Vulnerability of honey bee queens to heat-induced loss of fertility

Alison McAfee1, Abigail Chapman2, Heather Higo2, Robyn Underwood3, Joseph Milone1, Leonard J. Foster2,3, M. Marta Guarna4,5, David R. Tarpy1,6 and Jeffery S. Pettis5

All species need to reproduce to maintain viable populations, but heat stress kills sperm cells across the animal kingdom and rising frequencies of heat waves are a threat to biodiversity. Honey bees (Apis mellifera) are globally distributed microlivestock; therefore, they could serve as environmental biomonitors for fertility losses. Here, we found that queens have two potential routes of temperature-stress exposure: within colonies and during routine shipping. Our data suggest that temperatures of 15–38 °C are safe for queens at a tolerance threshold of 11.5% loss of sperm viability, which is the viability difference associated with queen failure in the field. Heat shock activates expression of specific stress-response proteins in the spermatheca, which could serve as molecular biomarkers (indicators) for heat stress. This protein fingerprint may eventually enable surveys for the prevalence of heat-induced loss of sperm viability in diverse landscapes as part of a biomonitoring programme.

Climate change is threatening biodiversity around the globe2–3 and one potential driver is through heat-induced reductions in fertility4–6. The impact of heat on fertility is far-reaching in the animal kingdom, affecting mammals7–11, birds12, fish13, nematodes14 and insects15–21. At temperatures of 40–43 °C, spermogenesis is compromised15, sperm viability drops16,17, sperm are less competitive18,19, motility is compromised20,21. Extreme weather events such as heat waves are increasing in frequency and severity22–24, which could have widespread effects on populations via reduced reproductive output15–24.

Insects and other ectothermic animals are especially vulnerable to changes in local temperatures because, unlike mammals, they are less able to thermoregulate. Insects are critical components of ecosystems and agriculture, with economic estimates placing the global value of insect pollination at approximately €153 billion annually25–27. All species need to reproduce to maintain viable populations, but heat stress kills sperm cells across the animal kingdom and rising frequencies of heat waves are a threat to biodiversity. Honey bees are globally distributed microlivestock; therefore, they could serve as environmental biomonitors for fertility losses. Here, we found that queens have two potential routes of temperature-stress exposure: within colonies and during routine shipping. Our data suggest that temperatures of 15–38 °C are safe for queens at a tolerance threshold of 11.5% loss of sperm viability, which is the viability difference associated with queen failure in the field. Heat shock activates expression of specific stress-response proteins in the spermatheca, which could serve as molecular biomarkers (indicators) for heat stress. This protein fingerprint may eventually enable surveys for the prevalence of heat-induced loss of sperm viability in diverse landscapes as part of a biomonitoring programme.

To establish an environmental temperature biomonitoring programme, we monitored temperature fluctuations in colonies under extreme weather conditions, establishing that damaging intrahive temperatures can occur. Next, we tested a range of temperatures and exposure durations to determine thresholds above which queen quality is likely to be compromised. We then investigated the biochemical basis of heat-induced sperm viability reduction in queens and drones using quantitative proteomics, which showed how heat stress alters protein expression of reproductive tissues. The specific set of upregulated proteins we identified may eventually serve as diagnostic tools to elucidate causes of queen failure and eventually enable regional surveys of heat stress as part of an environmental temperature biomonitoring programme.

Results and discussion

Sperm viability losses associated with queen failure. To establish how much of a reduction in sperm viability is associated with field-observable reduced reproductive output (and associated economic losses for beekeepers), we collected queens rated as ‘failing’ (n = 58) and ‘healthy’ (n = 55) by beekeepers and measured the queens’ stored sperm viability (Fig. 1a). We found that the failed and healthy viability data were normally distributed (Shapiro test, P = 0.18 and 0.11, respectively) and that failed queens had significantly lower sperm viability (Student’s t-test, P = 5.8 × 10⁻⁴, F = 28.16), with an average viability of 55.8 ± 17.2% and 65.5 ± 15.1%, respectively. A critical temperature threshold for heat stress was determined as 40 °C to be safe for queens at a tolerance threshold of 11.5% loss of sperm viability.
average drop of 11.5%. We then set an 11.5% viability drop as the tolerance threshold in subsequent experiments aimed at identifying critical temperatures beyond which queens are at risk of substantial loss of stored sperm viability.

Temperature spikes occur in colonies and shipments. To document if routine shipping poses a threat of adverse temperature exposure to queens, we tracked the temperatures of eight domestic Queen shipments (seven via ground transportation, one via air; Fig. 1b). We found that even in these shipments, which were not deliberately timed to occur during extreme weather events, one package experienced a temperature spike to 38 °C and one dropped to 4 °C. Since honey bees cannot adequately thermoregulate in queen cages, extreme ambient temperatures are a hazard for shipping. However, little is known about a whole colony’s ability to thermoregulate in the face of extreme heat.

To gain a more complete picture of temperature fluctuations within colonies, we recorded temperatures throughout the brood nest (loggers placed between each frame of three ten-frame hives) during extreme heat in August in El Centro, California. The ambient temperatures, measured in the shade beneath each hive, reached up to 45 °C (Fig. 1c). In all three hives, the temperature at the two outermost frames spiked to upwards of 40 °C for 2–5 h; in two of the hives, temperatures exceeded 38 °C even one or two frames closer to the core. Therefore, the colony’s ability to thermoregulate begins to break down in extreme heat and queens could be vulnerable to temperature stress inside the hive.

Defining critical exposure and duration thresholds. Previous research has shown that both cold (4 °C) and hot (42 °C) temperatures reduce stored sperm viability in queens, but refined tolerance thresholds and biologically relevant viability losses are not known. To determine critical temperature and duration thresholds, we compared stored sperm viability across a temperature and duration gradient (5, 10, 15, 25, 38, 40 and 42 °C, exposed for 1, 2 or 4 h followed by a 2 d recovery period; Fig. 2a). Not all data from experimental groups were normally distributed (Shapiro test, P < 0.05); therefore, we analysed them with a Kruskal–Wallis test for non-parametric data. There was a significant effect of temperature for the 2 h (X^2 = 15.6, P = 0.016) and 4 h (X^2 = 17.9, P = 0.0065) treatments, while not at 1 h (X^2 = 9.12, P = 0.17). A Dunnett’s post hoc test revealed that the only temperatures that were significantly different from the control (25 °C) were the 2 h, 10 °C treatment (P = 0.045) and the 4 h, 42 °C treatment (P = 0.00057), at a family-wise error rate of α = 0.05. The 2 h and 4 h data were then pooled and optimally fit to a cubic polynomial regression (R^2 = 0.092, P = 0.012; Fig. 2b) to find the temperature tolerance thresholds for queens, given a predefined maximum acceptable drop in sperm viability (11.5%). This model suggests that 15.2–38.2 °C is the suggested ‘safe zone’ with minimal loss of viability for 2–4 h exposures.

To test effects of heat on ejaculated sperm viability, we exposed single-drone ejaculates to 42 °C for 0, 2, or 4 h, followed by a 2-d recovery period at 25 °C (Fig. 2c). These data were normally distributed (Shapiro test, P = 0.18); therefore, we used a two-way analysis of variance (ANOVA) for analysis. We found that responses differed depending on the colony source but heat dramatically decreased viability by 35% after both 2 and 4 h (factors: time and colony, P = 1.2 × 10^-7, F(time) = 24, d.f.(time) = 2, P(colony) = 0.00015, F(colony) = 11, d.f.(colony) = 2). Heat shock therefore affects stored and ejaculated sperm viability at similar magnitudes. Drones could also be appropriate biomonitors of heat stress, as their sperm is also sensitive to changes in temperature. However, drones are not as long-lived as queens and are only seasonally available.

Sex biases in heat-stress survival. Sturup et al. previously reported that drones are mortally sensitive to heat; however, queens are typically tolerant of stressful conditions. As a biomonitor, a favourable feature would be to survive through heat stress while accumulating a physiological and molecular record of the stress event(s). We compared drone, queen and worker survival over time at 25, 38 and 42 °C, confirming that drones, but not queens, are mortally sensitive to heat (log-rank test, P < 0.00001; Fig. 3a and Supplementary Fig. 1). A total 54% of drones died over the course of 6 h at 42 °C, whereas every queen survived. We also found that drones are more sensitive to heat than workers, which have a similar lifespan to drones but are non-reproductive females.

In this experiment, drones and workers were from local colonies, while queens were either from a local origin or imported from Hawaii and Australia. If bees are to be used as a global biomonitor for temperature stress, an important experiment will be...
to determine if there are differences in survival, physiological response or biochemical response to heat between genetic stocks that may be adapted to hotter or cooler climates.

The sex-biased heat sensitivity of honey bees is puzzling because drones and queens spend most of their lives inside the hive and can avoid participating in mating flights during hot weather. They are exposed to similar environmental conditions, and on the basis of this, one might expect them to have similar physiological tolerance thresholds. However, the ability of a colony to survive and produce reproductive individuals directly depends on the survival of individual queens (there is only one per colony) and not individual drones (hundreds per colony). Natural selection may have favoured high survivorship of queens, as evidenced by not only their long lifespan (up to 5 yr) but also their high tolerance of heat stress, despite the death of their stored sperm. To our knowledge, sex biases in thermal tolerance for other social insects have not yet been investigated, as critical thermal tolerance studies have focused on workers or on male fertility alone46–55.

Next, we investigated sex biases in heat tolerance for two solitary insect species—fruit flies (Drosophila melanogaster) and brown marmorated stink bugs (Halyomorpha halys). The stink bugs have a similar body size and thermal tolerance range as honey bees56 and...
their global distribution has expanded to cover four continents, making them a potential candidate biomonitor. Fruit flies have a much smaller body size but are also widely distributed, so they could be candidate biomonitors too. We found that both male and female stink bugs and fruit flies readily die with heat (Fig. 3b,c), meaning that they cannot accumulate a physiological or molecular record of a heat-stress event and have limited use for observing impacts of heat on fertility through time. However, these data do indicate another worrying trend: that exceptionally extreme temperatures may reduce insect populations through direct kills, in addition to reducing reproductive output if they survive. We found no difference between the heat sensitivities of males and females for fruit flies and a slight opposite (female) biased heat sensitivity ($P = 0.015$, log-rank test). In contrast, the high survivorship of honey bee queens makes them a good candidate biomonitor, should they possess molecular signatures indicative of heat stress when it occurs.

**Heat activates specific proteins in the spermatheca.** Sperm longevity is enabled by molecular processes that reduce oxidative damage and maintain sperm in a quiescent metabolic state. For example, the spermatheca is a highly anaerobic environment, which helps prevent formation of reactive oxygen species (ROS)$^{57}$. Enzymes that further limit damage from ROS are also upregulated in mated queens compared to virgins$^{34,39}$ and ROS damage leads to infertility in mammals$^{60,61}$. Heat is well known to lead to oxidative stress$^{62}$; therefore, we proposed that queens may combat heat stress by upregulating enzymes that mitigate oxidative damage. Additionally, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) has been previously implicated in stored sperm longevity, since metabolic activity measurements of stored sperm supplemented with this enzyme's substrate (glyceraldehyde-3-phosphate) improved viability and produced the highest molar ratio of ATP via anaerobic catabolism$^{57}$. Heat-induced changes in GAPDH expression could therefore also impact sperm viability.

To test these hypotheses, we compared expression of ROS-mitigating enzymes (superoxide dismutases (SOD1, 2 and 3) and catalase) as well as GAPDH in heat-shocked and non-heat-shocked virgin spermathecae, mated spermathecae and ejaculated semen. Stored sperm are reportedly transcriptionally active in queens to some degree$^{57}$, although whether this is true in general has been a matter of debate across species$^{63}$. It is not known with certainty whether ejaculated sperm are transcriptionally active; nevertheless, analysing ejaculates, virgin spermathecae and mated spermathecae helped us to disentangle the male versus female origins of expression in sperm-filled, mated spermathecae. We found that heat shock did not upregulate expression of ROS-mitigating enzymes, nor GAPDH (Supplementary Fig. 1); rather, all enzymes were consistently (but

---

**Fig. 3** Sex-biased heat mortality in honey bees, stink bugs and fruit flies. Statistical differences were evaluated using a log-rank test. Lower case letters indicate significant differences. See Supplementary Fig. 4 for risk tables, including population sizes used to generate survival curves. **a,** Honey bees (drones, queens and workers) were held at different temperatures and mortality was recorded hourly. Drones are sensitive to heat, whereas queens and workers are not. **b,** Fruit flies ($D. melanogaster$; heat shock = 38 °C) and stink bugs ($H. halys$; heat shock = 42 °C) do not have male-biased heat sensitivity. For fruit flies, loss of motility was used as the endpoint. $P$ values indicate comparisons between heat-shocked males and females. Stink bugs exhibit small but significant female-biased heat sensitivity.
A multitude of other proteins could also be responsible for mitigating damage from heat—most obviously, members of the heat-shock protein (HSP) family. HSPs generally function as molecular chaperones that stabilize proteins, refold damaged proteins and prevent protein aggregation but they can have diverse functions in specific contexts and their precise role in the honey bee spermatheca is unknown. Some HSPs contain an ATPase domain and require ATP to function (for example, HSP70s, HSP90s and HSP110s), while others operate independently of ATP (for example, HSP10s, HSP20s and HSP60s). In addition, since glyceraldehyde-3-phosphate yields the most ATP per unit via anaerobic catabolism of any tested substrate and significantly increased sperm longevity, we hypothesized that heat shock should induce expression of sHSPs in the spermatheca.

Not significantly downregulated with heat (by 10–30%). Queens must therefore use other strategies, if any, to combat heat stress.

To determine which HSPs were upregulated with heat shock, we compared global protein expression profiles in heat-shocked and non-heat-shocked mated and virgin spermathecae, ejaculates and ovaries (Fig. 4a and Supplementary Fig. 2). We proposed that the heat-shock response should both mitigate ROS production and conserve ATP in the spermatheca but this ATP-conservation may not be as critical in the ovaries where the ATP economy is not expected to be as tightly controlled. Of the 2,778 protein groups identified in the spermathecae and ejaculates, only five were significantly upregulated with heat (5% false discovery rate, FDR), all of which were identified in the mated and virgin spermathecae. All five of the proteins were unique, ATP-independent sHSPs (accessions: XP_001120194.1, XP_001119884.1, XP_395659.1, XP_001120006.2 and XP_026294937.1) (Fig. 4b). In heat-shocked virgin spermathecae, two of the same sHSPs (XP_395659.1 and XP_001120006.2) were also significantly upregulated and all mirrored the expression patterns in the spermathecae of heat-shocked mated queen (even if not significant), indicating that this is a queen-derived and not a sperm-derived, response (Fig. 4c,d). By contrast, the most strongly upregulated protein in heat-shocked ovaries was HSP70 (an ATP-dependent HSP) and no sHSPs were upregulated in this tissue, supporting our initial ideas (Fig. 4e).
spermathecae (Fig. 5a,b). However, the only significantly enriched GO term in the semen analysis was related to the electron transfer activity (Fig. 5c), driven by a heat-induced downregulation of the proteins linked to this GO term (Fig. 5d), suggesting that the sperm may be less able to produce the large amounts of ATP necessary for flagellar beating. This is consistent with Gong et al., who found that heat stress at 42°C impaired mitochondrial function, reduced the activities of electron transport chain complexes and lowered total cellular ATP. Numerous proteins (none of which were HSPs) were downregulated in heat-shocked spermathecae but, on the basis of the GO term enrichment analysis and manual inspection of their functions, it is unclear what their biological significance is. It is possible that they are degradation products of heat-killed sperm but, since these proteins were largely absent in the semen samples, we cannot confirm this.

Twenty HSPs were identified in spermathecae overall and 13 were identified in the semen (Fig. 6); however, the precise functions of specific honey bee HSPs are largely unknown. To gain insight into potential roles of the five differentially expressed HSPs we identified, we compared their sequences to annotated sequences in other species and identified putative protein domains using the National Center for Biotechnology Information (NCBI) basic local alignment search tool (BLAST). All five of the HSPs contain one or more alpha crystallin domains, which is characteristic of sHSPs. Four of these HSPs are within the expected molecular weight range and one of the proteins (XP_026294937.1) is predicted to be 56.7 kDa (and contains two alpha crystallin domains instead of one), highlighting that HSPs should not necessarily be categorized on the basis of molecular weight alone. All five of the honey bee sHSPs are paralogous to the D. melanogaster gene (2) (also known as CRYAB). In Drosophila, upregulation of this gene causes increased lifespan of individual flies. CRYAB and other sHSPs are highly conserved in both vertebrates and invertebrates but is by far the best studied in humans. In humans, sHSP upregulation is associated with anti-apoptotic properties, as well as mitigating ROS production. Their upregulation in heat-shocked testes is thought to help compensate for the damaging effects of heat and we speculate that they are playing a similar role in the spermatheca. Queens with strongly upregulated sHSPs may therefore be better able to counteract ROS production, sperm death and ultimately maintain longevity. However, it remains unknown if the honey bee paralogues have functionally diverged from those in other species.

Other researchers have suggested that in mammals, some ATP-dependent HSPs (for example, HSP70s and HSP90s) may be important for maintaining male fertility via quality control of sperm. Unlike sHSPs, these HSPs appear to be pro-apoptotic factors and could theoretically help prevent damaged sperm from being able to fertilize an egg. Neither queen spermathecae nor drone ejaculates provide evidence supporting this strategy of quality control in honey bees; in this experiment, none of the HSPs that were upregulated with heat in spermatheca or semen contained an ATP-binding nor ATPase domain, which are characteristics of HSP70s.
and HSP90s but not the sHSPs (Figs. 5 and 6)\(^\text{36,37}\). Rather, HSP70 was only upregulated in the ovaries, which do not directly participate in sperm maintenance.

In mammals, sHSPs are upregulated in testes after heating but so are ATP-consuming proteins such as HSP105, HSP70–1, HSP70–2 and HSP90 (refs. \(^\text{37,57}\)). We speculate that the energetic cost associated with ensuring high sperm quality is advantageous in mammals because it helps reduce the risk of an egg going unfertilized, which would produce no progeny. However, the honey bee’s differing strategy of upregulating only the ATP-independent sHSPs in the spermatheca is consistent with a tightly controlled ATP usage economy, as has been previously suggested\(^\text{36,37}\). The ATP-independent sHSPs have the dual purpose of also limiting oxidative damage and conserving ATP. The significant enrichment for differentially expressed proteins involved in multiple nucleotide metabolic processes, including ATP, in the transition from virgin to mated spermathecae (Supplementary Fig. 2) supports the notion that regulating ATP production and consumption are critical for maintaining stored sperm viability.

Moreover, analysing heat-shocked mated queen ovaries revealed that HSP70 (an ATP-dependent HSP) was uniquely upregulated in this tissue (Figs. 5 and 6). This observation is consistent with selection for ATP-consuming quality control mechanisms when a failed fertilization event fails to produce offspring (in this case, when a non-viable egg meets a viable sperm). Overall, these data indicate that upregulation of specific sHSPs is not a general indicator of cellular stress, since it was not observed in the ovaries. Indeed, the ovaries appear not to express most sHSPs at all, whereas they were abundantly expressed even in non-heat-shocked spermathecae.

Finally, we compared protein expression in spermathecae, fat bodies and ovaries between 11 queens that failed due to unknown causes and 11 age-matched, apiary-matched healthy queens. We identified 1,219, 1,640 and 1,782 proteins, respectively, but did not find any significant expression differences at 10% FDR, indicating that there is not one universal signature of failure (Supplementary Fig. 3). Rather, we suspect that different stressors alter protein expression in these tissues in different ways and that no significant differences were found because it is highly unlikely that the queens all failed from the same cause.

We propose that the sHSP response of the spermathecae may serve as a post-queen-failure biomarker of heat stress which could help diagnose causes of colony failure in the field and simultaneously provide an indication of risks for heat-induced loss of fertility for other species. Beekeepers routinely replace queens every 1 or 2 yr and these queens could be banked for biomarker testing as part of a surveillance programme. Future experiments should include a blind heat-shock trial to determine: (1) if previously heat-shocked queens can be reliably distinguished from non-heat-shocked queens on the basis of these biomarkers, (2) how long the heat-shock proteomic signature lasts and (3) if other stressors produce proteomic signatures that overlap with the heat-shock signature.

**Conclusion**

Our experiments show that temperature stress deeply damages both stored and ejaculated honey bee sperm viability, that queens are vulnerable to temperature changes both in colonies and during transport, and that temperatures ranging between 15 °C and 38 °C for 2–4 h are generally safe for queens. Honey bees have a strong sex bias in heat tolerance, with females being highly tolerant—a bias that does not exist in the two solitary insect species we tested (*H. halys* and *D. melanogaster*). Future research should investigate if this sex-biased heat tolerance is also present in other Hymenopterans to better understand the evolutionary origin. Upon heat shock, queens upregulate ATP-independent HSPs in their spermathecae, which both minimizes ATP consumption and could provide beneficial anti-apoptotic properties. In contrast, HSP70 (an ATP-consuming HSP) is upregulated in ovaries. Once validated in field trials, these protein signatures could serve as biomarkers for heat stress enabling longitudinal surveys for the prevalence of heat-induced loss of sperm viability in diverse landscapes as part of a biomonitoring programme.

**Methods**

Sperm viability assays. Honey bee queens (Kona Queens supplied in a single shipment) were treated at one of seven different temperatures (5, 10, 15, 23, 38, 40 and 42 °C) for 1, 2, or 4 h, then held at 25 °C for 2 d. The temperature range was chosen because previous research showed that 4 and 8 °C were sufficiently cold to reduce sperm viability and that 40–42 °C was sufficiently hot. Therefore, more temperatures were chosen at the cold (5, 10 and 15 °C) and hot (38, 40 and 42 °C) extremes to try to capture the critical point at which the viability begins to drop. We chose 25 °C as the control temperature, rather than 35 °C, because these queens were obtained from a commercial supplier and queens are not held at hive temperature during transport. While a 35 °C-treatment group would be an appropriate control, it would not be meaningful in these circumstances.

Fourteen queens (replicates) were included in the 25 °C treatment (negative control), whereas eight or nine queens were included in all other temperatures and exposure durations (see Supplementary Table 1 for specific replication information). Following this, queens were beheaded and their spermathecae were dissected with fine forceps. The spermathecae were gently agitated with a pipet tip in 100 μl of Buffer D (17 mM d-glucose, 54 mM KCl, 25 mM NaHCO\(_3\), 83 mM Na\(_2\)C\(_6\)H\(_5\)O\(_7\)) to break them open and release the sperm. Sperm viability was determined using a live/dead sperm viability kit (Thermo) following the protocols of Collins and Donoghue\(^\text{77}\). Briefly, the SYBR14 dye was diluted 1:9 in dimethylsulfoxide and 2 μl of the diluted SYBR14 dye and 4 μl of propidium iodide were gently mixed with the sperm and incubated for 15 min in the dark. Then 2 μl of the mixture was added to a glass microscope slide and viewed under a fluorescent microscope. Live (green) and red (dead) sperm were counted until 100 cells were observed, covering multiple fields of view. Unless otherwise reported, all statistical analyses were performed in R (v.3.5.1). First, the data were tested for normality using a Shapiro test. The data were not normally distributed; we used a non-parametric test (Kruskal–Wallis) to investigate the effects of temperature on viability for each duration separately. When a significant effect of temperature was identified, we performed a post hoc Dunnets test to identify significant contrasts to the 25 °C control. To identify the best-fitting linear model, we pooled the 2 h
and 4h data and performed linear regressions using temperature as a continuous variable, testing exponents 1–4 to identify the optimal fit (highest $R^2$).

Drones were harvested from three different colonies headed by Kona queens with similar ages, kept at 37 °C, with 100% RH. The colders were kept at 3 °C for several hours (in some cases, overnight) at cooler temperatures. Control female flies were maintained for a further week after the experiment to confirm that most of them had successfully mated before the heat-stress tests. Male and female stink bugs were reared in the laboratory according to standard protocols. Stink bugs which transitioned from nymphs to adults 3–7 d prior were transferred to ventilated 500-ml volume plastic cages with a piece of moist paper towel (three to six stink bugs per cage). They were heat stressed at 42 °C (50 females) and 41 males) or held at 25 °C (36 females and 24 males) for 6 h. Only mated females carrying eggs were included in the experiment, as determined by poststress dissection.

In all cases, Kaplan–Meier survival curves were generated in R and compared using log–rank tests. See Supplementary Fig. 3 for risk tables and sample size information for survival tests for all species.

Heat shock for proteomics. Our experimental design for the proteomics experiments was to compare heat-shocked and non-heat-shocked treatments of the same Greek bees. Virgin queens (n = 5 each for each treatment) were given two, 8-min CO2 treatments on sequential days, then re-introduced to the queen's colonies. This process stimulates virgin queens to begin laying9 and we conducted these treatments to minimize the physiological differences between virgin and mated queens.

Virgin and mated queens were retrieved from their colonies and half of each (ten) were subjected to heat shock (42 °C, 2 h) and then maintained at 30 °C for 2 d. The other half were held only at 30 °C for 2 d. At 4–6 weeks after mating, the queens were anaesthetized with CO2, beheaded, then their spermathecae were removed with fine forceps. Both ovaries were also removed and weighed.

During the same period, 200 drones from a different colony in the same apiary were collected and maintained in the laboratory overnight at ambient temperature with excess syrup (50% sucrose). The next day, semen was harvested with glass capillaries according to the methods described earlier. Because many drones were not sexually mature, 60 semen samples (out of the 200 drones) were collected. Capillaries were placed in Petri dishes and half (30) were heat-shocked as described, then kept at 25 °C for 2 d. The other half were only kept at 25 °C for 2 d. Ten samples from each experimental group were used for sperm viability assays as described earlier.

Proteomics sample preparation. Semen and spermatheca samples were homogenized in 2-ml screw-cap tubes containing 100 µl of lysis buffer (6 M guanidinium chloride, 100 mM Tris, pH 8.5) and four ceramic beads. The Lysing Instrument (ProSciTech) was used to record the temperature every 10 min and were placed between each frame of the hive. Hives consisted of ten deep frames in a single wooden Langstroth brood box with a medium depth box on top of the brood chamber and a migratory wooden cover. The ambient temperatures were recorded with a temperature logger placed in the shade beneath each hive.

Hive temperature recording. To record the internal and external temperatures of colonies, we placed temperature loggers inside three standard, ten-frame colonies in El Centro, California, from 12 August to 6 September 2017. The hottest day (the day before the heat stress) was August 4. Loggers were turned on every 10 min and were placed between each frame of the hive. Hives consisted of ten deep frames in a single wooden Langstroth brood box with a medium depth box on top of the brood chamber and a migratory wooden cover. The ambient temperatures were recorded with a temperature logger placed in the shade beneath each hive.

Heat-shock survival. For honey bee heat-shock tests, drones and workers were collected by retrieving a frame of capped brood from a hive and allowing the bees to emerge in an incubator. Newly emerged bees were marked with a paint pen and returned to the colony to age for 1 week, at which time they were recaptured and caged in California field cages. For drone samples, bees were added per cage along with malleable candy (icing sugar-based) and five worker attendants. A total of 199 drones were caged in the following experimental groups: 42 °C heat (50 drones), 38 °C heat (50 drones) and 25 °C (99 drones) for 6 h (all at 60% relative humidity). Every hour, the number of drones that perished was recorded. Worker heat-stress tests were conducted independently with six workers per cage (54 workers per group) at 42 and 25 °C.

Queens were from a variety of sources (local, Hawaii and Australia) and their ages ranged from approximately 3 to 5 weeks. They were kept in California mini-cages with five worker attendants each and held at 42, 38 or 25 °C for either 1, 2 or 4 h as part of other experiments (for viability measurements and proteomics). Relative humidity varied between 40 and 80% but this did not have an effect on survival (100% of queens survived in both cases). Eight queens were held at 42 °C for 6 h at a constant 60% relative humidity specifically for survival analysis. See Supplementary Fig. 4 for sample sizes and risk tables for each treatment group. No queens died in these tests; therefore it was not necessary to account for geographic origin in statistical analyses.
Mass spectrometry data were searched using MaxQuant (v.1.5.3.30) using default parameters, except ‘match between runs’ was enabled. Peptide spectral matches, peptide identifications and protein identifications were controlled at 1% FDRs. The protein search database was the NCBI Identical Protein Groups database for *Apis mellifera* (downloaded 1 November 2018; 21,425 entries) plus all honey bee viral proteins contained within NCBI (a further 508 entries). Differential expression analysis was performed in Perseus (v.1.6.1.1) essentially as previously described 

5. Walsh, B. S. et al. The impact of climate change on fertility. *Nature Sustainability* 18, 647–653 (2016).

6. Hurley, L. L., McDiarmid, C. S., Friesen, C. R., Griffith, S. C. & Rowe, M. Experimental heatwaves negatively impact sperm quality in the zebra finch. *Proc. R. Soc. B* 285, 20172547 (2018).

7. Breckels, R. D. & Neff, B. D. The effects of elevated temperature on the sexual traits, immunology and survivorship of a tropical ectotherm. *J. Exp. Biol. 216*, 2664–2668 (2013).

8. Harvey, S. C. & Viney, M. E. Thermal variation reveals natural variation between isolates of Caenorhabditis elegans. *J. Exp. Zool. B Mol. Evol. 308*, 409–416 (2007).

9. Porcelli, D., Gaston, K. J., Butlin, R. K. & Snoook, R. R. Local adaptation of reproductive performance during thermal stress. *J. Evol. Biol. 30*, 422–429 (2017).

10. Goeman, J. L., Brik, M., Cuzick, J., Brown, M. D., Neuvirth, E. & Meuleman, P. Genomic control for multiple testing in cancer genomics. *Hum. Mol. Genet. 21*, 4282–4288 (2012).

11. Zhao, Y., Leek, J. T. & Storey, J. D. qvalue: an R package for multiple hypothesis testing. *Bioinformatics 27*, 1107–1108 (2011).

12. Perseus: a comprehensive and interactive omics platform for the analysis of mass spectrometry data. *Nature Sustainability* 18, 409–416 (2020).

13. Willms, K. & Schütz, M. S. Heat stress in bees: a review of the effects on life span and reproductive success. *Apidologie 41*, 495–513 (2010).

14. Bangs, O. D., Hoehn, M. A., van der Weijden, G. & Visscher, R. P. The effects of heat stress on the longevity and reproduction of *Apis mellifera* queens. *Apidologie 37*, 212–220 (2006).

15. Bachtler, V. & Stadler, K. The effects of heat on the testes of mammals. *Anim. Reprod. 3*, 81–91 (2006).

16. Thonneau, P., Bujan, L., Multigner, L. & Mieusset, R. Occupational heat exposure and male fertility: a review. *Hum. Reprod. 13*, 2122–2125 (1998).

17. Yaeram, J., Setchell, B. P. & Maddocks, S. Effect of heat stress on the fertility of honey bee queens in vitro. *Reprod. Fertil. Dev. 18*, 647–653 (2006).

18. Hall, G. C. & Marshall, P. J. Heat stress effects on the reproductive development of honey bee queens. *Proc. R. Soc. Lond. B* 278, 1212–1219 (2011).

19. Crowell, S. A. et al. Heat tolerance and fertility in honey bee queens. *Apidologie 38*, 309–322 (2007).

20. Baer, S. A. et al. The effects of heat stress on the life span and reproduction of honey bee queens. *Apidologie 39*, 385–396 (2008).

21. Luber, G. & McGeehin, M. Climate change and extreme heat events. *Am. J. Prev. Med. 30*, 429–435 (2006).

22. Grulke, P. et al. The effects of heat stress on the life span and reproduction of honey bee queens. *Apidologie 39*, 323–332 (2008).

23. Hallmann, C. A. et al. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE 11*, e0185809 (2017).

24. Yang, R. C., Shen, M. R., Chiang, P. H., Yang, S. L. & Chen, S. S. A possible role of heat shock proteins in human sperm motility. *Gaongxi Yi Xue Ke Xue Za Zhi 8*, 299–305 (1992).

25. Luber, G. & McGeehin, M. Climate change and extreme heat events. *Am. J. Prev. Med. 30*, 429–435 (2006).

26. Grulke, P. et al. The effects of heat stress on the life span and reproduction of honey bee queens. *Apidologie 39*, 323–332 (2008).

27. Hayhoe, K., Sheridan, S., Kalkstein, L. & Greene, S. Climate change, heat waves, and mortality projections for Chicago. *J. Great Lakes Res. 36*, 65–73 (2010).

28. Beaudry, P. et al. Heat stress effects on the life span and reproduction of honey bee queens. *Appl. Sci. 8*, 108 (2018).

29. Hallmann, C. A. et al. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE 11*, e0185809 (2017).

30. Cameron, S. A. et al. Patterns of widespread decline in North American bumble bees. *Proc. Natl. Acad. Sci. USA 108*, 662–667 (2011).

31. Rasmont, P. & Mersch, P. First estimation of faunistic drift by bumblebees of Belgium (Hymenoptera: Apidae). *Apidologie 31*, 141–147 (1988).

32. Winfree, R., Aguilar, R., Vázquez, D. P., LeBuhn, G. & Aizen, M. A. Meta-analysis of bees’ responses to anthropogenic disturbance. *Ecology 90*, 2068–2076 (2009).

33. Ricketts, T. H. et al. Landscape effects on crop pollination services: are there general patterns? *Ecol. Lett. 11*, 499–515 (2008).

34. Biesmeijer, J. C. et al. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science 313*, 351–354 (2006).

35. Whitehorn, P. R., Tinsley, M. C., Brown, M. J., Darvill, B. & Goulson, D. Genecide diversity; parasite resistance and community in wild bumblebees. *Proc. R. Soc. Lond. B* 278, 1195–1202 (2010).

36. McDermott, H. & Dobson, A. Disease, habitat fragmentation and conservation. *Proc. Biol. Sci. 269*, 2041–2049 (2002).

37. Baer, B., Collins, J., Maalaps, K. & den Boer, S. P. Sperm use economy of honeybee (*Apis mellifera*) queens. *Ecol. Evol. 6*, 2877–2885 (2016).

38. Delaney, D. A., Keller, J. J., Aron, R. & Tarpy, D. R. The physical, insemination, and reproductive quality of honey bee queens (*Apis mellifera L.*). *Apidologie 42*, 1–13 (2011).

39. Tarpy, D. R. & Olivaraz, R. Jr Measuring sperm viability over time in honey bee queens to determine patterns in stored-sperm and queen longevity. *J. Apic. Res. 53*, 493–495 (2014).

40. Stabenheinert, A., Kovac, H. & Broschneider, R. Honeybee colony thermoregulation—regulatory mechanisms and contribution of individuals in dependence on age, location and thermal stress. *PLoS ONE 5*, e9867 (2010).
41. Alattal, Y. Impact of temperature extremes on survival of indigenous and exotic honey bee subspecies, *Apis mellifera*, under desert and semiarid climates. *Bull. Insectol.* 68, 219–222 (2015).

42. Fahrendohl, L., J. I., & Schröder, U. Thermal investigations of a honey bee colony: thermoregulation of the hive during summer and winter and heat production of members of different bee castes. *J. Comp. Physiol. B* 159, 551–560 (1989).

43. Bordier, C. et al. Colony adaptive response to simulated heat waves and consequences at the individual level in honeybees (*Apis mellifera*). *Sci. Rep.* 7, 3760 (2017).

44. Villa, J. D., Gentry, C. & Taylor, O. R. Preliminary observations on thermoregulation, clustering, and energy utilization in African and European honey bees. *J. Kans. Entomol. Soc.* 60, 4–14 (1987).

45. Smith, K. E. et al. Honey as a biomonitor for a changing world. *Front. Genet.* 10, 551–560 (2019).

46. Clémencet, J., Cournault, L., Odent, A. & Yanoviak, S. P. Variation in thermal tolerance of North American ants. *J. Insect Physiol.* 107, 65–68 (2017).

47. Macías-Macías, J. O. et al. Comparative temperature tolerance in stingless bee *Heterotermes mossambicus* (Hagen). *J. Insect Physiol.* 39, 523–528 (1993).

48. Chirault, M. et al. A combined approach to heat stress effect on male fertility in *Apis mellifera*. *Apidologie* 48, 769–782 (2017).

49. Verble-Pearson, R. M., Gifford, M. E. & Yanoviak, S. P. Variation in thermal tolerance of *Hymenoptera: Apidae: Meliponini*. *Apidologie* 42, 670–689 (2011).

50. Abou-Shaara, H. F., Al-Ghamdi, A. A. & Mohamed, A. A. Tolerance of two genotypes in the parasitic wasp *Nasonia vitripennis* (*Hymenoptera: Braconidae*). *Environ. Exp. Biol.* 78, 870–880 (2013).

51. Scaccini, D., Duso, C. & Pozzebon, A. Lethal effects of high temperatures on *Drosophila melanogaster* (*Diptera: Drosophilidae*). *Cell. Mol. Life Sci.* 73, 35–43 (2016).

52. Baudier, K. M., Mudd, A. E., Erickson, S. C. & O'Donnell, S. Microhabitat and worker thermal environmental and temperature gradient for a temperate parasite bee *Trigona rufa* (Hymenoptera: Apidae: Eucera). *BMC Bioinf.* 6, 27 (2005).

53. Li, Y., Chen, X. & Yang, Q. Structural integrity of developing muscles, and proper muscle and heart development in *Drosophila melanogaster*. *Dev. Biol.* 301, 539–554 (2007).

54. Smith, K. E. et al. Honey as a biomonitor for a changing world. *Front. Genet.* 10, 351 (2019).

55. McAfee, A., Pettis, J. S., Tarpy, D. R. & Foster, L. J. Feminizer and doublesex knock-outs cause honey bee survival experiments. We also thank Ashurst Bee Company for help with colony management and Kettle Valley Queen Honey, Wildwood Honey, Honey, Heather Meadows Honey Farm, Six Legs Good Apiaries, Wildwood Queens, Cariboo Honey and Worker Bee Honey Company for donating failed and healthy queens for this research.

**A. C. and M. O. conducted the failed queen survey, with assistance from H.H. and M.M.G. H.H. and M.M.G. contributed intellectually.**

**Acknowledgements**

This work was supported by Natural Sciences and Engineering Research Council of Canada Discovery grant no. 31164-11 and grants from Genome Canada and Genome British Columbia awarded to J.L.F; Project Apis m grants awarded to A.M., M.M.G. and J.S.P.; and USDA NIFA grant no. 2016-07962 awarded to J.S.P. and D.R.T. We thank A. Sébastien and M. J. for providing us with stink bugs and fruit flies, respectively, for the survival experiments. We also thank Ashurst Bee Company for help with colony management and Kettle Valley Queen Honey, Wildwood Honey, Honey, Heather Meadows Honey Farm, Six Legs Good Apiaries, Wildwood Queens, Cariboo Honey and Worker Bee Honey Company for donating failed and healthy queens for this research.

**Author contributions**

A.M. wrote the first draft of the manuscript and revisions, conducted all data analysis, made the figures and performed the proteomics experiments. A.C. and M. O. conducted the failed queen survey, with assistance from H.H. and M.M.G. H.H. and M.M.G. executed the queen shipment temperature tracking. J.M. performed the survival experiments. M.M.G. and J.S.P. performed the drone sperm viability analyses. R.U. contributed the age-matched failed and healthy queens. J.S.P. performed the queen sperm viability measurements across the range of temperatures and measured internal hive temperatures. Grants to D.R.T., J.S.P., M.M.G., A.M. and J.L.F. supported this research. All authors contributed intellectually.

**Competing interests**

J.S.P. owns a honey bee consulting business. All other authors have no competing interests.

**Additional information**

Supplementary information is available for this paper at https://doi.org/10.1038/s41893-020-0493-x.

Correspondence and requests for materials should be addressed to J.L.F., M.M.G. or D.R.T.

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. © The Author(s), under exclusive licence to Springer Nature Limited 2020.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a
- [ ] The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- [ ] A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- [ ] The statistical test(s) used AND whether they are one- or two-sided
- [ ] Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- [ ] A description of all covariates tested
- [ ] A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- [ ] A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- [ ] For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  - Give P values as exact values whenever suitable.
- [ ] For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- [ ] For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- [ ] Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection
- Mass spectrometry data was acquired using Hystar (v3.2.44.0) and microTOF acquisition software (v4.0.49.3245).

Data analysis
- R (v3.5.1), Perseus (v1.6.1.1), MaxQuant (v1.5.3.30), BLAST2GO (v4.1.9)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw mass spectrometry data, databases, and search results are available on PRIDE ProteomeXchange (accession: PXD013728). Figures with associated raw mass spectrometry data include Figs. 4, 5, 6, and S4. Data used to generate other figures will be made available upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences
Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | No a priori sample size calculations were performed. For proteomics data, sample sizes are typically in the 3-5 biological replicates range, and our replication (n = 5-10) exceeds the norm. Sample sizes (number of queens) used for stored sperm viability measurements were chosen based on results from Pettis et al. (2016), where n = 8-14 queens was sufficient to identify strong significant differences in temperature-dependent viability. Sample size sufficiency for Kaplan Meier curves was not tested. |
| Data exclusions | Proteins identified in fewer than three biological replicates per experiment were excluded from the analysis because they are more likely to be spurious identifications and have limited statistical utility. Proteins which were likely contaminants (e.g. keratin), reverse hits, and only identified by site were also excluded. No other data were excluded. |
| Replication | We have since replicated the proteomics results of the heat-shock experiments, determining that heat-stressed queens from an entirely different origin (Australia) still upregulate the same heat-shock proteins described in this paper. The stored sperm viability analysis partially replicates a previous study by JSP, where he found that 4C and 40C temperatures significantly reduced stored sperm viability after 2 h exposures. No further attempts have yet been made to replicate the results. |
| Randomization | Queen assignment to temperature and time experimental groups was performed randomly. Drone assignment was also random. Sample order of injection for MSMS analysis was randomized. |
| Blinding | AM remained blind to the health status of the ‘healthy’ and ‘failing’ queens sampled by RU and JSP until MSMS data acquisition was complete. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a | n/a |
| ☒ Antibodies | ☒ ChIP-seq |
| ☒ Eukaryotic cell lines | ☒ Flow cytometry |
| ☒ Palaeontology | ☒ MRI-based neuroimaging |
| ☒ Animals and other organisms | |
| ☒ Human research participants | |
| ☒ Clinical data | |

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

- **Laboratory animals**: This study did not involve laboratory animals
- **Wild animals**: This study did not involve wild animals
- **Field-collected samples**: Queen honey bees for the stored sperm viability experiments were purchased from a commercial queen rearing operation in Hawaii (Kona Queens). Queen honey bees used for the proteomics experiments were reared locally by HH using standard queen rearing techniques in the summer of 2018 and open-mated in Langley, British Columbia. Drones for viability analyses were collected from honey bee colonies in Beaverlodge, Alberta, and Vancouver, British Columbia during the summer of 2018. Drones for survival tests were collected from a single colony in Raleigh, North Carolina, in April of 2019. Workers for survival tests were collected from Vancouver, BC, in April, 2019. Queens for survival tests were of mixed origins, including Australia, Hawaii, and Local BC. Queens that were rated as healthy and failing in the field were collected from multiple experimental apiaries in Pennsylvania, where each queen pair was age- and apiary-matched. In all cases, honey bees were collected and experimented on during daylight hours (between 8:00 am and 6:00 pm). Photoperiod within this range is not known to affect the parameters tested. In all cases, honey bees were housed in standard Langstroth beekeeping equipment, allowed to self-thermoregulate and were maintained (fed, tested and treated for diseases, split, etc.) according to standard beekeeping practices as needed. Queens for the healthy vs failed survey were collected from multiple beekeepers and queen breeders throughout BC between June and August 2019.

- **Ethics oversight**: As insects, honey bees, stink bugs, and fruit flies are not subject to animal ethics guidelines at the University of British Columbia.

Note that full information on the approval of the study protocol must also be provided in the manuscript.