

X 染色体上の遺伝子および半倍数性生物における連鎖不平衡と集団の有効な大きさ

平成 28 年 5 月 19 日受付

野村 哲郎^{1,2)}

高橋 純一^{1,2)}

¹⁾ 京都産業大学総合生命科学部

²⁾ 京都産業大学ミツバチ産業科学研究センター

要 旨

X 染色体および半倍数性生物の遺伝子座間の連鎖不平衡から、過去の集団の有効な大きさ (N_e) を推定するための理論を開発した。得られた理論の応用として、セイヨウミツバチ (*Apis mellifera*) の N_e の歴史的変化を推定した。

キーワード：X 染色体、半倍数性生物、連鎖不平衡、集団の有効な大きさ、セイヨウミツバチ

1. はじめに

連鎖不平衡の大きさ (LD) は、集団の有効な大きさ (N_e) の推定に利用することができる^{1,2,3)}。一般に、小さな N_e は相対的に高い LD を生じる。さらに、強く連鎖した遺伝子座間の LD は過去の N_e に関する情報を与え、逆に弱く連鎖した遺伝子座間の LD は比較的最近の N_e に関する情報を与える^{2,3)}。したがって、ある時点でのサンプルから過去の N_e の変化を推定することができる。

LD からの N_e の推定理論は、これまでに主として常染色体上の遺伝子について開発され、高密度の SNP チップが利用できるヒトや家畜について、 N_e の歴史的変化が明らかにされてきた^{3,4)}。それらの研究から、ヒトでは過去から現在に至るまでに人口の爆発的増加を反映して N_e に急激な拡大が見られる^{3,4)}。これに対して、多くの家畜種では家畜化に伴うボトルネック効果および家畜化後の選抜によって N_e の急激な縮小が報告されている^{3,4)}。

X 染色体上の遺伝子の LD と N_e については、理論ならびにその応用研究はほとんどなされ

てこなかった。X染色体上の遺伝子伝達は半倍数性生物における遺伝子伝達と形式的に同一であり、X染色体上の遺伝子の LD と N_e に関する理論はミツバチをはじめとする半倍数性生物の進化を説明する上で有効である。

本研究では、X染色体上の遺伝子の LD から N_e を推定するための理論を導くとともに、その応用としてセイヨウミツバチ (*Apis mellifera*) の N_e の歴史的变化について一知見を述べる。

2. 連鎖不平衡の定常値

(1) 仮定と記号の定義

雌はXX、雄はXYの性決定様式を持つ生物を想定し、X染色体上の2つの遺伝子座間の LD について考える。X染色体上の遺伝子の伝達様式(図1)は、半倍数性生物における遺伝子の伝達様式と同一であるので、導かれた理論は半倍数生物にも適用できる。用いる主な記号とその定義を表1に示す。

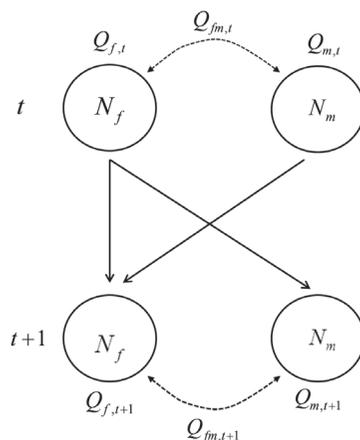


図1 遺伝子の伝達様式

表1 記号とその定義

記号	定義
m, f (添字)	m = 雄親、 f = 雌親
t (添字)	世代
N	個体数
Q	一方の遺伝子座上の2つの遺伝子が同祖 (identical by descent) であるとき、もう一方の遺伝子座上の2つの遺伝子が同一の経路で同祖となる条件付き確率
r	2つの遺伝子座上の遺伝子の頻度の相関係数 $E[r^2] = Q$
c	組換え価 $c = \frac{2}{3} c_f$

(2) 定常値の導出

連続した2世代 ($t, t+1$) のLD は、図1を参考にすれば、

$$\left. \begin{aligned} Q_{f,t+1} &= \frac{1}{4} \left[\frac{1}{2N_f} + \left(1 - \frac{1}{2N_f}\right) (1-c_f)^2 Q_{f,t} \right] \\ &\quad + \frac{1}{4} \left[\frac{1}{N_m} + \left(1 - \frac{1}{N_m}\right) Q_{m,t} \right] + \frac{1}{2} (1-c_f) Q_{fm,t} \\ Q_{m,t+1} &= \frac{1}{2N_f} + \left(1 - \frac{1}{2N_f}\right) (1-c_f)^2 Q_{f,t} \\ Q_{fm,t+1} &= \frac{1}{2} \left[\frac{1}{2N_f} + \left(1 - \frac{1}{2N_f}\right) (1-c_f)^2 Q_{f,t} \right] + \frac{1}{2} (1-c_f) Q_{fm,t} \end{aligned} \right\} \quad (1)$$

と表される。式 (1) において、

$$\begin{aligned} x &= \left(1 - \frac{1}{2N_f}\right) (1-c_f)^2 \approx 1 - 2c_f - \frac{1}{2N_f} \\ y &= 1 - \frac{1}{N_m} \\ z &= 1 - c_f \end{aligned}$$

とおけば、近似的に

$$\begin{bmatrix} Q_{f,t+1} \\ Q_{m,t+1} \\ Q_{fm,t+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{4}x & \frac{1}{4}y & \frac{1}{2}z \\ x & 0 & 0 \\ \frac{1}{2}x & 0 & \frac{1}{2}z \end{bmatrix} \begin{bmatrix} Q_{f,t} \\ Q_{m,t} \\ Q_{fm,t} \end{bmatrix} + \begin{bmatrix} \frac{1}{8N_f} + \frac{1}{4N_m} \\ \frac{1}{2N_f} \\ \frac{1}{4N_f} \end{bmatrix}$$

あるいは

$$\mathbf{q}_{t+1} = \mathbf{C}\mathbf{q}_t + \mathbf{n}$$

が得られる。定常状態での \mathbf{q} を \mathbf{q}_∞ とすれば、

$$\mathbf{q}_\infty = (\mathbf{I} - \mathbf{C})^{-1} \mathbf{n} \quad (2)$$

である。ここで、 \mathbf{I} は 3×3 の単位行列である。式 (2) より、 Q の定常値は

$$Q_{f,\infty} = Q_{m,\infty} = \frac{1}{\frac{8}{3}N_e c_f + 1}$$

として求められる。ここで、

$$N_e = \frac{9N_f N_m}{2N_f + 4N_m}$$

は X 染色体上の遺伝子あるいは半倍数性生物の集団の有効な大きさである⁵⁾。 r の定常値を r_∞

とすれば、 $E[r_\infty^2] = Q_{f,\infty} = Q_{m,\infty}$ であるから¹⁾、

$$E[r_\infty^2] = \frac{1}{\frac{8}{3}N_e c_f + 1} \quad (3)$$

より、観測された連鎖不平衡 r^2 が定常状態に達しているものと仮定して N_e が推定できる。

図2は、コンピュータシミュレーションによって発生させた種々の組換え価 c_f における r^2 の定常値 (sim) と式 (3) から得られた定常値の近似解 (approx) を比較したものである。図には、式 (1) から求めた定常値の厳密解 (雌: exact_f、雄: exact_m) も合わせて示してある。式 (3) による近似解は厳密解とよく一致しており、またシミュレーション値にも極めて近い値を与えている。

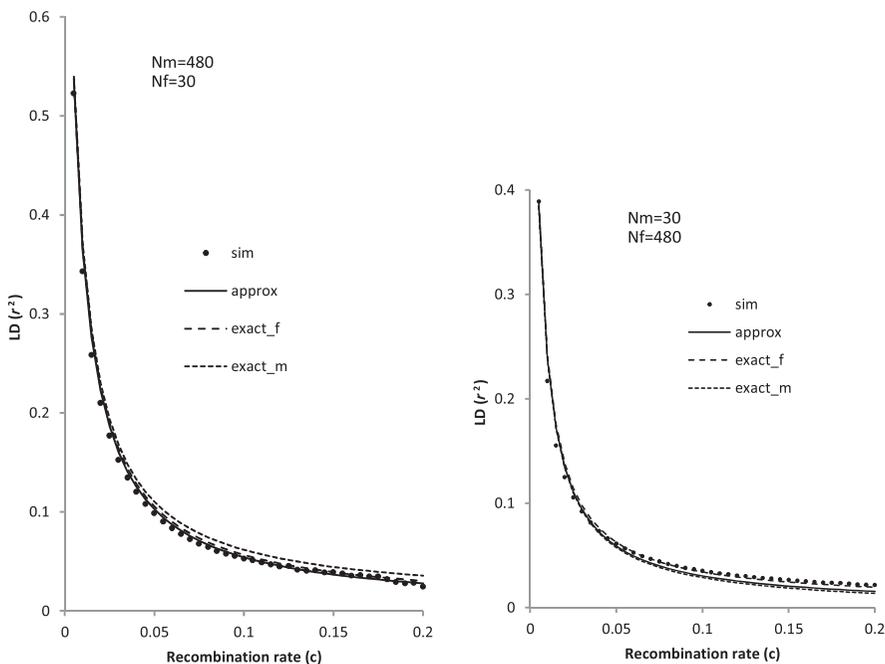


図2 種々の組換え価における r^2 の定常値のシミュレーションでの観測値 (sim)、式 (3) による近似解 (approx) および式 (1) による厳密解 (exact_f, exact_m)

3. 過去の N_e の推定

現在を世代0として、 t 世代前の N_e が t の線形関数 $2N_{e,t} = \alpha + \beta t$ であるものとする。Hayes ら³⁾にしたがえば、 r^2_{∞} は近似的に

$$r^2_{\infty} = \frac{1}{1 + \beta + 2c\alpha} \quad (4)$$

と表される。ここで、 c は雌雄をプールした組換え価である。雄の X 染色体あるいは半倍数性生物の雄では組換えは生じないので、 $c = \frac{2}{3}c_f$ である。式 (3) と (4) より、

$$\frac{1}{1 + \beta + \frac{2}{3}c_f\alpha} = \frac{1}{1 + \frac{8}{3}N_{e,t}c_f}$$

であるから、

$$2N_{e,t} = \frac{3}{4c_f} \beta + \alpha = \alpha + \beta t$$

を得る。したがって、定常状態を仮定して式 (3) から得られた N_e の推定値は、

$$t = \frac{3}{2c_f} \quad (5)$$

世代前の N_e の推定値を与える。

図3は、集団の有効な大きさが実線で示したような変化をしたときに、1Morganの染色体上に均等に配置した201個の遺伝子座に生じる連鎖不平衡をコンピュータシミュレーションにより発生させ、式(5)からの世代の推定値に対する式(3)からの N_e の推定値をプロットしたものである。(a) および (b) は、 N_e が線形に増加あるいは減少した場合である。また (c) は指数関数的に N_e が増加あるいは減少した場合である。いずれの場合も、式(5)による世代の推定精度は高いとは言えないが、少なくとも N_e が過去に増加傾向にあったのか、減少傾向にあったのかはとらえられている。

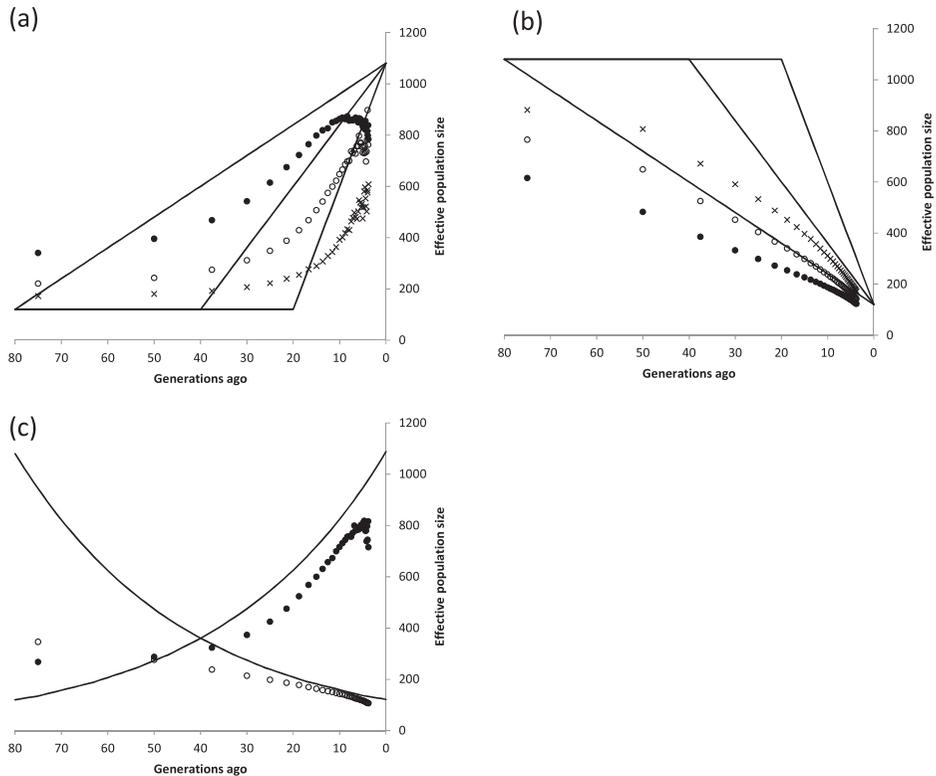


図3 集団の有効な大きさが実線の変化をしたとき、式(3)と式(5)からの変化の予測。(a) 線形増加、(b) 線形減少、(c) 指数関数的増加および減少。

4. セイヨウミツバチの N_e の歴史的変化の推定

野生動物の家畜化 (domestication) は、野生集団の一部を利用して行われるため、家畜化にともない N_e が急激に減少し遺伝的多様性が低下する⁶⁾。このような現象は、家畜化にともなうボトルネック (domestication bottleneck) と呼ばれている。さらに、家畜化後の近親交配や選抜によって、家畜集団の遺伝的多様性は低下し続けている。近年、高密度 SNP チップが利用できるウシ、ブタ、ニワトリなどの家畜では、連鎖不平衡から N_e の歴史的変化を調べる研究が行われており、家畜化にともなうボトルネックと家畜化後の N_e の継続した低下が報告されている^{3,4)}。

エジプトやメソポタミアでは、少なくとも紀元前 7,000 年にはミツバチを家畜化してハチミツ、蜜ろうなどを利用していたことを洞窟壁画から知ることができる^{7,8)}。現在、養蜂に用いられているセイヨウミツバチ (*Apis mellifera*) の多くの群は、ヨーロッパのセイヨウミツバチ群に起源を持つと考えられている。ヨーロッパのセイヨウミツバチ群は、西部および北部ヨーロッパの亜種からなる M グループと東部および南部ヨーロッパの亜種からなる C グループに分けられる。これら 2 つのグループは、氷河期の後にアフリカから独立にヨーロッパに侵入したものとされている^{8,9)}。現在の商業ベースで養蜂に利用されているセイヨウミツバチ群は、これら 2 系統の混合 (admixture) 集団であると考えられており、DNA 解析の結果もそれを支持している^{7,8)}。

ここでは式 (3) と (5) を用いて、セイヨウミツバチの N_e の歴史的変化を調べた。材料は、Wallberg ら⁸⁾ が報告した連鎖不平衡のデータ (Supplementary Figure 4) である。このデータには、M および C グループに加えて、セイヨウミツバチの祖先と考えられるアフリカの亜種からなる A グループ、中近東の亜種からなる O グループ、さらにアメリカミネソタ州の養蜂群 (Managed US) とスウェーデンの養蜂群 (Managed EU) についての連鎖不平衡の推定値が含まれる。Wallberg ら⁸⁾ にしたがって、世代間隔は 1 年と仮定した。ミツバチでは他の動物種に比べて、全ゲノムにわたって例外的高い頻度で組換えが起こることが知られている。ここでは Beye ら¹⁰⁾ の報告から、1Mb あたり 23.2cM の組換え価を採用した。解析結果は図 4 に示すとおりである。A グループが最も大きな N_e を示した。このグループから生じたと考えられる M および C グループの N_e は、A グループの 4 分の 1 以下であり、これら 2 つのグループの成立に際して強いボトルネック効果が働いたことが示唆される。中近東の亜種からなる O グループの N_e は、A グループよりもはるかに小さな値を示した。セイヨウミツバチの起源を O グループとする説もあるが⁹⁾、 N_e の大きな A グループが N_e の小さな O グループから生じたとは考え難く、セイヨウミツバチの起源を O グループとする説は今回の結果からは支持されない。養蜂群の 2 つサンプルの N_e は、いずれも過去から現在に至るまで最も小さい値を示し、とくに家畜化 (約 9,000 年前) 以降も、他の家畜に見られるような継続した低下が認められた。

セイヨウミツバチの家畜化およびその後の飼育管理は、他の家畜と同様に N_e および遺伝的多様性の低下を引き起こしたことが報告されてきた¹¹⁾。このような遺伝的多様性の低下にともなう蜂群の脆弱性が、近年問題となっているミツバチの大量死 (CCD: Colony Collapse Disorder) の一因となっている可能性も示唆されている⁹⁾。一方、Harpur ら⁶⁾はカナダおよびフランスのコマーシャル群について 358 個の SNP を解析し、これらの群では他の家畜とは異なり、広範な蜂群の混合 (admixture) が遺伝的多様性の増加をもたらしたことを示している。

本研究における試算では、セイヨウミツバチでも他の家畜と同様に家畜化およびその後の飼育管理によって低下したことを示唆する結果が得られた。ただし、式 (5) を導くにあたって設けた仮定や近似にはさらに検討が必要である。この問題については、今後も継続して検討する予定である。

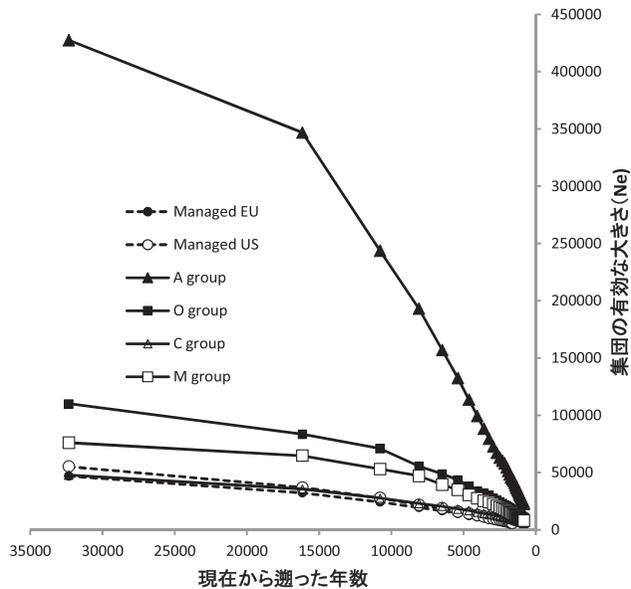


図4 セイヨウミツバチの集団の有効な大きさの歴史的变化の推定

謝 辞

本研究は京都産業大学先端科学技術研究所の研究活動によるものである。また、本研究の一部は、平成 27 年度農林水産業・食品産業科学技術研究推進事業【発展融合ステージ】27013B の支援を受けて行った。

参考文献

- 1) J.A. Sved (1971). *Theor. Pop. Biol.* 2: 125–141.
- 2) W.G. Hill (1981). *Genet. Res.* 38: 209–216.
- 3) B.J. Hayes, P.M. Visscher, H.C. McPartlan, M.E. Goddard (2007). *Genome Res.* 13: 635–643.
- 4) The Bovine HapMap Consortium (2009). *Science* 324: 528–532.
- 5) S. Wright (1933). *Proc. Nat. Acad. Sci. USA* 19: 411–420.
- 6) 野澤 謙 (1975). *日畜会報* 46: 549–557.
- 7) B.A. Harpur, S. Minaei, C.F. Kent, A. Zayed (2012). *Mol. Ecol.* 21: 4414–4421.
- 8) A. Wallberg et al. (2014). *Nature Genet.* 46: 1081–1088.
- 9) B. Oldroyd (2012). *Mol. Ecol.* 21: 4409–4411.
- 10) M. Beye et al. (2016). *Genome Res.* 16: 1339–1344.
- 11) P. De La Rúa, Jafee, I. Munoz, J. Serrano, R.F.A. Moritz, F.B. Kraus (2013). *Mol. Ecol.* 22: 3208–3210.

Linkage disequilibrium and effective population size in X-linked genes and haplodiploids

Tetsuro NOMURA
Jun-ichi TAKAHASHI

Abstract

A theory for estimating historical effective population size in X-linked genes and haplodiploids from linkage disequilibrium was developed. As an application of the obtained theory, historical change in the effective population of honey bees, *Apis mellifera*, was estimated.

Keywords: X chromosome, haplodiploids, linkage disequilibrium, effective population size,
Apis mellifera