

Jost's D と G_{ST} の論争に関するメモ書き

Jost (2008) G_{ST} and its relatives do not measure differentiation. *Mol Ecol* 17, 4015-4026.

遺伝的分化の指標としてよく使われる G_{ST} に物言いをつけた Jost's D に関する論文。ちなみに Jost さんは genetic な指数だけでなく種多様性の指数についてもいろいろ論文・コメントを書いているみたい。

遺伝的分化指数のほとんどは F_{ST} や G_{ST} とその親戚たちを使用
これらの指数は 0 だと遺伝的分化なし、1 だと完全に遺伝的分化

その開祖は

Wright (1951) The genetical structure of populations. *Annals of Eugenics* 15, 323-354.

Lewontin (1972) The apportionment of human diversity. *Evolutionary Biology* 6, 381-398

Nei (1973) Analysis of gene diversity in subdivided populations. *PNAS* 70, 3321-3323

などで、当時は多型性が低いマーカーなどが中心であった。

多型性や分化が大きい場合には G_{ST} のような指標は意味のある結果を導かない
マーカーの多型性が高いと、集団間で対立遺伝子を共有していない場合でも G_{ST} は 1 にならず、多型性が上がるにつれて、対立遺伝子を共有していない場合の G_{ST} の値は小さくなる。

→同じ G_{ST} の値であっても、マーカーの多型性によって意味する分化の程度は異なる！

$$H_T = H_S + D_{ST} < 1$$

なぜならば H (ヘテロ接合度の期待値) = $1 - \sum p_i^2$ なので、 $H = 1$ になることはありえない
この関係式から、 H_S と D_{ST} は非独立であり、 H_S が大きいと自ずと D_{ST} は小さくなる、すなわち集団間の多様性は小さくなる。

$G_{ST} = D_{ST} / H_T$ を考えれば、 H_S が大きくなると G_{ST} も小さくなることは容易に想像がつく。
ヘテロ接合度や Shannon の多様度のような比率尺度の比較は意味のある結果を出さない、なぜならば分集団と分集団全体の多様性の関係が線形ではないから。

分集団内と分集団間の要素が真に独立した多様性の尺度が必要。

そのキーとなる尺度は以下の論文で示されている。

Hill (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* 54, 427-432.

「対立遺伝子を共有していない分集団がプールされて一つの集団として扱われた場合、プールした集団の多様性は元々の分集団の多様性の 2 倍になる」

これを基に考えたのが Diversity Δ で Jost (2008) の式 4 で表される。

$q=0$ のとき Δ は対立遺伝子数に、 $q \rightarrow 1$ のとき Δ は Shannon のエントロピーに、 $q=2$ のとき Δ は Nei の遺伝子多様度の逆数である Kimura & Crow の有効対立遺伝子数になる。

真の多様性尺度は解釈が簡単であるべき。

多様性が 20% 低下したというときに、「20%の対立遺伝子が失われた」であれば解釈はシンプルだが、「20% ヘテロ接合度が下がった」は解釈が難しい。

真の多様性 Δ_T は分集団内の多様性 Δ_S と分集団間の多様性 Δ_{ST} の積であるべき

$$\rightarrow \Delta_T = \Delta_S * \Delta_{ST}$$

Δ_{ST} は分集団間の分化の程度を示す絶対値なので、 G_{ST} に替わる指標として、1~0 の間を取るような値を考案する必要がある。

$$D \equiv [(\Delta_S/\Delta_T) - 1] / [(1/n) - 1]$$

Jost (2008) の論文に対するコメント①

Ryman and Leimar (2009) G_{ST} is still a useful measure of genetic differentiation - a comment on Jost's D . Mol Ecol 18, 2084-2087.

Jost (2008) は、「 D はヘテロ接合度 H とは独立した分化の指数であり、 G_{ST} よりも優れている」と主張しているが、場合によっては D は G_{ST} よりも H の影響を受けるし、突然変異の影響も G_{ST} と D では様々なので一概に G_{ST} を棄却して D を使えばいいわけではない。また、 G_{ST} は対象生物の複数の遺伝子座で観察された分化の平均値として取り扱われるが、 D は個々の遺伝子座ごとに記述しないとイケない。研究者の興味が特定の遺伝子座の突然変異や多様性である場合は D は有効であるが、遺伝的浮動や移住などのデモグラフィックなプロセスである場合は D は適切ではない。そういう意味で D には弱点があり、また R_{ST} や G_{ST} にも弱点がある。

(個人的なコメント：なので、どれがいいということはないってこと？興味の対象によってどれを使うか考えろ！ってことですいいんですかね？)

Jost (2008) に対するコメント②

Heller and Siegmund (2009) Relationship between three measures of genetic differentiation G_{ST} , D_{est} and G'_{ST} : how wrong have we been. Mol Ecol 18, 2080-2083

G_{ST} が H_S に影響され、 $G_{ST} < 1-H_S$ であることはすでに Hedrick (1999) が指摘している。 G_{ST} の問題を解決するために Hedrick (2005) の G'_{ST} と Jost (2008) の D がそれぞれ異なるアプ

ローチで考案された。

G'_{ST} は G_{ST} を標準化して 0-1 に納まるようにしたもの。

D は多様性 H を分ける完全に新しい方法を採用し、直観的・数学的にわかりやすい指標を考案。

Mol Ecol に出版されている G_{ST} とその親戚たちのデータをメタ解析し、 G'_{ST} と D_{est} を算出して比べてみました。

G_{ST} , G'_{ST} , D_{est} はそれぞれ強い相関があったけど、 G_{ST} は G'_{ST} , D_{est} に対して常に過少評価。どの指数も H_S と負の相関があった。 G'_{ST} と D_{est} を比べてみると G'_{ST} の方が高い値になりがちではあるがそんなに大きな違いはない（相関係数は $\tau = 0.92$ ）。

マイクロサテライトのような高い H_S のマーカーを使うときは G_{ST} は不適なので G'_{ST} か D_{est} を使うのがいいよ。