File S1. Supplementary Notes

Note 1. Competitive Advantage of a Mutant in Sperm Competition

In the main text, we derived a master equation for the probability that a mutant sperm with allele A at a focal locus wins in a competition among N_A mutant sperm and N_B wild-type sperm (with allele B):

$$
\begin{aligned}
& P\left[\text{winner} = A \mid N_A \; A's \; \& \; N_B \; B's\right] \\
= & \int_{-\infty}^{+\infty} dx \frac{d}{dx} \left\{ P\left[\max\{X_1^A, ..., X_{N_A}^A\} < x\right]\right\} \; P\left[\max\{X_1^B, ..., X_{N_B}^B\} < x\right] \\
= & \int_{-\infty}^{+\infty} dx \; N_A f_A(x) \left(1 - P\left[X^A > x\right]\right)^{N_A - 1} \left(1 - P\left[X^B > x\right]\right)^{N_B} \\
= & \; 1 - \int_{-\infty}^{+\infty} dx \; \left(1 - P\left[X^A > x\right]\right)^{N_A} N_B f_B(x) \left(1 - P\left[X^B > x\right]\right)^{N_B - 1} , \qquad \text{(S1)}
\end{aligned}
$$

which is identical to Eq. 2 in the main text. Here $f_A(x)$ and $f_B(x)$ are the probability density functions of a sperm competitiveness measure x for allele A and allele B, respectively. When $N_A, N_B \gg 1$, it can be approximated as:

$$
\begin{aligned}
& P\left[\text{winner} = A \mid N_A \; A's \; \& \; N_B \; B's\right] \\
\approx & \int_{-\infty}^{+\infty} dx \; N_A f_A(x) \exp\left\{ - \left(N_A P\left[X^A > x\right] + N_B P\left[X^B > x\right]\right)\right\} \\
\approx & \; 1 - \int_{-\infty}^{+\infty} dx \; N_B f_B(x) \exp\left\{ - \left(N_A P\left[X^A > x\right] + N_B P\left[X^B > x\right]\right)\right\} . \qquad \text{(S2)}
\end{aligned}
$$

The exact master equation Eq. S1 provides the winning probability:

$$
P\left[\text{winner} = A \mid N_A \; A's \; \& \; N_B \; B's\right] = \frac{N_A}{N_A + N_B} ,
$$

when the mutation (from B to A) is exactly neutral, i.e. $f_A(x) \equiv f_B(x)$.

(1-0) Perturbation formula in general case

In population genetics, it is very common to deal with a situation where the effect of the mutation is fairly small at a generation but could become large when accumulated through

K. Ezawa and H. Innan

generations. To deal with such situations, let us assume that the probability density functions (PDFs) for the two alleles are almost identical:

$$f_A(x) \equiv f(x) + \delta f(x) , \quad f_B(x) \equiv f(x) , \quad \text{with} \quad \int_{-\infty}^{+\infty} dx |\delta f(x)| \ll 1.$$

And let X_f denote a random variable conforming to the PDF $f(x)$. Then the probability, Eq. S1, that the winner has allele A is rewritten and approximated up to $O(\delta f)$ as:

$$
\begin{aligned}
& P\,[\text{winner} = A \mid N_A\ A's\ \&\ N_B\ B's\,] \\
= {}& 1 - \int_{-\infty}^{+\infty} dx\ N_B f(x) \left[1 - \int_x^{+\infty} d\xi\ (f + \delta f)(\xi)\right]^{N_A} \left[1 - \int_x^{+\infty} d\xi\ f(\xi)\right]^{N_B - 1} \\
\approx {}& 1 - N_B \int_{-\infty}^{+\infty} dx\ f(x) \left[1 - \int_x^{+\infty} d\xi\ f(\xi)\right]^{N_A + N_B - 1} \\
& + N_B \int_{-\infty}^{+\infty} dx \left\{ f(x) \left[1 - \int_x^{+\infty} d\xi\ f(\xi)\right]^{N_A + N_B - 2} N_A \int_x^{+\infty} d\xi\ \delta f(\xi) \right\} \\
= {}& 1 - \frac{N_B}{N_A + N_B} + \frac{N_A\ N_B}{N_A + N_B - 1} \int_{-\infty}^{+\infty} dx \left\{ \frac{d}{dx} \left[1 - \int_x^{+\infty} d\xi\ f(\xi)\right]^{N_A + N_B - 1} \int_x^{+\infty} d\xi\ \delta f(\xi) \right\} \\
= {}& \frac{N_A}{N_A + N_B} + \frac{N_A\ N_B}{N_A + N_B - 1} \int_{-\infty}^{+\infty} dx \left[1 - \int_x^{+\infty} d\xi f(\xi)\right]^{N_A + N_B - 1} \delta f(x) . \quad\quad (\text{S3})
\end{aligned}
$$

Partial integration was used to achieve the last equation. Now, define a quantity:

$$
\begin{aligned}
\psi_A(N_A,\ N_B) & \equiv \frac{(N_A + N_B)^2}{N_A + N_B - 1} \int_{-\infty}^{+\infty} dx \left[1 - \int_x^{+\infty} d\xi f(\xi)\right]^{N_A + N_B - 1} \delta f(x) \\
& = \frac{N_A + N_B}{N_A + N_B - 1} \int_0^{N_A + N_B} dy \left(1 - \frac{y}{N_A + N_B}\right)^{N_A + N_B - 1} \delta \ln f(x(y)) \quad (\text{S4})
\end{aligned}
$$

The latter equation results from changing dummy integration variables from x to $y \equiv (N_A + N_B)P\,[X_f > x]$ and introducing the notation, $\delta \ln f(x) \equiv \frac{\delta f(x)}{f(x)}$. When $N_A + N_B \gg 1$, it is approximated as:

$$
\begin{aligned}
\psi_A(N_A,\ N_B) & \approx (N_A + N_B) \int_{-\infty}^{+\infty} dx\ \exp\left(-(N_A + N_B)P\,[X_f > x]\right) \cdot \delta f(x) \\
& = \int_0^{+\infty} dy\ e^{-y}\ \delta \ln f(x(y)) . \quad\quad (\text{S5})
\end{aligned}
$$

The first approximate equation gives exactly Eq. 4 in the main text.

With Eq. S4 (or Eq. S5), Eq. S3 can be rewritten and rearranged as:

$$P \left[\text{winner} = A \mid N_A \ A's \ \& \ N_B \ B's \right] \approx \frac{N_A}{N_A + N_B} \left\{ 1 + \frac{N_B}{N_A + N_B} \psi_A \right\}$$

$$\approx \frac{N_A}{N_A + N_B(1 - \psi_A)}$$

$$\approx \frac{N_A(1 + \psi_A)}{N_A(1 + \psi_A) + N_B} \ , \tag{S6}$$

which is referred to as Eq. 3 in the main text. Here we omitted the dependence of ψ_A on N_A and N_B for notational convenience. The approximate equations Eq. S6 demonstrate that allele A has an advantage as much as ψ_A, as defined in Eq. S4, over allele B in the sperm competition. Thus the problem boils down to estimating the competitive advantage $\psi_A(N_A, \ N_B)$.

Let us now calculate the competitive advantage in several specific cases.

(1-1)Increased mean in exponential distribution

First we consider a simplest example, where the measure x follows an exponential distribution and the mutation slightly increases the mean:

$$f(x) = \exp(-x) \ , \quad \text{and} \quad f(x) + \delta f(x) = (1 - \delta\tau) \exp(-(1 - \delta\tau)x) \ .$$

Here we rescaled x so that the mean is 1 for the wild-type.

In this case, $y = (N_A + N_B) \exp(-x)$, and

$$\delta \ln f(x) \approx \delta\tau(x - 1) = \delta\tau \left\{ \ln\left(\frac{N_A + N_B}{y}\right) - 1 \right\} \ .$$

Substituting this into Eq. S5, we get:

$$\psi_A \approx \int_0^{+\infty} dy \ e^{-y} \ \delta\tau \left\{ \ln\left(\frac{N_A + N_B}{y}\right) - 1 \right\}$$

$$= \delta\tau \left\{ \ln(N_A + N_B) - 1 + \gamma \right\} \ , \tag{S7}$$

where $\gamma \equiv - \int_0^{+\infty} dy\, e^{-y} \ln y = 0.57721...$ is Euler's constant. Thus, ψ_A in this case roughly scales as $\ln(N_A + N_B)$.

(1-2) Shift of normal distribution

Next let us consider a case where the measure x is governed by a normal distribution and the mutation shifts the mean of the distribution:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) , \quad \text{and} \quad f(x) + \delta f(x) = f(x - \delta m) .$$

Here, we rescaled and shifted x so that its mean and variance become 0 and 1, respectively, for the wile-type.

In this case, $\delta \ln f(x) \approx \delta m\, x$, and

$$y(x) = (N_A + N_B) \int_x^{+\infty} \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \approx \frac{N_A + N_B}{\sqrt{2\pi}\, x} e^{-\frac{x^2}{2}} .$$

The right-most hand side is the leading term of an asymptotic expansion for $x \gg 1$. Solving it for x iteratively, we get:

$$
\begin{aligned}
x(y) \;\approx\;& \sqrt{2 \ln\left(\frac{N_A + N_B}{\sqrt{2\pi}\, y\, x(y)}\right)} \\[2mm]
\approx\;& \sqrt{2 \ln\left(\frac{N_A + N_B}{\sqrt{2\pi \times 2 \ln\left(\frac{N_A+N_B}{\sqrt{2\pi}\, y\, x(y)}\right)}}\right) - 2 \ln y} \\[2mm]
\approx\;& \sqrt{2 \ln\left(\frac{N_A + N_B}{\sqrt{2\pi \left\{2 \ln\left(\frac{N_A+N_B}{\sqrt{2\pi}\, x(y)}\right) - 2 \ln y\right\}}}\right) - 2 \ln y} \\[2mm]
\approx\;& \sqrt{2 \ln\left(\frac{N_A + N_B}{\sqrt{4\pi\, \ln\left(\frac{N_A+N_B}{\sqrt{4\pi}}\right)}}\right) - \frac{\ln y}{\sqrt{2 \ln\left(\frac{N_A+N_B}{\sqrt{4\pi \cdot \ln\left(\frac{N_A+N_B}{\sqrt{4\pi}}\right)}}\right)}}} . \quad \text{(S8)}
\end{aligned}
$$

Substituting the last approximation into Eq. S5, we have:

$$
\begin{aligned}
\psi_A &\approx \int_0^{+\infty} dy \ e^{-y} \ \delta m \ x(y) \\
&\approx \delta m \ \sqrt{2\ln\left(\frac{N_A + N_B}{\sqrt{4\pi \ \ln\left(\frac{N_A+N_B}{\sqrt{4\pi}}\right)}}\right)} \left\{ 1 + \frac{\gamma}{2\ln\left(\frac{N_A+N_B}{\sqrt{4\pi \ \ln\left(\frac{N_A+N_B}{\sqrt{4\pi}}\right)}}\right)} \right\} .
\end{aligned} \tag{S9}
$$

Because $\ln(N_A + N_B)$ is fairly small compared to $N_A + N_B (\gg 1)$ yet considerably larger than 1, we see that the competitive advantage ψ_A roughly scales as $\sqrt{\ln(N_A + N_B)}$ in this case.

(1-3) Variance increase in normal distribution

Once again, we assume that the measure x behaves according to a normal distribution: $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$. This time, we consider that the mutant increased the variance of the distribution:

$$
f(x) + \delta f(x) = \frac{1 - \delta\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\{(1-\delta\sigma)x\}^2}{2} \right) .
$$

In this case, the dummy variable y, and consequently the function $x(y)$ as well, are the same as in the last subsection. Regarding $\delta \ln f(x)$, we have:

$$
\delta \ln f(x) \approx \delta\sigma(x^2 - 1) .
$$

Substituting these approximations into Eq. S5, we get:

$$
\begin{aligned}
\psi_A &\approx \int_0^{+\infty} dy \ e^{-y} \ \delta\sigma \left[(x(y))^2 - 1 \right] \\
&\approx \delta\sigma \left\{ 2\ln\left(\frac{N_A + N_B}{\sqrt{4\pi \ \ln\left(\frac{N_A+N_B}{\sqrt{4\pi}}\right)}}\right) + 2\gamma - 1 \right\} .
\end{aligned} \tag{S10}
$$

Thus, the competitive advantage ψ_A roughly scales as $\ln(N_A + N_B)$ in this case.

K. Ezawa and H. Innan

(1-4) Increased maximum of an upper-bounded distribution

So far, the distribution of the measure x for the sperm performance was either exponential or normal, neither of which is bounded from above.

In actual sperm competitions, however, it may be more natural to assume a performance measure bounded from above by a positive maximum value. We could consider that all characters influencing the sperm competitiveness could be integrated into a single measure, which we call the "velocity", which is the reciprocal of the total time from the start (ejaculation) till the completion of the fertilization. No matter how good the sperm performance is, the total time remain finite and can never be zero, thus there must always be a finite non-zero upper bound in the "velocity". Considering this way, a natural form of the probability distribution $f(x)$ near the upper-boundary x_M would be:

$$f(x) \propto \begin{cases} (\alpha + 1)(x_M - x)^\alpha & \text{for} \quad x < x_M , \\ 0 & \text{for} \quad x \geq x_M , \end{cases} \tag{S11}$$

where the exponent $\alpha > 0$ determines the steepness of the distribution. At this point, the asymptotic distribution (Eq. S11) still has a freedom of a multiplication factor. For later convenience, we choose such a factor that the functional form should be valid in the entire region, $0 \leq x$, and we also rescale x so that x_M will be 1:

$$f(x) = \begin{cases} (\alpha + 1)(1 - x)^\alpha & \text{for} \quad 0 \leq x < 1 , \\ 0 & \text{for} \quad x \geq 1 . \end{cases} \tag{S12}$$

In this case, the dummy integration variable becomes:

$$y = \begin{cases} (N_A + N_B)(1 - x)^{\alpha+1} & \text{for} \quad x < 1 , \\ 0 & \text{for} \quad x \geq 1 . \end{cases}$$

Because y is zero all across $x \geq 1$, the second equation in Eq. S5 needs a slight modification if $\delta f(x) > 0$ in $x > 1$:

$$\psi_A = \int_0^{+\infty} dy \, e^{-y} \, \delta \ln f(x(y)) + (N_A + N_B) \int_1^{+\infty} dx \, (f(x) + \delta f(x)) . \tag{S13}$$

Now, let us consider a particular case where the mutation slightly widen the region of x by increasing the upper-bound:

$$f(x) + \delta f(x) = \frac{\alpha + 1}{1 + \delta x_M} \left(1 - \frac{x}{1 + \delta x_M} \right)^{\alpha}.$$

Then, for $x < 1$, we have

$$\delta \ln f(x) \approx \delta x_M \left(\frac{\alpha \, x}{1 - x} - 1 \right) \approx \delta x_M \left[\alpha \left(\frac{N_A + N_B}{y} \right)^{\frac{1}{\alpha + 1}} - (\alpha + 1) \right]. \qquad \text{(S14)}$$

Substituting the above two equations into Eq. S13 yields:

$$
\begin{aligned}
\psi_A \;\approx\;& \delta x_M \int_0^{+\infty} dy \; e^{-y} \left[\alpha \left(\frac{N_A + N_B}{y} \right)^{\frac{1}{\alpha + 1}} - (\alpha + 1) \right] \\
& + (N_A + N_B) \int_1^{1 + \delta x_M} dx \; \frac{\alpha + 1}{1 + \delta x_M} \left(1 - \frac{x}{1 + \delta x_M} \right)^{\alpha} \\
\;=\;& \delta x_M \left[(N_A + N_B)^{\frac{1}{\alpha + 1}} \; \alpha \; \Gamma \left(\frac{\alpha}{\alpha + 1} \right) - (\alpha + 1) \right] + (N_A + N_B) \left(\frac{\delta x_M}{1 + \delta x_M} \right)^{\alpha + 1} \\
\;\approx\;& \delta x_M (N_A + N_B)^{\frac{1}{\alpha + 1}} \; \alpha \; \Gamma \left(\frac{\alpha}{\alpha + 1} \right). \qquad \text{(S15)}
\end{aligned}
$$

The last approximation holds because we are now considering δx_M that is small enough to give $\delta x_M \, (N_A + N_B)^{\frac{1}{\alpha + 1}} \ll 1$, and because we now consider $(N_A + N_B)^{\frac{1}{\alpha + 1}} \gg 1$. The approximate equation Eq. S15 states that the competitive advantage ψ_A roughly scales as $(N_A + N_B)^{\frac{1}{\alpha + 1}}$ in this case.

K. Ezawa and H. Innan

Note 2. Calculation of Advantage in One-on-One Competition

Here we derive a formula for competitive advantage in one-on-one competitions, then apply it to the aforementioned four particular cases. They will serve as a basis for assessing the enhancement of the advantage by fierce competitions among numerous sperm.

(2-0) General formula

As in the main text (or in Supplementary Materials and Methods), consider the case where the mutation from B to A changed the distribution only slightly:

$$f_B(x) = f(x) , \quad f_A(x) = f(x) + \delta f(x) .$$

Then, the probability that allele A wins in a one-on-one competition with B is:

$$
\begin{aligned}
P\left[\text{winner} = A \mid 1\ A\&\ 1\ B\right] &= \int_{-\infty}^{+\infty} dx \left[(f(x) + \delta f(x)) \int_{-\infty}^{x} d\xi f(\xi) \right] \\
&= \frac{1}{2} + \int_{-\infty}^{+\infty} dx \left[\delta f(x) \int_{-\infty}^{x} d\xi f(\xi) \right] .
\end{aligned}
\tag{S16}
$$

If we set

$$\psi_A \equiv 4 \int_{-\infty}^{+\infty} dx \left[\delta f(x) \int_{-\infty}^{x} d\xi f(\xi) \right] , \tag{S17}$$

the above equation is rearranged as:

$$
\begin{aligned}
P\left[\text{winner} = A \mid 1\ A\&\ 1\ B\right] &= \frac{1}{2}\left(1 + \frac{1}{2}\psi_A\right) \\
&\approx \frac{1}{1 + (1 - \psi_A)} \\
&\approx \frac{1 + \psi_A}{(1 + \psi_A) + 1} .
\end{aligned}
\tag{S18}
$$

Thus ψ_A is interpreted as the competitive advantage of allele A over allele B. [Actually, these equations are special cases of Eqs. S3, S4 and S6 when $N_A = N_B = 1$.] Now we will calculate Equation S17 for specific cases.

(2-1) Increased mean in exponential distribution

In this case, $f(x) = \exp(-x)$ and $\delta f(x) \approx \delta\tau \; (x-1)\exp(-x)$ (for $x \geq 0$). Substituting them into Equation S17, we have:

$$
\begin{aligned}
\psi_A &\approx 4\int_0^{+\infty} dx \left[\delta\tau \; (x-1)\exp(-x) \int_0^x d\xi \exp(-\xi)\right] \\
&= 4\,\delta\tau \int_0^{+\infty} dx \left[(x-1)\exp(-x)(1-\exp(-x))\right] \\
&= 4\,\delta\tau \left[\Gamma(2) - \Gamma(1) - \left(\frac{1}{2}\right)^2\Gamma(2) + \left(\frac{1}{2}\right)\Gamma(1)\right] \\
&= \delta\tau \; .
\end{aligned}
\tag{S19}
$$

(2-2) Shift of normal distribution

In this case, $f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ and $\delta f(x) \approx \delta m \; x f(x) = -\delta m \; \frac{d}{dx}f(x)$. Substituting them into Equation S17, we have:

$$
\begin{aligned}
\psi_A &\approx 4\int_{-\infty}^{+\infty} dx \left[-\delta m \frac{d}{dx}f(x)\int_{-\infty}^x d\xi f(\xi)\right] \\
&= -4\,\delta m \left[f(x)\int_{-\infty}^x d\xi f(\xi)\right]_{x=-\infty}^{x=+\infty} + 4\,\delta m \int_{-\infty}^{+\infty} dx \, (f(x))^2 \\
&= 4\,\delta m \int_{-\infty}^{+\infty} dx \, \{f(x)\}^2 \; .
\end{aligned}
$$

Actually, equations up to this point hold for an infinitesimal constant shift *of any distribution* that is differentiable in the interval $-\infty < x < +\infty$. Now, substituting $f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ into the rightmost hand side, we get:

$$
\begin{aligned}
\psi_A &\approx 4\frac{\delta m}{2\pi}\int_{-\infty}^{+\infty} dx \exp(-x^2) \\
&= \frac{2}{\sqrt{\pi}}\,\delta m \; .
\end{aligned}
\tag{S20}
$$

K. Ezawa and H. Innan

(2-3) Variance increase in normal distribution

In this case, $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$ and $\delta f(x) \approx \delta\sigma(x^2 - 1)f(x) = \delta\sigma\left(\frac{d}{dx}\right)^2 f(x)$. Substituting them into Equation S17, we have:

$$
\begin{aligned}
\psi_A &\approx 4\int_{-\infty}^{+\infty} dx \left[\delta\sigma\left(\frac{d}{dx}\right)^2 f(x) \int_{-\infty}^{x} d\xi f(\xi)\right] \\
&= 4\,\delta\sigma\left[\frac{d}{dx}f(x)\int_{-\infty}^{x} d\xi f(\xi)\right]_{x=-\infty}^{x=+\infty} - 4\,\delta\sigma\int_{-\infty}^{+\infty} dx \left(\frac{d}{dx}f(x)\right)f(x)\ . \\
&= -4\,\delta\sigma\int_{-\infty}^{+\infty} dx \left[\frac{1}{2}\frac{d}{dx}\left(f(x)\right)^2\right] \\
&= -4\,\frac{\delta\sigma}{2}\left[\left(f(x)\right)^2\right]_{x=-\infty}^{x=+\infty} \\
&= 0\ .
\end{aligned}
\tag{S21}
$$

Therefore, just increasing the variance of a normal distribution gives no competitive advantage (of $O(\delta\sigma)$) to the mutant as far as one-on-one competitions are concerned.

(2-4) Increased maximum of an upper-bounded distribution

In this case, $f(x) = (\alpha + 1)(1 - x)^\alpha$ (for $0 \le x \le 1$) and $\delta f(x) \approx \delta x_M \left(\frac{\alpha x}{1-x} - 1\right) f(x)$. Substituting them into Equation S17, we have:

$$
\begin{aligned}
\psi_A &\approx 4\int_{0}^{1} dx \left[\delta x_M \left(\frac{\alpha x}{1 - x} - 1\right)(\alpha + 1)(1 - x)^\alpha \int_{0}^{x} d\xi\,(\alpha + 1)(1 - \xi)^\alpha\right] \\
&= 4\,\delta x_M\,(\alpha + 1)\int_{0}^{1} dx \left\{\left[\alpha x(1 - x)^{\alpha-1} - (1 - x)^\alpha\right]\left[1 - (1 - x)^{\alpha+1}\right]\right\} \\
&= 4\,\delta x_M(\alpha + 1)\left[\alpha\,B(2,\,\alpha) - \frac{1}{\alpha + 1} - \alpha\,B(2, 2\alpha + 1) + \frac{1}{2\alpha + 2}\right] \\
&= 4\,\delta x_M\left[\alpha(\alpha + 1)\frac{1}{\alpha(\alpha + 1)} - 1 - \alpha(\alpha + 1)\frac{1}{(2\alpha + 1)(2\alpha + 2)} + \frac{1}{2}\right] \\
&= \frac{2(\alpha + 1)}{2\alpha + 1}\,\delta x_M\ .
\end{aligned}
\tag{S22}
$$

Note 3. Enhancement Factor of Sperm Competitive Advantage

Now that we know the competitive advantage both for one-on-one competition and for competition among numerous competitors, we can calculate the enhancement factor for the specific cases.

(3-1) Increased mean in exponential distribution

In this case, $\psi_A(N_A = N_B = 1) \approx \delta\tau$, and $\psi_A(N_A, N_B \gg 1) \approx \delta\tau \{\ln(N_A + N_B) - 1 + \gamma\}$. Thus, we have:

$$R[\psi_A](N_A, N_B) \equiv \frac{\psi_A(N_A, N_B)}{\psi_A(N_A = N_B = 1)} \approx \ln(N_A + N_B) - 1 + \gamma . \qquad \text{(S23)}$$

(3-2) Shift of normal distribution

In this case, $\psi_A(N_A = N_B = 1) \approx (2\delta m)/\sqrt{\pi}$ and

$$\psi_A(N_A, N_B \gg 1) \approx \delta m \sqrt{2 \ln \left(\frac{N_A + N_B}{\sqrt{4\pi \ \ln \left(\frac{N_A + N_B}{\sqrt{4\pi}} \right)}} \right)} .$$

Taking the ratio of these two yields:

$$R[\psi_A](N_A, N_B) \approx \sqrt{\frac{\pi}{2} \ \ln \left(\frac{N_A + N_B}{\sqrt{4\pi \ \ln \left(\frac{N_A + N_B}{\sqrt{4\pi}} \right)}} \right)} . \qquad \text{(S24)}$$

(3-3) Variance increase in normal distribution

In this case, $\psi_A(N_A = N_B = 1) \approx 0$ and

$$\psi_A(N_A, N_B \gg 1) \approx \delta\sigma \left\{ 2 \ln \left(\frac{N_A + N_B}{\sqrt{4\pi \ \ln \left(\frac{N_A + N_B}{\sqrt{4\pi}} \right)}} \right) + 2\gamma - 1 \right\} .$$

K. Ezawa and H. Innan

Thus, $R[\psi_A]$ is $+\infty$, because there is no advantage in an one-on-one competition.

(3-4) Increased maximum of an upper-bounded distribution

In this case, $\psi_A(N_A = N_B = 1) \approx \delta x_M \times 2(\alpha + 1)/(2\alpha + 1)$ and

$$\psi_A(N_A, N_B \gg 1) \approx \delta x_M \, (N_A + N_B)^{\frac{1}{\alpha+1}} \, \alpha \, \Gamma\left(\frac{\alpha}{\alpha+1}\right) \ .$$

Taking the ratio, we have:

$$
\begin{aligned}
R[\psi_A](N_A, N_B) &\approx (N_A + N_B)^{\frac{1}{\alpha+1}} \frac{\alpha(2\alpha + 1)}{2(\alpha + 1)} \Gamma\left(\frac{\alpha}{\alpha+1}\right) \\
&= (N_A + N_B)^{\frac{1}{\alpha+1}} \frac{2\alpha + 1}{2} \Gamma\left(\frac{2\alpha + 1}{\alpha + 1}\right) \ . \quad\quad \text{(S25)}
\end{aligned}
$$

Note 4. Population Genetic Behavior of Mutant Frequency under Paternity-Sharing Sperm Competition

In the previous subsection, we examined the effect of an allelic difference on each instance of sperm competition, which we expressed in terms of the probability that a better allele will win. In population genetics, such competitions will take place here and there throughout the population. Thus, we expect that even a small competitive advantage could accumulate through generations to make a big difference.

Here, we want to focus on the effect of competitions among sperm sharing paternity, which have been overlooked in the previous studies. For this purpose, we consider an extreme situation in which a population consists of individuals that are strictly monogamous (and especially mono-androus). In this situation, there will *never* be post-copulatory competitions, including sperm competitions, *between different males*.

As in the previous section (or in the main text), we focus on a single locus (or site) and assume that the locus has two alleles, A and B, which are selectively neutral except in sperm competition. Here the locus is assumed to be on an autosome. We also assume that there are no further mutations at the locus (or site) and that the two alleles were present from the beginning in the current generation. Let $P^{(P)}[Z \mid Z_1 Z_2]$ denote the probability that a sperm with allele Z (= A or B) wins a successful competition among sperm ejaculated by a male individual with the genotype $Z_1 Z_2$ (Z_1, Z_2 = A or B). Obviously, for homozygous males, we have:

$$
\begin{aligned}
P^{(P)}[A \mid AA] &= P^{(P)}[B \mid BB] = 1 \;, \\
P^{(P)}[A \mid BB] &= P^{(P)}[B \mid AA] = 0 \;.
\end{aligned}
\tag{S26}
$$

For heterozygous males, we can use Eq. S6. Assuming that there are an equal number of sperm with alleles A and B , $N_A = N_B$, and assuming that the allele difference has only a

K. Ezawa and H. Innan

small effect, we have:

$$P^{(P)}[A \mid AB] = \frac{1}{2}\left(1 + \frac{\psi_A}{2}\right), \quad P^{(P)}[B \mid AB] = \frac{1}{2}\left(1 - \frac{\psi_A}{2}\right). \tag{S27}$$

Let $p_t^{(P)}(Z_1 Z_2)$ be the frequency of paternal genomes with the genotype $Z_1 Z_2$ (Z_1, Z_2 = A or B) at the locus in the current (*i.e.*, the t-th) generation. Then, the expected frequency, $p_{t+1}^{(P)}(Z)$, of allele Z (= A or B) *of paternal origin* at the next (*i.e.*, the $(t+1)$-th) generation is in general:

$$p_{t+1}^{(P)}(Z) = \sum_{Z_1 Z_2 = AA, AB, BB} P^{(P)}[Z \mid Z_1 Z_2]\, p_t^{(P)}(Z_1 Z_2).$$

This equation, after substituting Eq. S26 and Eq. S27 into it, reduces to:

$$p_{t+1}^{(P)}(A) = p_t^{(P)}(AA) + \frac{1}{2}\left(1 + \frac{\psi_A}{2}\right) p_t^{(P)}(AB), \tag{S28}$$
$$p_{t+1}^{(P)}(B) = 1 - p_{t+1}^{(P)}(A).$$

In the deterministic limit, the diploid frequencies $p_t^{(P)}(Z_1 Z_2)$ are given by the Hardy-Weinberg principle (see e.g. section 2.2 of **?**):

$$p_t^{(P)}(AA) = p_t^{(P)}(A)\, p_t^{(M)}(A),$$
$$p_t^{(P)}(AB) = p_t^{(P)}(A)\, p_t^{(M)}(B) + p_t^{(P)}(B)\, p_t^{(M)}(A),$$
$$p_t^{(P)}(BB) = p_t^{(P)}(B)\, p_t^{(M)}(B), \tag{S29}$$

where $p_t^{(M)}(Z)$ is the frequency of allele Z *of maternal origin* at the current (i.e. the t-th) generation. Substituting Eq. S29 into Eq. S28, we get:

$$p_{t+1}^{(P)}(A) = \frac{1}{2}\left\{p_t^{(P)}(A) + p_t^{(M)}(A)\right\} + \frac{\psi_A}{4}\left\{p_t^{(P)}(A)\, p_t^{(M)}(B) + p_t^{(P)}(B)\, p_t^{(M)}(A)\right\}, \tag{S30}$$

and $p_{t+1}^{(P)}(B) = 1 - p_{t+1}^{(P)}(A)$.

Let us next consider the evolution of the maternal allele frequency. If we assume that the alleles A and B have the same probability of transmission to the next generation, the reasoning leading to paternal allele frequency also applies here, with $\psi_A = 0$. The result is:

$$p_{t+1}^{(M)}(A) = \frac{1}{2}\left\{p_t^{(P)}(A) + p_t^{(M)}(A)\right\}, \tag{S31}$$

and $p_{t+1}^{(M)}(B) = 1 - p_{t+1}^{(M)}(A)$.

Taking the arithmetic mean of Eq. S30 and Eq. S31, and ignoring terms of $O((\psi_A)^2)$, we get a simple recursion relation:

$$p_{t+1}(A) = p_t(A) + \frac{\psi_A}{4} p_t(A)(1 - p_t(A)) , \tag{S32}$$

and $p_{t+1}(B) = 1 - p_{t+1}(A)$. In the main text, they are Eqs. 9 and 10, respectively. Here,

$$p_t(Z) \equiv \frac{1}{2} \left\{ p_t^{(P)}(Z) + p_t^{(M)}(Z) \right\} ,$$

is the gender-averaged frequency of allele Z (= A or B) at the t-th generation.

Although we have ignored genetic drift so far, taking account of genetic drift is not so difficult. For this purpose, it is sufficient to notice that our recursion equation, Eq. S32, is equivalent to the deterministic recursion equation of the allele frequency:

$$p_{t+1}(A) = p_t(A) + s \, p_t(A)(1 - p_t(A)) ,$$

when allele A has a selective advantage of s ($\ll 1$) over allele B. Therefore, the diffusion theory framework such as unfolded in section 8.8.3 of (**?**) applies also here, if s is replaced by $\frac{\psi_A}{4}$. Thus we have the fixation probability $u(p)$ of allele A when its initial frequency is p:

$$u(p) \approx \frac{1 - \exp(-N_e \psi_A \, p)}{1 - \exp(-N_e \psi_A)} , \tag{S33}$$

where N_e is the effective population size. This is Eq. 11 in the main text. The initial frequency of a new mutation should be $p = 1/(2N)$, where N is the actual population size. If $N_e = N$, the equation is approximated as:

$$u(p) \approx \frac{\psi_A/2}{1 - \exp(-N\psi_A)} , \tag{S34}$$

which in turn reduces to $u(p) \approx \psi_A/2$ when $\exp(N\psi_A) \gg 1$. The effect of sperm competitive advantage is 1/4-fold smaller than that of selective advantage of the same intensity (i.e. when $s = \psi_A$). A multiplicative factor of 1/2 comes from the neutrality of the alleles in the maternal transmission, and the other multiplicative factor of 1/2 originates from the fact that thecompetition among paternity-sharing sperm is effective only when the male is heterozygous.

K. Ezawa and H. Innan