across
208 all chromosomes, with the exception of a few specific chromosomal regions
209 (Supplemental Fig S5). Random distribution of essential and growth-restricting genes
210 is also observed in mammalian cells(Yilmaz et al. 2018).

Functional categorization by Gene Ontology (GO) analysis revealed that the essential genes were primarily related to core cell components (cell part, binding, and catalytic activity), and metabolism (Supplemental Fig S6A). A majority of basic Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were significantly enriched in essential genes (Fig. 4A), including Ribosome (42.1%), RNA transport (29.5%), mRNA surveillance (37.5%), Spliceosome (26.2%), Ribosome biogenesis in eukaryotes (27.5%), and Pyrimidine metabolism (25.0%). These pathways are mainly involved in the fundamental biological processes that maintain cell growth, proliferation, and survival. For example, three of the most enriched categories pertained to DNA processing (11 genes, “DNA replication”; 6 genes, “mismatch repair”), RNA processing (36 genes, “RNA transport”; 24 genes, “mRNA surveillance”; 32 genes, “spliceosome”; 14 genes, “RNA degradation”), and protein processing (53 genes, “Ribosome”; 10 genes, “proteasome”). We also found that some signaling pathways, such as mTOR and FoxO, were significantly enriched. Inactivation of TOR in both *Drosophila* and mammalian cells has shown that TOR controls both cell size and proliferation in early embryos (Murakami et al. 2004). FoxO1 has also been found to be an essential regulator of pluripotency in human embryonic stem cells and developing embryos (Yu et al. 2018). Therefore, we suspect that mTOR and FoxO signaling pathways were enriched because BmE is an embryonic cell line of *B. mori*. Unlike the essential genes, the growth-restricting genes were primarily enriched in lysosome and diverse metabolic pathways which are dispensable for cell survival but critical for growth control and secondary metabolism (Supplemental Fig S6B).

Subcellular localization analysis showed that the majority of the essential genes encoded proteins that localized to nucleus (38.2%), cytoplasm (34.4%), and mitochondrion (7.9%), while the rest encoded secreted proteins or proteins that localized to other cellular compartments, such as the plasma membrane, endoplasmic reticulum membrane, and mitochondrial membrane (Fig. 4B). To further investigate the importance of subcellular localization of selected genes, we analyzed the percentage of essential genes among all genes assigned to different categories of subcellular localization. We found that significantly more essential than growth-restricting genes were localized in the nucleus and mitochondrion, while a higher proportion of growth-restricting genes than essential genes were assigned to cellular compartments related to the extracellular space such as plasma membrane and secreted proteins (Fig. 4C). The observed biased cellular localizations of essential and growth-restricting genes has also been observed in mammalian cells (Yilmaz et al. 2018).

Use of *B. mori* CRISPR screens to identify genes responsible for temperature challenge

We next asked whether our BmEGCKLib library could be used as a CRISPR screening platform to identify genes that respond to environmental stimuli. As a proof-of-principle investigation, we chose temperature challenge. Each group of about 4×10^7 BmEGCKLib cells was exposed to 4˚C and 30˚C for 20 days; cells grown at 27℃ were used as control (Fig. 5A). Only a few cells survived in the 4˚C and 30˚C groups, while cells in the control group grew well. The sgRNAs of surviving cells

were analyzed for enrichment and depletion using similar methods to those used to identify the essential and growth-restricting genes. Apparently, several KEGG pathways were significantly enriched in cells challenged at 4˚C and overlapped highly with some pathways for essential genes. The overlapped pathways were categorized to RNA protein processing. However, none of the pathways pertaining to DNA processing that were enriched in essential genes were found in the 4˚C challenged group (Fig. 5B; Supplemental Table S6). We also found that genes with depleted sgRNAs in the 4˚C group were significantly enriched in the steroid biosynthesis pathway and genes with depleted sgRNAs in the 30˚C group were highly enriched in fatty acid biosynthesis and fatty acid metabolism pathways (Fig. 5B and C; Supplemental Table S6). Both steroids and fatty acids are known to be key regulators of cell membrane fluidity (Singer and Nicolson 1972; Ipsen et al. 1987). These results are consistent with their reported role in membrane fluidity (Ipsen et al. 1987) as well as their protective roles at low temperatures(Xu and Siegenthaler 1997). Comparing the proportions of enriched genes on cytomembrane, subcellular organ membranes, and other non-membrane subcellular organs, we found genes that anchored on biomembranes, especially on the cytomembrane, were likely to be associated with depleted sgRNAs in both the 4℃ and 30℃ groups (Fig. 5D). This suggested that biomembrane systems were more likely to be destroyed by extreme temperatures. We also identified more genes with enriched sgRNAs compared to those with depleted sgRNAs in the mitochondria of cells from both temperature challenge groups (Fig. 5D; Supplemental Table S6). This indicated that decreased energy metabolism may help cells survive against extreme ambient temperatures.

Use of *B. mori* CRISPR screens to identify host-pathogen interactions

Lastly, we aimed to take advantage of the hypothesis-free application of the BmEGCKLib screen platform to identify genes involved in host-pathogen interactions between BmNPV and *B. mori*. BmNPV, a typical species of Baculoviridae, is a natural *B. mori* pathogen which causes enormous economic losses in the sericulture industry every year. BmEGCKLib cells were infected with BmNPV 4 times at a high level of infection once every 2 days. The few cells surviving among ~ 4×10^7 cells infected were harvested for analysis of host-pathogen interactions. Using a p-value < 0.05 as a threshold, we identified a positive selection of 811 genes and negative selection of 809 genes, in which most highly ranked genes were represented by multiple independent sgRNAs. Many genes previously reported to play anti-BmNPV roles were found in the group of positively selected genes (Fig. 6A; Supplemental Table S7), suggesting that the BmEGCKLib can be used to identify genes involved in *B. mori*-BmNPV interactions and that the 1,614 selected genes may be truly responsible for BmNPV infection.

We then mapped the selected genes to KEGG pathways (Fig. 6B). We observed enrichment for genes involved in phagosome and notch signaling pathways, both of which are represented in genome-wide screens for influenza virus and have been demonstrated by several previous studies to be involved in the host response to influenza virus, vesicular stomatitis virus, and *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) (Volkman and Goldsmith 1985; Maxfield and Yamashiro 1987; Sun et al. 2005). To investigate further the possible mechanism of BmNPV entry into *B. mori* cells, we analyzed in detail the genes in the most enriched pathway, the phagosome pathway. We found, 5 of 12 genes were the vacuolar-type

H^+-ATPase (V-ATPase) subunits (Fig. 6C; Supplemental Table S7). As a key enzyme of endosomal acidification, V-ATPase has been demonstrated to play an important role during influenza virus entry into mammalian cells (Rossmann and Rao 2012). Previously, V-ATPase was found to exhibit a higher expression level in NPV-resistant *B. mori* strains (Lu et al. 2013). A more recent study showed that inhibition of endosome acidification by ammonium chloride treatment in *B. mori* cells results in significant inhibition of virus reproduction (Feng et al. 2018). In addition to V-ATPase, five other components of the endocytosis pathways, including TUBB, Sec61, RAB5, Rac, and PIKFYVE, were significantly enriched (Fig. 6D). These results suggest that endocytosis, especially V-ATPase mediated endosomal acidification, plays a critical role during the entry of BmNPV into *B. mori* cells.

To reveal further the genes and pathways involved in host cell defense against virus infection, we also analyzed negatively selected genes. These genes were primarily classified in the Wnt signaling and MAPK signaling pathways (Fig. 6B). Wnt is a cellular pathway related to cell cycle signaling which has also been shown to be affected by diverse human viruses, such as Epstein-bar virus(Birdwell et al. 2018), human herpesvirus(Liu et al. 2016), human papillomavirus(Bello et al. 2015), and hepatitis B virus(Yin et al. 2017). Thus, the manipulation of Wnt signaling is recognized as a critical generalized process for viral pathogenesis(Hao et al. 2015). The Wnt pathway has also been identified in several genome-wide genetics screens in human cells or *Drosophila* against Rift Valley fever virus and Sendai virus(Baril et al. 2013; Harmon et al. 2016). However, whether the Wnt pathway has antiviral activity against BmNPV or whether BmNPV modulates Wnt signaling to evade this antiviral host response remains unknown. Our results showed that a core component of the

Wnt pathway, β-catenin, was highly ranked among the negatively selected genes (Fig. 6D). This supports the supposition that the Wnt pathway may play an important role in *B. mori* for host response against BmNPV.

To validate whether the screened genes play important roles in the BmNPV infection, we randomly chose 9 genes from the top30 positively screened genes and 2 genes encoding v-ATPase subunits, and constructed 11 knockout cell lines. BmNPV infection were performed with MOI=1 in 11 cell lines, BmE cells were used as control. The fluorescent microscopy observation and flow cytometry analysis at 72 hpi both revealed that BmNPV positive cells were significantly fewer in all 11 knockout cell lines than that in control BmE cells (Fig. 7A and B; Supplemental Fig S7). Quantification of viral DNA in infected cells using qPCR showed that the DNA content of BmNPV were significantly lower in all knockout cell lines (Fig. 7C). These results indicated the crucial roles for 11 genes under BmNPV infection and the accuracies of our screens results.

Discussion

Genome-wide CRISPR-based screening technologies have greatly accelerated the functional annotation of genomes and provided a hypothesis-free, cost-effective pipeline to uncover target genes in certain contexts. Delivery of genome-wide CRISPR libraries is achieved mostly by lentiviral vectors in mammalian cells and recently by site-specific recombination in *Drosophila* cells(Shalem et al. 2014; Viswanatha et al. 2018). However, these approaches are currently inapplicable for *B. mori* and many other non-model organisms, possibly as a result of their low efficiency.

In the current study, we showed that the *piggy*Bac transposon could be used as an alternative strategy to deliver multiplexed DNA libraries. The coverage, accuracy, and sgRNA distribution of the newly constructed BmEGCKLib were sufficient to perform the screening of essential and growth-restricting genes under normal conditions, as well as context-specific screening under abiotic (such as temperature challenge) and biotic (such as host-pathogen interactions) stresses. The *piggy*Bac transposon, a mobile genetic element present in a diverse range of species, can efficiently transpose between vectors and chromosomes in nearly all organisms tested. Since *piggy*Bac transposes via a "cut and paste" mechanism, the inserted copy number per cell can be easily controlled by quantifying the transfected donor plasmid. In the present study, we observed that > 95% cells in the BmEGCKLib harbored only a single sgRNA insertion as revealed by single cell sequencing (Supplemental Fig S8). Furthermore, the *piggy*Bac transposon has the largest reported cargo capacity, up to 207 kb (Li et al. 2013), for these kinds of vectors, allowing co-delivery of other functional elements with a CRISPR library in complex screens. Thus, we believe that the *piggy*Bac strategy developed in our study possesses several advantages over lentiviral or recombinase based strategies and will be versatile for diverse cell types and organisms.

B. mori is among the most important economically beneficial insects and a powerful model system for Lepidoptera because of its finely decoded genomic sequence (Xia et al. 2008), rich genetic resources(Goldsmith et al. 2005) and multiple available tools for genetic manipulation (Tamura et al. 2000; Takasu et al. 2010; Ma et al. 2012; Wang et al. 2013). After the completion of *B. mori* genome sequencing, the systematical functional investigation of all *B. mori* genes has been a major

challenge. Transgenic and genome editing technologies have been established in an effort to achieve this goal. However, the function of over 80% of the *B. mori* genes remains unclear, partially because of the low efficiency of hypothesis-driven research strategies, which heavily rely on previous knowledge and well-grounded hypotheses. In the current study, we presented the first hypothesis-free genetic screen approach to rapidly investigate the function of *B. mori* genes on a genome-wide scale. A total of 1,006 genes essential for cell viability, 838 genes that restricted cell growth, 3,013 genes responsible for ambient temperature change, and 1,614 genes involved in BmNPV infection were identified using our genetic screen approach.

The identification and analysis of genes essential for cell viability in *B. mori* also implies some interesting aspects of cell biology in general. First, we found that the essential genes in *B. mori* shared considerable overlap with all seven eukaryotic species compared, whereas very limited overlap existed with the prokaryote *B. thuringiensis*, indicating an evolutionary conservation of basic biological processes for eukaryotic cell survival. Second, essential genes primarily encoded proteins involved in fundamental biological processes such as DNA, RNA, and protein processing, and overall were evenly distributed across all chromosomes, suggesting that the random chromosomal distribution of essential genes may have a protective role in the core biological processes. Third, these results indicated that essential genes leading to cell survival under normal conditions were located primarily in the nucleus, cytoplasm and mitochondrion, while genes located on the cell membrane and extracellular matrix were much lower in the numbers detected by our screen. These has also been observed in mammalian cells (Yilmaz et al. 2018), suggests that: (1) essential genes may have different roles in regulating cell growth; (2) essential genes

may have different levels of functionally redundant genes in the different cellular compartments; and (3) most proteins secreted outside cells are not crucial to cell viability and their loss may even promote cell growth.

Since the ambient temperature has a great impact on the ectotherm like insect. We performed the CRISPR screening to identify genes that respond to environmental stimuli in BmE cells. Our screening showed, DNA maintenance and replication were the core events of the cells and their disruption probably threatens cell survival, however two other fundamental biological processes crucial to the maintenance of cell growth in normal conditions, RNA processing and protein processing were not. In addition, gene function clustering revealed steroids and fatty acids played important roles in membrane fluidity and protected cells from extreme ambient temperatures. Additionally, subcellular localization analyzing suggested that biomembrane systems were very fragile at extreme temperatures, and the decreased of energy metabolism may help cells survive against extreme ambient temperatures. Since the global climate changes has become the most serious environmental issue, the extremely temperatures appear more frequently. Therefore, the study of the genes responsible for temperature challenge in BmE cells is helpful to understand the effects of climate changes on insects. Owing to insects are the main agricultural pests, our research also provided enlightenment for pest control under the global climate changes.

BmNPV is the most important pathogen of *B. mori*. The research of the interaction mechanism between BmNPV and host *B. mori* was important not only for sericulture but also for pest control. Taken together, our BmEGCKLib screening confirmed previously described biological processes, such as phagosome and notch signaling pathways, are involved in the interaction between BmNPV and host *B. mori*

cells and uncovered some novel processes, such as dorsal-ventral axis formation and the Wnt pathways associated with the infection of BmNPV. In addition, we identified over 493 significantly selected genes whose functions are currently unknown due to their limited homology to identified genes of model organisms. We propose that some of the selected genes may also be involved in the interaction between BmNPV and host *B. mori* cells and may provide a reliable resource for defining new targets for antiviral drugs and breeding silkworm strains genetically resistant to this widespread pathogen. It's the first time to uncover host-pathogen interactions in insect through the genome-scale knockout library. Our strategy can provide new ideas for insect antiviral research.

We developed a functional screening platform in *B. mori* using *piggy*Bac delivered CRISPR libraries. We showed that CRISPR screening is a powerful tool for the rapid investigation of gene functions on a large scale and may serve as a powerful resource for investigating long-standing questions in *B. mori* and entomology. The value of our platform was demonstrated by the discovery and highlighting of a large set of new genes and pathways that participate in cell variability, cell growth, ambient temperature stimuli, and BmNPV infection. We believe that further validation and investigation of the candidate target genes revealed by our screening will shed more light on the molecular mechanisms of host response to ambient temperature stimuli and BmNPV infection. More importantly, the present study lays the ground for further genetic screens against diverse biological processes.

Methods

447 **Design and construction of the vectors**

448 The *piggy*Bac transposase expression vector, A3-helper was from stocks stored in
449 our laboratory. The CRISPR library delivery vector, pB-CRISPR, was constructed
450 with the process available in the Supplemental Materials. The target vectors used to
451 test the knockout efficiency of pB-CRISPR, was constructed using library
452 construction methods. The EGFP expression cassette was delivered by *Minos*
453 transposons, the vector named PUC57-Mi-puro-EGFP. *Minos* transposase gene
454 expression cassette was synthesized and inserted into pUC57-T-simple and the
455 plasmid named Mi-helper. The details were available in the Supplemental Materials.
456 The whole sequences were available in Supplemental Materials.

457 **sgRNA library design, synthesis and construction**

458 The sgRNAs were designed using the CasFinder method (John et al. 2014). All
459 sgRNAs selected had a 5´G added to improve the U6 transcription efficiency. The
460 library of sgRNAs were encoded within 70-nt oligonucleotides and synthesized on the
461 94K arrays using the services of BGI. *U6* promoter, multiple sgRNAs and sgRNA
462 scaffolds were linked together using overlap PCR, then cloned into the AscI/NheI site
463 of pB-CRISPR vector to construct the CRISPR knockout plasmid library
464 (pB-CRISPR library). The details methods were available in the Supplemental
465 Materials.

466 **Genome-scale screening in *B. mori***

467 For essential or growth-restricting genes screening, three parts of 4×10^7
468 BmEGCKLib cells were harvested at three time points. For the screening of genes
469 responsible for temperature challenging, three sets of 4×10^7 cells of the

BmEGCKLib in complete medium were exposed for 20 days to different temperatures. For the screening to identify genes involved in host-pathogen interactions, 4×10^7 cells of the BmEGCKLib were supplemented with BmNPV (with an EGFP tag, stored in our laboratory) for four rounds of infection. The genomic DNA was extracted from each group and the sgRNA distribution analyzed by second-generation sequencing (Mega Genomic, China). The details methods were available in the Supplemental Materials.

Pooled sgRNA sequences and data analysis

All of the genomic DNA extracted from the cell libraries was PCR amplified. The PCR products were gel-purified and sequenced by Illumina. The raw data were uploaded to a server in our laboratory and filtered. The paired sequences were then spliced using flash software. Then the number of reads for each sgRNA was calculated using the Bowtie 2 (Langmead and Salzberg 2012) and performed the next analysis using MAGeCK. The details methods were available in the Supplemental Materials.

The analysis of EGFP knockout efficiency of pB-CRISPR

To test the knockout efficiency of pB-CRISPR, three sgRNAs targeting EGFP were constructed into pB-CRISPR, and two NS (nonspecific) sgRNAs were used as controls (Supplemental Fig S2; Supplemental Table S1). Each vector was co-transfected with A3-Helper into BmE-EGFP cell line. After two months of selection with zeocin, all the samples were performed with flow cytometry analysis, fluorescence imaging, and Sanger sequencing. The details were available in the Methods section of the Supplemental Materials.

Data access:

The raw data generated in this study have been submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA602216.

Acknowledgment:

The authors would thank Prof. Marian R. Goldsmith for her valuable advice and help in the preparation of this manuscript. This work was supported by grants from the National Natural Science Foundation of China (No. 31802011, 31530071), and Chongqing Research program of basic Research and Frontier Technology (No. cstc2017jcyjAX0349, cstc2018jcyjAX0471). None of these funding plays roles in the design of the study, and collection, analysis, and interpretation of data and in writing the manuscript.

Author Contributions:

JC and SM designed the experiments. JC and KY performed most of the experiments with help from RW, YL, RS, and XW. RW performed the functional validation experiments during the revision process. TZ and XC performed the bioinformatics analysis. JC, and SM wrote the manuscript with help from QX. All authors read and approved the final manuscript.

Disclosure declaration

The authors declare that they have no competing interests

References

Baril M, Es-Saad S, Chatel-Chaix L, Fink K, Pham T, Raymond VA, Audette K, Guenier AS, Duchaine J, Servant M et al. 2013. Genome-wide RNAi Screen Reveals a New Role of a WNT/CTNNB1 Signaling Pathway as Negative Regulator of Virus-induced Innate Immune Responses. *Plos Pathog* **9:** (6) (2013).

Bassett AR, Kong LS, Liu JL. 2015. A Genome-Wide CRISPR Library for High-Throughput Genetic Screening in *Drosophila* Cells. *J Genet Genomics* **42:** 301-309.

Bello JOM, Nieva LO, Paredes AC, Gonzalez AMF, Zavaleta LR, Lizano M. 2015. Regulation of the Wnt/beta-Catenin Signaling Pathway by Human Papillomavirus E6 and E7 Oncoproteins. *Viruses-Basel* **7:** 4734-4755.

Birdwell CE, Prasai K, Dykes S, Jia Y, Munroe TGC, Bienkowska-Haba M, Scott RS. 2018. Epstein-Barr virus stably confers an invasive phenotype to epithelial cells through reprogramming of the WNT pathway. *Oncotarget* **9:** 10417-10435.

Boutros M, Kiger AA, Armknecht S, Kerr K, Hild M, Koch B, Haas SA, Paro R, Perrimon N, Consortium HFA. 2004. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303:** 832-835.

Carette JE, Guimaraes CP, Varadarajan M, Park AS, Wuethrich I, Godarova A, Kotecki M, Cochran BH, Spooner E, Ploegh HL et al. 2009. Haploid Genetic Screens in Human Cells Identify Host Factors Used by Pathogens. *Science* **326:** 1231-1235.

Chen SD, Sanjana NE, Zheng KJ, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissleder R et al. 2015. Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. *Cell* **160:** 1246-1260.

Feng M, Zhang JJ, Xu WF, Wang HP, Kong XS, Wu XF. 2018. *Bombyx mori* nucleopolyhedrovirus utilizes a clathrin and dynamin dependent endocytosis entry pathway into BmN cells. *Virus Res* **253:** 12-19.

Goldsmith MR, Shimada T, Abe H. 2005. The genetics and genomics of the silkworm, *Bombyx mori*. *Annu Rev Entomol* **50:** 71-100.

Hao HP, Wen LB, Li JR, Wang Y, Ni B, Wang R, Wang X, Sun MX, Fan HJ, Mao X. 2015. LiCl inhibits PRRSV infection by enhancing Wnt/beta-catenin pathway and suppressing inflammatory responses. *Antivir Res* **117:** 99-109.

Harmon B, Bird SW, Schudel BR, Hatch AV, Rasley A, Negrete OA. 2016. A Genome-Wide RNA Interference Screen Identifies a Role for Wnt/beta-Catenin Signaling during Rift Valley Fever Virus Infection. *J Virol* **90:** 7084-7097.

Hart T, Brown KR, Sircoulomb F, Rottapel R, Moffat J. 2014. Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Mol Syst Biol* **10:** (7), 733.

Hou PP, Wu C, Wang YC, Qi R, Bhavanasi D, Zuo ZX, Dos Santos C, Chen SL, Chen Y, Zheng H et al. 2017. A Genome-Wide CRISPR Screen Identifies Genes Critical for Resistance to FLT3 Inhibitor AC220. *Cancer Res* **77:** 4402-4413.

559 Ipsen JH, Karlstrom G, Mouritsen OG, Wennerstrom H, Zuckermann MJ. 1987.
560 Phase-Equilibria in the Phosphatidylcholine-Cholesterol System. *Biochim*
561 *Biophys Acta* **905:** 162-172.
562 John Aach, Prashant Mali, George M. Church. 2014. CasFinder: Flexible algorithm for
563 identifying specific Cas9 targets in genomes. bioRxiv doi: 10.1101/005074.
564 Kim HS, Lee K, Kim SJ, Cho S, Shin HJ, Kim C, Kim JS. 2018. Arrayed CRISPR screen
565 with image-based assay reliably uncovers host genes required for
566 coxsackievirus infection. *Genome Res* **28:** 859-868.
567 Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, Ishihara G, Kawaoka S,
568 Sugano S, Shimada T et al. 2014. A single female-specific piRNA is the primary
569 determiner of sex in the silkworm. *Nature* **509:** (7502), 633.
570 Koike-Yusa H, Li YL, Tan EP, Velasco-Herrera MD, Yusa K. 2014. Genome-wide
571 recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide
572 RNA library. *Nat Biotechnol* **32:** 267-273.
573 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat*
574 *Methods* **9:** 357-U354.
575 Li RB, Zhuang Y, Han M, Xu T, Wu XH. 2013. *piggy*Bac as a high-capacity transgenesis
576 and gene-therapy vector in human cells and mice. *Dis Model Mech* **6:** 828-833.
577 Li W, Xu H, Xiao TF, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS.
578 2014. MAGeCK enables robust identification of essential genes from
579 genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* **15:** 554.
580 Liu X, Gallay C, Kjos M, Domenech A, Slager J, van Kessel SP, Knoops K, Sorg RA,
581 Zhang JR, Veening JW. 2017. High-throughput CRISPRi phenotyping identifies
582 new essential genes in *Streptococcus pneumoniae*. *Mol Syst Biol* **13:** (5), 931.
583 Liu Y, Hancock M, Workman A, Doster A, Jones C. 2016. beta-Catenin, a Transcription
584 Factor Activated by Canonical Wnt Signaling, Is Expressed in Sensory Neurons
585 of Calves Latently Infected with Bovine Herpesvirus 1. *J Virol* **90:** 3148-3159.
586 Lu P, Xia HC, Gao L, Pan Y, Wang Y, Cheng X, Lu HG, Lin F, Chen L, Yao Q et al. 2013.
587 V-ATPase Is Involved in Silkworm Defense Response against *Bombyx mori*
588 Nucleopolyhedrovirus. *Plos One* **8:** (6), e64962.
589 Ma HM, Dang Y, Wu YG, Jia GX, Anaya E, Zhang JL, Abraham S, Choi JG, Shi GJ, Qi L et
590 al. 2015. A CRISPR-Based Screen Identifies Genes Essential for
591 West-Nile-Virus-Induced Cell Death. *Cell Rep* **12:** 673-683.
592 Ma SY, Chang JS, Wang XG, Liu YY, Zhang JD, Lu W, Gao J, Shi R, Zhao P, Xia QY.
593 2014. CRISPR/Cas9 mediated multiplex genome editing and heritable
594 mutagenesis of BmKu70 in *Bombyx mori*. *Sci Rep-Uk* **4:** 4489.
595 Ma SY, Liu Y, Liu YY, Chang JS, Zhang T, Wang XG, Shi R, Lu W, Xia XJ, Zhao P et al.
596 2017. An integrated CRISPR *Bombyx mori* genome editing system with
597 improved efficiency and expanded target sites. *Insect Biochem Molec* **83:**
598 13-20.
599 Ma SY, Zhang SL, Wang F, Liu Y, Liu YY, Xu HF, Liu C, Lin Y, Zhao P, Xia QY. 2012.
600 Highly Efficient and Specific Genome Editing in Silkworm Using Custom TALENs.
601 *Plos One* **7:** (9), e45035.
602 Maxfield FR, Yamashiro DJ. 1987. Endosome acidification and the pathways of
603 receptor-mediated endocytosis. *Advances in experimental medicine and*
604 *biology* **225:** 189-198.
605 Mohr SE, Smith JA, Shamu CE, Neumuller RA, Perrimon N. 2014. RNAi screening
606 comes of age: improved techniques and complementary approaches. *Nat Rev*
607 *Mol Cell Bio* **15:** 591-600.
608 Murakami M, Ichisaka T, Maeda M, Oshiro N, Hara K, Edenhofer F, Kiyama H,
609 Yonezawa K, Yamanaka S. 2004. mTOR is essential for growth and

proliferation in early mouse embryos and embryonic stem cells. *Mol Cell Biol* **24:** 6710-6718.

Omenetto FG, Kaplan DL. 2010. New Opportunities for an Ancient Material. *Science* **329:** 528-531.

Parnas O, Jovanovic M, Eisenhaure TM, Herbst RH, Dixit A, Ye CJ, Przybylski D, Platt RJ, Tirosh I, Sanjana NE et al. 2015. A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell* **162:** 675-686.

Patel SJ, Sanjana NE, Kishton RJ, Eidizadeh A, Vodnala SK, Cam M, Gartner JJ, Jia L, Steinberg SM, Yamamoto TN et al. 2017. Identification of essential genes for cancer immunotherapy. *Nature* **548:** (7669), 537-542.

Rossmann MG, Rao VB. 2012. Viruses: Sophisticated Biological Machines. *Viral Molecular Machines* **726:** 1-3.

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG et al. 2014. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* **343:** 84-87.

Singer SJ, Nicolson GL. 1972. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science* **175:** (4023), 720-31.

Song CQ, Li YX, Mou HW, Moore J, Park A, Pomyen Y, Hough S, Kennedy Z, Fischer A, Yin H et al. 2017. Genome-Wide CRISPR Screen Identifies Regulators of Mitogen-Activated Protein Kinase as Suppressors of Liver Tumors in Mice. *Gastroenterology* **152:** (5), 1161-1173.e1.

Sun XJ, Yau VK, Briggs BJ, Whittaker GR. 2005. Role of clathrin-mediated endocytosis during vesicular stomatitis virus entry into host cells. *Virology* **338:** 53-60.

Takasu Y, Kobayashi I, Beumer K, Uchino K, Sezutsu H, Sajwan S, Carroll D, Tamura T, Zurovec M. 2010. Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection. *Insect Biochem Molec* **40:** 759-765.

Tamura T, Thibert C, Royer C, Kanda T, Abraham E, Kamba M, Komoto N, Thomas JL, Mauchamp B, Chavancy G et al. 2000. Germline transformation of the silkworm *Bombyx mori* L-using a *piggy*Bac transposon-derived vector. *Nat Biotechnol* **18:** 81-84.

Viswanatha R, Li ZC, Hu YH, Perrimon N. 2018. Pooled genome-wide CRISPR screening for basal and context-specific fitness gene essentiality in *Drosophila* cells. *Elife* **7:** e36333.

Volkman LE, Goldsmith PA. 1985. Mechanism of Neutralization of Budded Autographa-Californica Nuclear Polyhedrosis-Virus by a Monoclonal-Antibody - Inhibition of Entry by Adsorptive Endocytosis. *Virology* **143:** 185-195.

Wang T, Wei JJ, Sabatini DM, Lander ES. 2014. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* **343:** 80-84.

Wang TM, Guan CG, Guo JH, Liu B, Wu YA, Xie Z, Zhang C, Xing XH. 2018. Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat Commun* **9:** (1), 2475.

Wang YQ, Li ZQ, Xu J, Zeng BS, Ling L, You L, Chen YZ, Huang YP, Tan AJ. 2013. The CRISPR/Cas System mediates efficient genome engineering in *Bombyx mori*. *Cell Res* **23:** 1414-1416.

Xia QY, Guo YR, Zhang Z, Li D, Xuan ZL, Li Z, Dai FY, Li YR, Cheng DJ, Li RQ et al. 2009. Complete Resequencing of 40 Genomes Reveals Domestication Events and Genes in Silkworm (*Bombyx*). *Science* **326:** 433-436.

Xia QY, Li S, Feng QL. 2014. Advances in Silkworm Studies Accelerated by the Genome Sequencing of *Bombyx mori*. *Annual Review of Entomology, Vol 59, 2014* **59:** 513-536.

Xia QY Wang J Zhou ZY Li RQ Fan W Cheng DJ Cheng TC Qin JJ Duan J Xu HF et al. 2008. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochem Molec* **38:** 1036-1045.

Xu CL, Qi XL, Du XG, Zou HY, Gao F, Feng T, Lu HX, Li SL, An XM, Zhang LJ et al. 2017. *piggy*Bac mediates efficient in vivo CRISPR library screening for tumorigenesis in mice. *P Natl Acad Sci USA* **114:** 722-727.

Xu YN, Siegenthaler PA. 1997. Low temperature treatments induce an increase in the relative content of both linolenic and Delta(3)-trans-hexadecenoic acids in thylakoid membrane phosphatidylglycerol of squash cotyledons. *Plant Cell Physiol* **38:** 611-618.

Yilmaz A, Peretz M, Aharony A, Sagi I, Benvenisty N. 2018. Defining essential genes for human pluripotent stem cells by CRISPR-Cas9 screening in haploid cells. *Nat Cell Biol* **20:** (5), 610-619.

Yin YQ, Li F, Li SL, Cai JJ, Shi J, Jiang YH. 2017. TLR4 Influences Hepatitis B Virus Related Hepatocellular Carcinoma by Regulating the Wnt/beta-Catenin Pathway. *Cell Physiol Biochem* **42:** 469-479.

Yu F, Wei R, Yang J, Liu JL, Yang K, Wang HN, Mu YM, Hong TP. 2018. FoxO1 inhibition promotes differentiation of human embryonic stem cells into insulin producing cells. *Exp Cell Res* **362:** 227-234.

Figure Legends

Fig. 1 The *piggy*Bac library delivery method for gene knockout in *B. mori* cells.

(*A*) Schematic of the *piggy*Bac library delivery vector (pB-CRISPR). ITR, inverted terminal repeats of *piggy*Bac; IE2, *IE2* promoter; Zeocin, zeocin selection marker gene; Ser1, Sericin 1 poly(A); gRNA, sgRNA and Scaffolds; U6, polymerase III *U6*- promoter; Hsp70, *hsp70* promoter; SpCas9, *B. mori* codon-optimized SpCas9 coding sequence; SV40, SV40 poly(A). (*B*) Flow cytometry analysis (upper) and statistical analysis (lower) of BmE-Mi-EGFP cells transduced with pB-CRISPRs (mock, no vector; NS-1, pB-CRISPR-NS-1; NS-2, pB-CRISPR-NS-2; EGFP-1, pB-CRISPR- EGFP-1; EGFP-2, pB-CRISPR-EGFP-2; EGFP-3, pB-CRISPR-EGFP-3). (*C*) Fluorescent images of BmE-Mi-EGFP knockout experiments. (*D*) Sanger sequence analysis of the ratio of indels for EGFP-1 (left), EGFP-2 (middle), and EGFP-3 (right) target regions of BmE-Mi-EGFP cells.

Fig. 2 Design and construction of the CRISPR sgRNA library for *B. mori* cells.

(*A*) Flowchart for constructing the CRISPR sgRNA library for *B. mori* cells. (*B*) Design for all sgRNAs in all *B. mori* gene exons (lower). The sgRNAs selected for the subsequent experiment (upper); see supplementary Table 2. (*C*) Second-generation sequence analysis of the sgRNA distribution for the pB-CRISPR library (orange) and BmEGCKLib (green). (*D*) Pie chart showing the percent distribution of the sgRNAs. (*E*) Overlap in sgRNA distribution between the pB-CRISPR library and BmEGCKLib. (*F*) Correlation analysis between the pB-CRISPR library and BmEGCKLib.

Fig. 3 Screening of essential and growth-restricting genes in BmE cells.

(*A*) and (*B*) Changes in sgRNAs contained in BmEGCKLib over time. Approximately 4×10^7 cells of BmEGCKLib were harvested at each of three time points: immediately (BmE-Lib1), one month (BmE-Lib2), and two months BmE-Lib3) after completion of zeocin selection. Overall the sgRNAs were gradually depleted but some of them were enriched over time. (*C*) Percentages of essential genes and growth-restricting genes in BmE cells. (*D*) Expression levels among all genes, essential genes, and growth-restricting genes in BmE cells and silk gland cells. (*E*) The overlap of essential genes among *Bombyx mori* and diverse eukaryotic species (*Homo sapiens, Mus musculus, Drosophila melanogaster, Saccharomyces cerevisiae, Danio rerio, Caenorhabditis elegans, and Arabidopsis thaliana*), and a prokaryotic species (*Bacillus thuringiensis*). Asterisks: the intersection of 1006 essential genes with *B. mori* orthologs of essential genes in the eight species; box plots: the intersection of 1000 random genes with *B. mori* orthologs of essential genes in the eight species.

Fig. 4 Analysis of essential genes for BmE cells.

(*A*) Top 30 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment of essential genes. (*B*) Distribution of essential genes assigned to subcellular localization. (*C*) Percent distribution of essential and growth-restricting genes assigned to subcellular localization within the total number of genes.

719 **Fig. 5** Screening of genes responsible for temperature challenge.

720 (*A*) Flowchart for the screening of genes involved in a response to temperature challenges. (*B*)
721 and **c** KEGG pathway enrichment for the genes in cells cultured at 4˚C (*B*) and 30˚C (*C*). (*D*)
722 Percentage of major types of subcellular localization genes within the total number of genes
723 at 4˚C and 30˚C (Mit, mitochondrion; ERM, endoplasmic reticulum membrane; GOLM,
724 Golgi apparatus membrane; PM, plasma membrane).

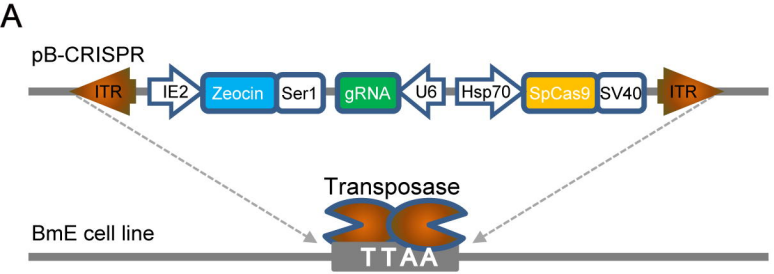
725 **Fig. 6** Screening of genes involved in the interaction between BmE cells and *B. mori*
726 nucleopolyhedrovirus (BmNPV).

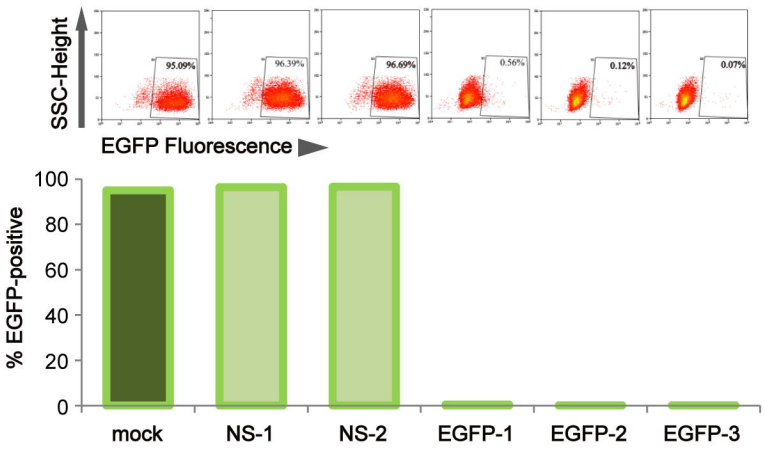
727 (*A*) Plot of genes ranked by negative score. The portion of genes that associated with virus
728 resistance are highlighted. (*B*) Genes for BmE-BmNPV interactions uncovered by KEGG
729 pathway analysis. (*C*) sgRNA score of genes in the phagosome pathway. (*D*) Schematic of
730 *Bombyx mori* cell interactions with BmNPV.

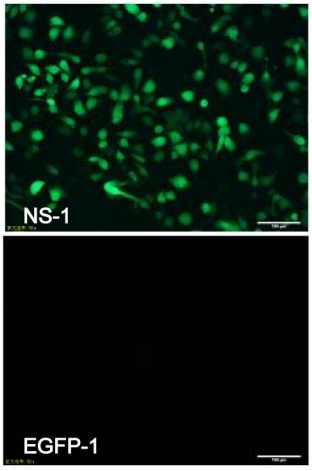
731 **Fig. 7** Functional analysis of genes involved in *B. mori*-BmNPV interactions.

732 (*A*) Fluorescent images analysis (upper) and flow cytometry analysis (lower) of the BmE
733 knockout cells infected with BmNPV at an MOI of 1. (*B* and *C*) Statistical flow cytometry
734 analysis (*B*) and relative BmNPV DNA level (*C*) in 11 BmE knockout cells infected with
735 BmNPV at an MOI of 1. The statistical significant differences were determined by Student's
736 *t*-test (* $p < 0.05$, ** $p < 0.01$, *** $p<0.001$).

Figure 1


A
pB-CRISPR
ITR
IE2
Zeocin
Ser1
gRNA
U6
Hsp70
SpCas9
SV40
ITR
Transposase
BmE cell line
TTAA


B
SSC-Height
EGFP Fluorescence
95.09%
96.39%
96.69%
0.56%
0.12%
0.07%
% EGFP-positive
100
80
60
40
20
0
mock
NS-1
NS-2
EGFP-1
EGFP-2
EGFP-3


C
NS-1
EGFP-1

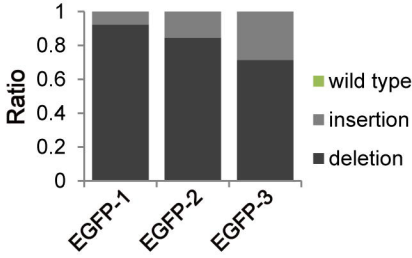

D
Ratio
1
0.8
0.6
0.4
0.2
0
wild type
insertion
deletion
EGFP-1
EGFP-2
EGFP-3

Figure 2

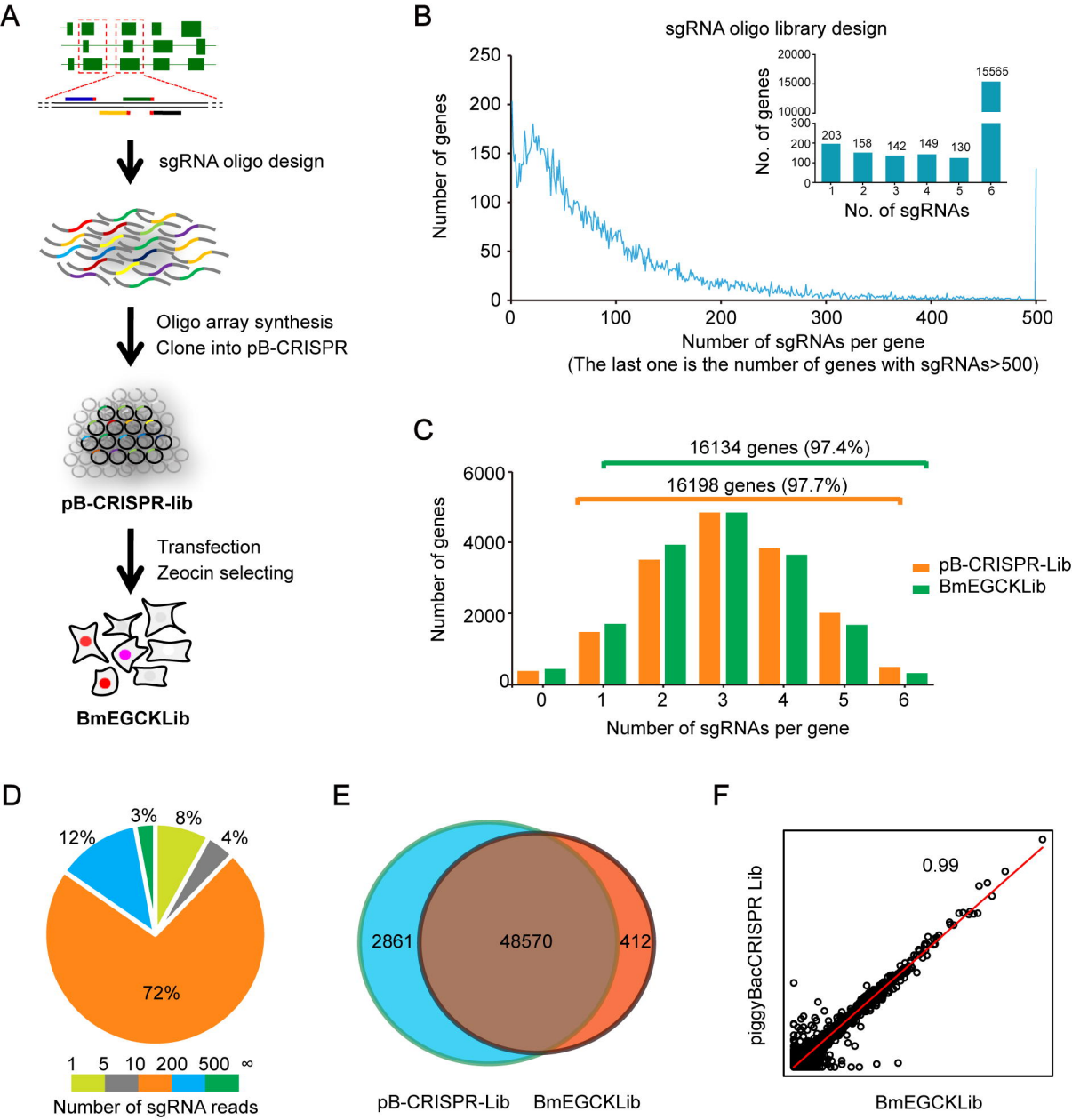

A
sgRNA oligo design
Oligo array synthesis
Clone into pB-CRISPR
pB-CRISPR-lib
Transfection
Zeocin selecting
BmEGCKLib
B
sgRNA oligo library design
Number of genes
Number of sgRNAs per gene
(The last one is the number of genes with sgRNAs>500)
No. of genes
No. of sgRNAs
203
158
142
149
130
15565
C
16134 genes (97.4%)
16198 genes (97.7%)
Number of genes
Number of sgRNAs per gene
pB-CRISPR-Lib
BmEGCKLib
D
12%
3%
8%
4%
72%
1 5 10 200 500 ∞
Number of sgRNA reads
E
2861
48570
412
pB-CRISPR-Lib
BmEGCKLib
F
0.99
piggyBacCRISPR Lib
BmEGCKLib

Figure 3

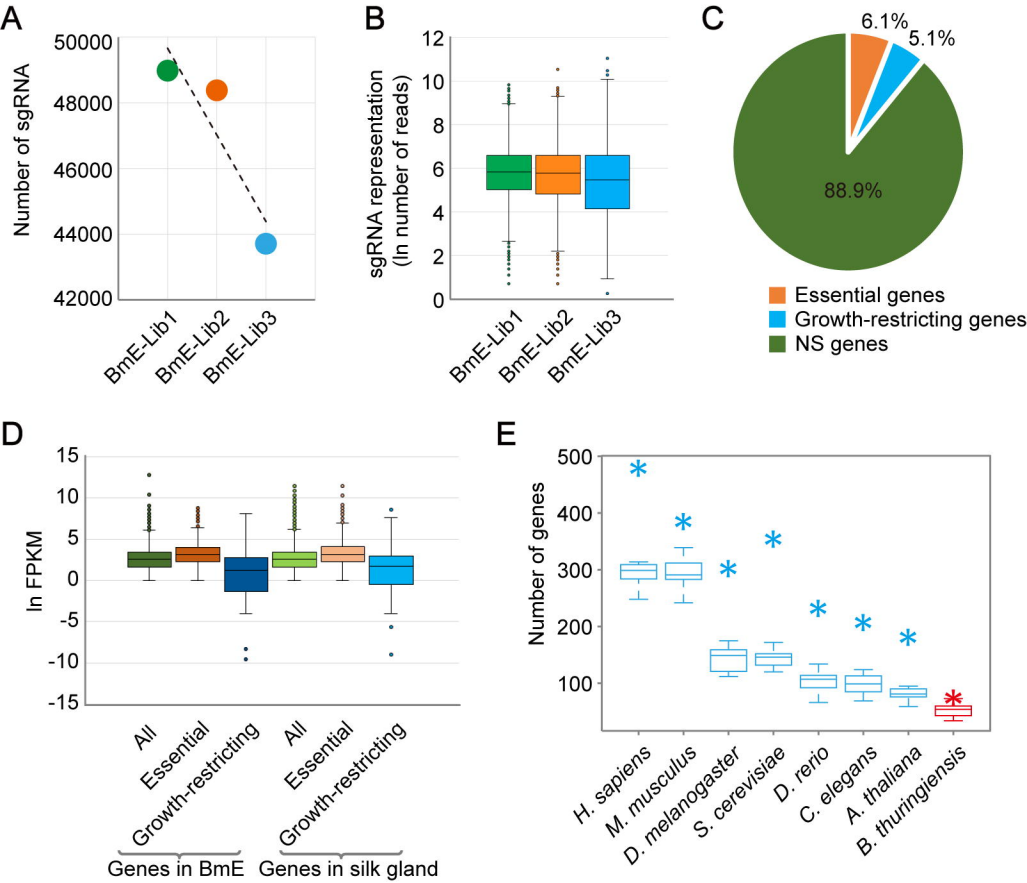

A
Number of sgRNA
50000
48000
46000
44000
42000
BmE-Lib1
BmE-Lib2
BmE-Lib3
B
sgRNA representation
(ln number of reads)
12
10
8
6
4
2
0
BmE-Lib1
BmE-Lib2
BmE-Lib3
C
6.1%
5.1%
88.9%
Essential genes
Growth-restricting genes
NS genes
D
ln FPKM
15
10
5
0
-5
-10
-15
All
Essential
Growth-restricting
All
Essential
Growth-restricting
Genes in BmE
Genes in silk gland
E
Number of genes
500
400
300
200
100
H. sapiens
M. musculus
D. melanogaster
S. cerevisiae
D. rerio
C. elegans
A. thaliana
B. thuringiensis

Figure 4

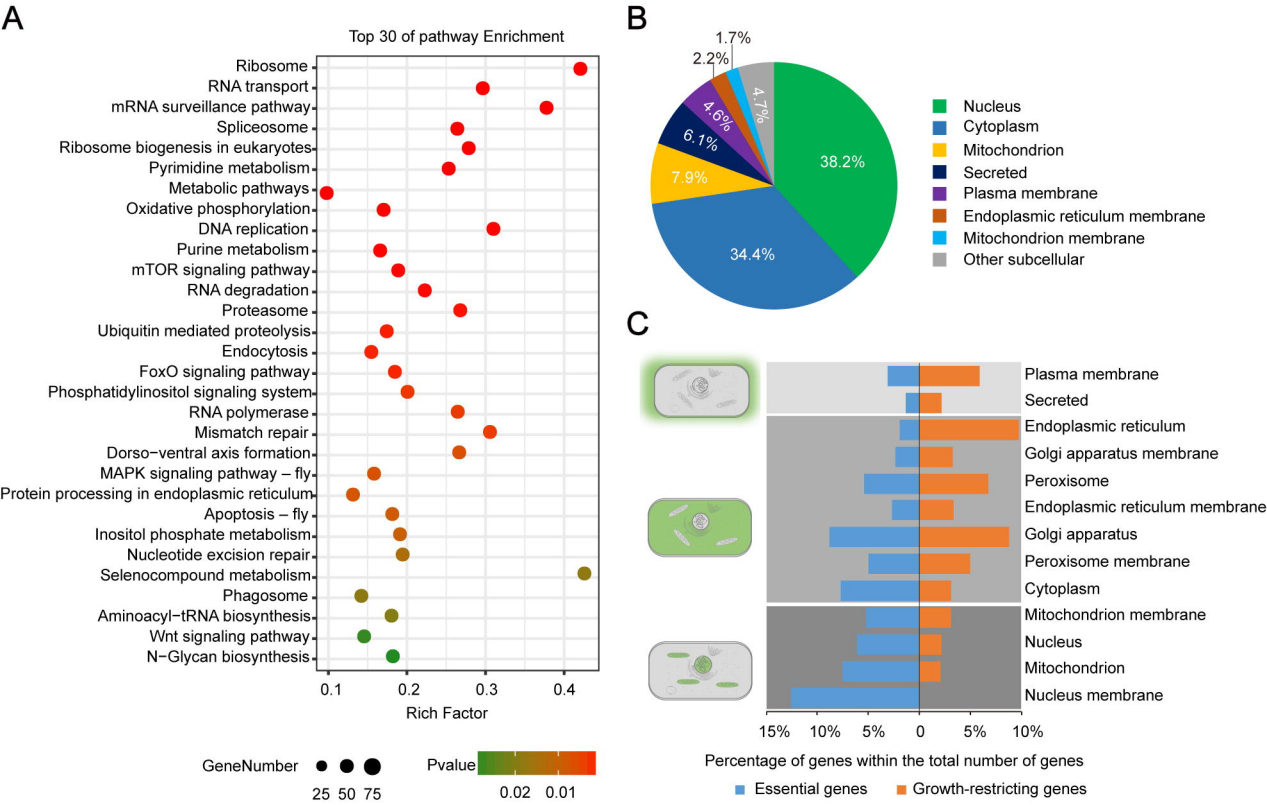
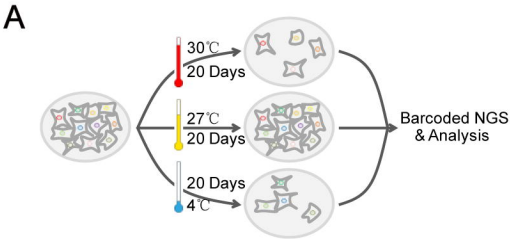
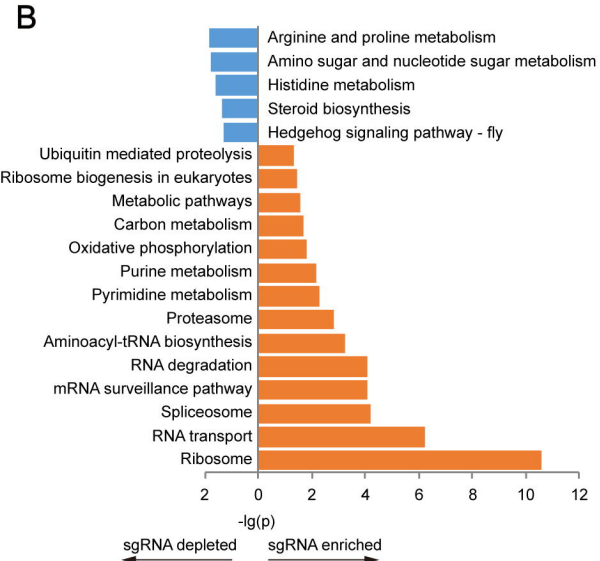
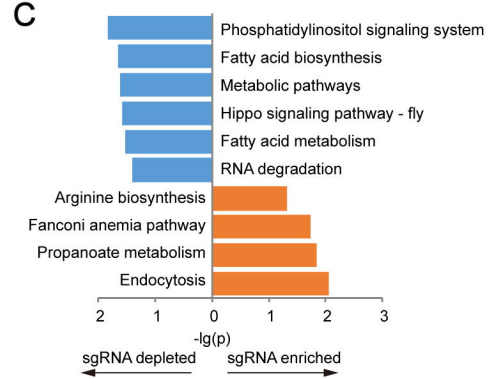

A
Top 30 of pathway Enrichment
Ribosome
RNA transport
mRNA surveillance pathway
Spliceosome
Ribosome biogenesis in eukaryotes
Pyrimidine metabolism
Metabolic pathways
Oxidative phosphorylation
DNA replication
Purine metabolism
mTOR signaling pathway
RNA degradation
Proteasome
Ubiquitin mediated proteolysis
Endocytosis
FoxO signaling pathway
Phosphatidylinositol signaling system
RNA polymerase
Mismatch repair
Dorso−ventral axis formation
MAPK signaling pathway – fly
Protein processing in endoplasmic reticulum
Apoptosis – fly
Inositol phosphate metabolism
Nucleotide excision repair
Selenocompound metabolism
Phagosome
Aminoacyl−tRNA biosynthesis
Wnt signaling pathway
N−Glycan biosynthesis
0.1
0.2
0.3
0.4
Rich Factor
GeneNumber
25 50 75
Pvalue
0.02 0.01
B
1.7%
2.2%
4.7%
4.6%
6.1%
7.9%
38.2%
34.4%
Nucleus
Cytoplasm
Mitochondrion
Secreted
Plasma membrane
Endoplasmic reticulum membrane
Mitochondrion membrane
Other subcellular
C
Plasma membrane
Secreted
Endoplasmic reticulum
Golgi apparatus membrane
Peroxisome
Endoplasmic reticulum membrane
Golgi apparatus
Peroxisome membrane
Cytoplasm
Mitochondrion membrane
Nucleus
Mitochondrion
Nucleus membrane
15% 10% 5% 0 5% 10%
Percentage of genes within the total number of genes
Essential genes
Growth-restricting genes

Figure 5


A
30℃
20 Days
27℃
20 Days
20 Days
4℃
Barcoded NGS
& Analysis


B
Arginine and proline metabolism
Amino sugar and nucleotide sugar metabolism
Histidine metabolism
Steroid biosynthesis
Hedgehog signaling pathway - fly
Ubiquitin mediated proteolysis
Ribosome biogenesis in eukaryotes
Metabolic pathways
Carbon metabolism
Oxidative phosphorylation
Purine metabolism
Pyrimidine metabolism
Proteasome
Aminoacyl-tRNA biosynthesis
RNA degradation
mRNA surveillance pathway
Spliceosome
RNA transport
Ribosome
2 0 2 4 6 8 10 12
-lg(p)
sgRNA depleted
sgRNA enriched


C
Phosphatidylinositol signaling system
Fatty acid biosynthesis
Metabolic pathways
Hippo signaling pathway - fly
Fatty acid metabolism
RNA degradation
Arginine biosynthesis
Fanconi anemia pathway
Propanoate metabolism
Endocytosis
2 1 0 1 2 3
-lg(p)
sgRNA depleted
sgRNA enriched

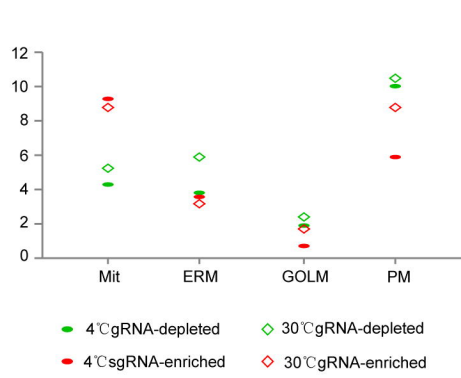

D
12
10
8
6
4
2
0
Mit
ERM
GOLM
PM
4℃gRNA-depleted
30℃gRNA-depleted
4℃sgRNA-enriched
30℃gRNA-enriched

Figure 6

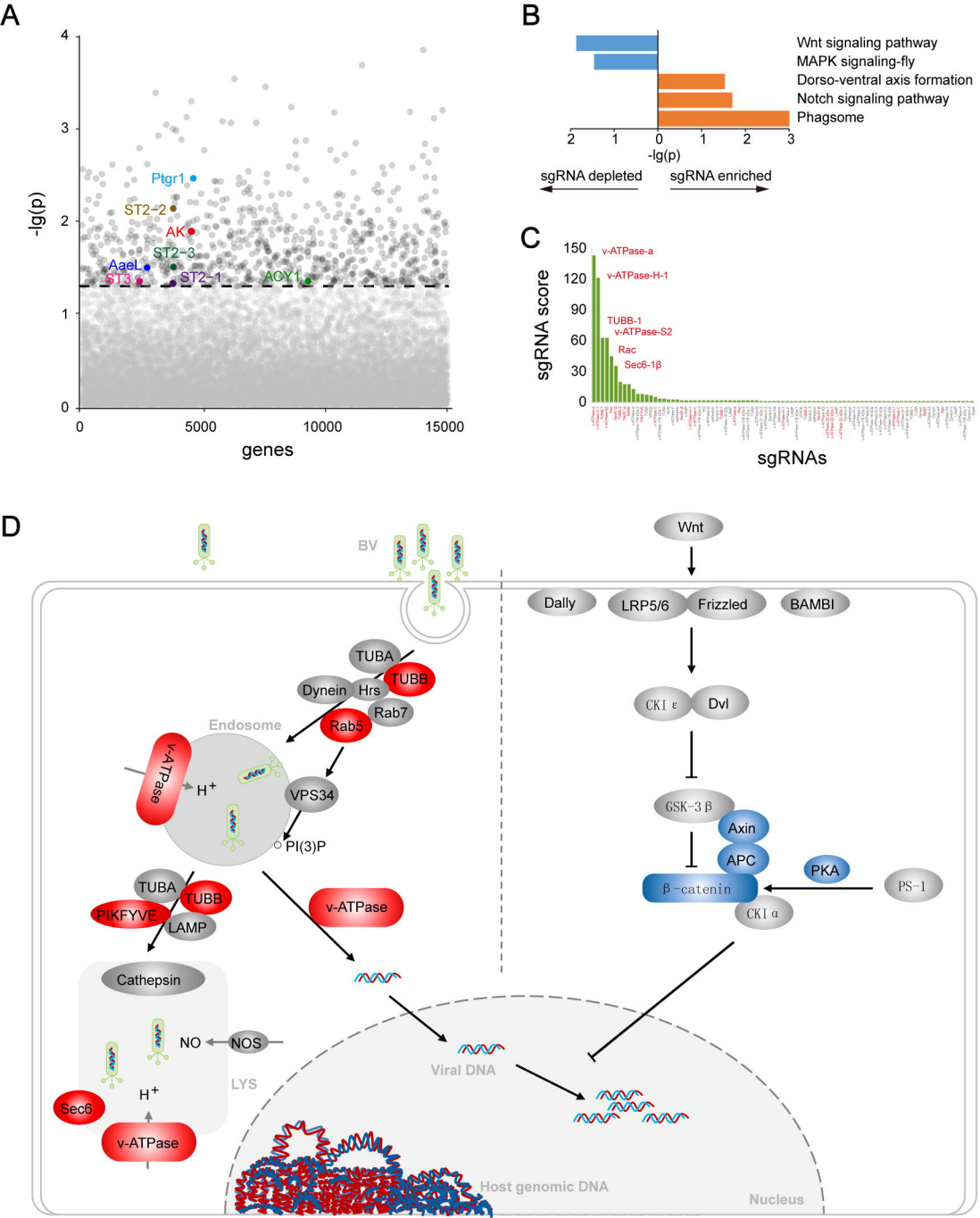
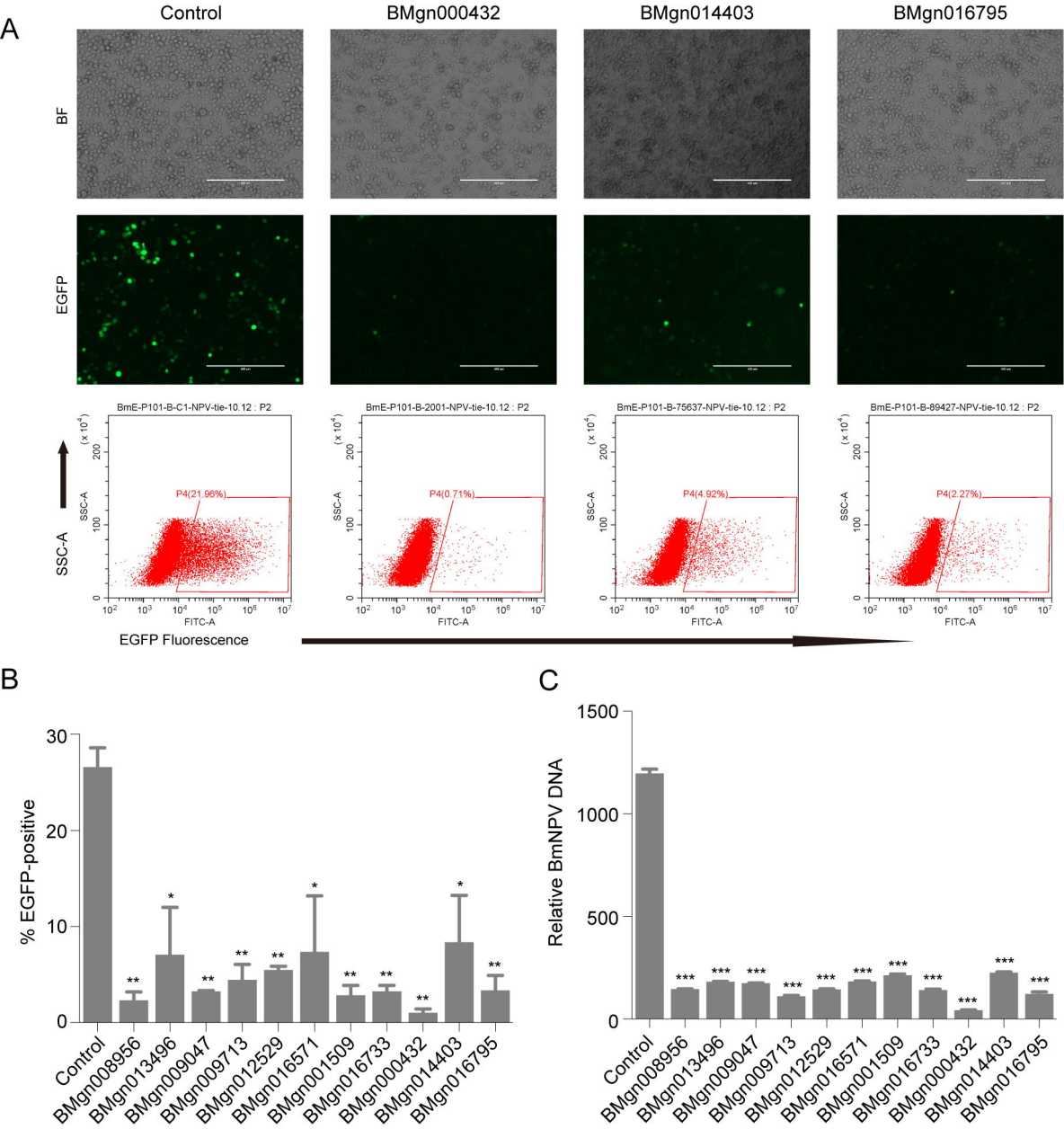

A
-lg(p)
genes
Ptgr1
ST2-2
AK
ST2-3
AaeL
ST3
ST2-1
ACY1
B
Wnt signaling pathway
MAPK signaling-fly
Dorso-ventral axis formation
Notch signaling pathway
Phagsome
-lg(p)
sgRNA depleted
sgRNA enriched
C
sgRNA score
sgRNAs
v-ATPase-a
v-ATPase-H-1
TUBB-1
v-ATPase-S2
Rac
Sec6-1β
D
BV
Wnt
Dally
LRP5/6
Frizzled
BAMBI
TUBA
TUBB
Dynein
Hrs
Rab5
Rab7
Endosome
v-ATPase
H+
VPS34
PI(3)P
CKI ε
Dvl
GSK-3 β
Axin
APC
PKA
PS-1
β-catenin
CKI α
TUBA
TUBB
PIKFYVE
LAMP
v-ATPase
Cathepsin
NO
NOS
LYS
H+
Sec6
v-ATPase
Viral DNA
Host genomic DNA
Nucleus

Figure 7


A
Control
BMgn000432
BMgn014403
BMgn016795
BF
EGFP
BmE-P101-B-C1-NPV-tie-10.12 : P2
BmE-P101-B-2001-NPV-tie-10.12 : P2
BmE-P101-B-75637-NPV-tie-10.12 : P2
BmE-P101-B-89427-NPV-tie-10.12 : P2
P4(21.96%)
P4(0.71%)
P4(4.92%)
P4(2.27%)
SSC-A
FITC-A
EGFP Fluorescence
B
% EGFP-positive
C
Relative BmNPV DNA
Control
BMgn008956
BMgn013496
BMgn009047
BMgn009713
BMgn012529
BMgn016571
BMgn001509
BMgn016733
BMgn000432
BMgn014403
BMgn016795