

北里大学院理学研究科

2023 年度博士論文

両生類の異種交配における
トランスポゾン・piRNA を介した非対称サブゲノム進化

須田 皓介 (DS-21901)

指導教授 遺伝子機能発現学 松尾拓哉

目次

1. 要旨	1
2. 序論	2~4
3. 材料と方法	5~9
4. 結果および個々の考察	10~18
5. 全体の考察	19~22
6. 参考文献	23~27
7. 謝辞	28
8. 表	29~39
9. 図	40~51

1. 要旨

ゲノム重複を伴う異種交配は、遺伝的多様性の増大を介して、系統進化や種の多様化に大きな貢献を果たしてきた。興味深いことに、異質四倍体化では、両親由来のサブゲノムの非対称的な分子進化が動物界で知られている。例えば、両生類のツメガエル属では、異質四倍体の子孫のアフリカツメガエルのサブゲノム L と S において、S 染色体が L に比べて全て短くなっている。またトノサマガエル属では、ワライガエルとコガタガエルの雑種のヨーロッパトノサマガエルにおいて、その体細胞は両親由来のサブゲノムが存在するが、生殖系列細胞ではコガタガエル由来のサブゲノムが選択的に排除される。しかし、これらの選択的 DNA 欠失の分子機構は全く未解明である。私はこの問題に対し、異種ゲノム間における自己・非自己の認識機構の存在を想定し、生殖隔離された集団のゲノム内で独自に進化する「トランスポゾン-piRNA システム」の不和合性が、サブゲノム選択的 DNA 欠失を誘引するという作業仮説を考えた。piRNA はトランスポゾンを抑制する small RNA である。本研究では、この作業仮説の検証を目的として、上述の両生類の 2 つの異種交配系において、サブゲノム間の比較解析を行った。ツメガエル属の異種交配に関しては、異質四倍体の 2 種（アフリカツメガエル・キタアフリカツメガエル）の L/S サブゲノムと、対照として二倍体のネッタイツメガエルのゲノムとの比較解析の結果、1,800-1,700 万年前に起った異種交配直後から、L に比べ S サブゲノムで、より多くの DNA トランスポゾンの活性化、遺伝子・遺伝子間領域の欠失、および大規模の染色体内逆位が起ったことが推定された。更に特定のトランスポゾンの活性化と S サブゲノムに偏った DNA 欠失との相関が認められた。一方、雑種ヨーロッパトノサマガエルでは、異種交配が現行され、トランスポゾンと piRNA の発現解析が可能という利点がある。そこで親ガエル 2 種のゲノム配列を新たに決定後、親ガエルと雑種の精巣の mRNA および piRNA 発現量の比較解析を行った。その結果、雑種精巣において、生殖細胞で欠失するコガタガエルゲノムに由来するトランスポゾンおよび piRNA 発現の変動が、残存するワライガエルゲノム由来のそれぞれと比べ、より大きいことが検出された。更にコガタガエルゲノム由来の 2 種の DNA トランスポゾンに対応する piRNA の精巣での発現量が、コガタガエルと比較して雑種で低下しており、雑種精巣ではコガタガエルゲノム由来の 2 種のトランスポゾン活動の脱抑制が示唆された。

以上の異種交配系での解析から、私は以下のモデル・仮説を提案する。(1)異種交配後の非対称サブゲノム進化にはトランスポゾンが直接関与する。片親種のサブゲノムに偏った DNA 欠失は、そのサブゲノム由来 DNA トランスポゾンの活性化を介した DNA 鎖切断が関与する可能性がある。(2)ヨーロッパトノサマガエルでの解析から、雑種の生殖系列細胞では、異種サブゲノムの混合状態が、コガタガエルの「トランスポゾン-piRNA システム」の大きな変動や一部崩壊をもたらし、コガタガエル由来 DNA トランスポゾンの脱抑制が、そのサブゲノムの完全欠失の起因となった。(3)両生類における異種ゲノム間の自己・非自己システムの実体は、種・集団内で独立に進化した「トランスポゾン-piRNA システム」の不和合性である。

2. 序論

近縁 2 種の交雑による異種交配は、ゲノム多様化を伴い、系統進化や種の多様化に大きな貢献を果たしてきたと考えられる。ゲノムが異なる異種間の交雑種は、多くの場合、異質倍数体化を伴うゲノム重複が起こるので、遺伝的多様性が増大される。これにより、環境への適応力が大きくなるだけでなく、“進化能力 (Evolvability)”が増大したゲノムの獲得、言い換えれば、“形質進化許容可能な潜在的ゲノムの獲得”につながる。この“進化能力”は、同種間における雑種である雑種強勢と比べ、異種交配の方が遺伝的多様性という点で大きく、異種交配が、有性生殖を介した生命進化の大きな原動力の 1 つと考えることができる。半世紀以上も前のゲノム情報がまだなかった時代に、進化学者の大野乾は、脊椎動物進化では二回の全ゲノム重複 (2R-WGD [whole genome duplication]) が起こったことという仮説を提唱した (Ohno 1970)。この 2R-WGD 仮説は、脊椎動物の祖先の、形質進化許容可能な潜在的ゲノムの獲得を示唆している。即ち、2R-WGD が、脊椎動物進化において、環境適応に対応する器官形成を含む形態進化に深く関わってきたのではないかと予想されてきた (Holland et al., 1994; Spring 1997)。最近の脊椎動物を含む脊索動物の各種ゲノムを用いたシテニー解析により、脊椎動物の共通祖先では同質型の WGD が、その後、有顎類の共通祖先で異種交配を介した異質型の WGD が起きたことが示唆され、2R-WGD という大野仮説が支持された (Simakov et al., 2020; Nakatani et al., 2022)。2 回目の WGD の異種交配によるゲノム倍数化は、脊椎動物の有顎類の多様化に大きく貢献してきたと考えられる。

上述したように、異種交配は生命進化に多大な貢献をしてきたと予想されるが、生命進化過程では、頻繁に多様化に貢献してきたわけではない。一般的には、異種交配では、個体発生過程の異常や、個体形成はできても次世代 (子孫) ができない (生殖能力がない)。この異種交配個体の生殖能力欠如という現象は、節足動物の昆虫や脊椎動物のいくつかの進化系統群で報告されている。研究が進んでいるショウジョウバエでは、異なるゲノム由来のトランスポゾンと piwi-interacting RNA (piRNA) の不和合性が雑種の不妊の原因となる可能性が提唱されている (Kotov et al., 2019)。トランスポゾンはゲノム内を転移・増殖するが、多くが利己的な DNA 配列で、生殖系列の細胞で活性化すると、転移・増殖が抑制されなければ、遺伝子の破壊を含むゲノムの不安定化を引き起こし、次世代に影響を及ぼすことになる (Belancio et al., 2008)。この次世代へ影響を及ぼすゲノムの不安定化を抑制するために、多細胞動物では、生殖系列細胞において piRNA という small RNA によってトランスポゾンの活動を抑制するシステム ([トランスポゾン-piRNA システム]) を構築してきた (Iwasaki et al., 2015)。例えば、ある動物の有性生殖集団で、新たにトランスポゾンが生殖系列細胞で分子進化したとすると、これを抑制する piRNA は新たに分子進化してくることになる。即ち、生殖隔離された集団は、独自の [トランスポゾン-piRNA システム] がゲノム内に構築されて生き残ってきたことになる。生殖隔離から経過した時間が長くなるほど、トランスポゾンと piRNA の構成と種類は変化し、明確に識別 (→ゲノムの自己・非自己認識) されることが考えられる。このアイデアが、本研究における作業仮説の根底にある。

異種のゲノムを引き継ぐ異質倍数体化した種では、サブゲノムが非対称的に分子進化してい

ることが報告されてきた (Rodriguez and Arkhipova. 2018)。植物の異種間雑種では、一方のサブゲノム由来の遺伝子発現量が、偏って多くなるという“サブゲノム優勢”という現象が知られている (Cheng et al., 2018 ; Alger et al., 2020; Wang et al., 2022)。一方、動物の異種間雑種では、一方のサブゲノムに偏った遺伝子欠失が、無尾両生類の異質四倍体アフリカツメガエルや、2 回目の WGD による異質倍数化後の有顎類子孫系統で報告されている (Session et al., 2016; Shimacov et al., 2020; Nakatani et al., 2021)。前者の異種間雑種の L と S のサブゲノムをもつアフリカツメガエルに関しては、アフリカツメガエルおよび近縁種のゲノム比較解析から、まず L と S の二倍体ゲノムのツメガエル祖先種が、約 3400 万年前に生殖隔離により L あるいは S の二倍体ゲノムをもつ種に分化したのち、約 1800-1700 万年前に、異種として再び出会い、異質四倍体化した集団が誕生したことが示唆されている (Session et al., 2016)。興味深いことに、異種交配後から現在までの間に、S サブゲノムの染色体は、対応する L サブゲノムの染色体に比べ、全て短いことが報告された (Session et al., 2016)。一方、同じ無尾両生類の異種交配において、異質四倍体化せずに、一方のゲノムだけが欠失する例も知られている。ヨーロッパ各地で生息する雑種ヨーロッパトノサマガエルにおける“雑種生成 (hybridogenesis)”と呼ばれるもので、親種 (ワライガエルとコガタガエル) の雑種であるヨーロッパトノサマガエルは、生殖系列細胞でコガタガエル由来のゲノムが選択的に排除されるという現象である (Uzzell et al., 1980; Dedukh et al., 2019; Miura et al., 2021)。これらの異種交配における、一方のサブゲノムに特異的な部分 DNA 欠失やゲノム排除という非対称的サブゲノム進化の分子機構は、現在、未解明である。これらのツメガエルとトノサマガエルの 2 タイプの異種交配では、片親のサブゲノム DNA が部分欠失あるいは全欠失という異なる現象であるが、片親種サブゲノムに対する特異性がある点で共通性がある。その共通の分子機構としては、異種ゲノム間における自己・非自己の認識機構の存在を考えることが可能かもしれない。そこで、私は、これら 2 つの異種交配におけるサブゲノム非対称的進化の原因は、生殖隔離中にそれぞれの集団で独自に進化した [トランスポゾン-piRNA システム] 間の不和合性にあるという仮説を考えた。具体的には、異種ゲノム混合状態により、システム間の不和合性が起こり、欠失あるいは排除されるサブゲノム内のトランスポゾン制御システムが機能低下し、片方種サブゲノムのトランスポゾンの脱抑制をもたらすという仮説である。本研究は、この不和合性仮説の検証を大きな目的として、①約 1800-1700 年前に異種交配した異質四倍体ツメガエル、②現存の雑種ヨーロッパトノサマガエルに関して、ゲノム・サブゲノムの比較解析およびトランスポゾン-piRNA の解析を行った。

ツメガエル (*Xenopus*) 属の異種交配に関しては、二倍体種のネッタイツメガエル (*X. tropicalis*) と異質四倍体のアフリカツメガエル (*X. laevis*) とキタアフリカツメガエル (*X. borealis*) 2 種、計 3 種のゲノムの比較解析を行った結果、異種交配後、S サブゲノムに偏った ①遺伝子欠失、②遺伝子間領域の欠失、③DNA トランスポゾン活性化、④大規模の染色体逆位、が示唆された。更に、アフリカツメガエルでのゲノム解析から、特定の DNA トランスポゾンの活性化の S サブゲノムに偏った DNA 欠失の関与が示唆された。異種交配時期の L/S サブゲノムに関わらない大規模な DNA トランスポゾンの活性化は、以前、私らが論文 (私が筆頭著者: Suda

et al., 2022) として発表していたが、今回は、最新のデータベースで同様の結果を確認し、新たな観点から比較解析を行い、L サブゲノムとは異なる S サブゲノムに特徴的な DNA トランスポゾンの活性化を検出した。

ワライガエルとコガタガエルの異種交配による雑種のヨーロッパトノサマガエルに関しては、親ガエル 2 種のゲノム配列を決定後、親ガエルと雑種の精巣の mRNA 及び piRNA 発現の比較解析を行い、雑種の精巣においてコガタガエルゲノム由来のトランスポゾン発現が大きく変動していることがわかった。さらに興味深いことに、雑種精巣で発現しているコガタガエルゲノム由来のいくつかのトランスポゾンに対する piRNA の発現量が、コガタガエルと比較して低下していた。これらの結果から、雑種ヨーロッパトノサマガエルの精巣では、特定のトランスポゾンに対応する piRNA の発現が低下により、コガタガエルゲノム由来のトランスポゾンが脱抑制されていることが示唆された。

以上、無尾両生類における 2 タイプの異種交配系での解析から、異種交配後の非対称サブゲノム進化には、トランスポゾンが直接関わる可能性が示唆された。考察の項の最後で、片親種サブゲノムの偏った部分欠失、あるいは完全欠失（ゲノム排除）に関する分子モデルを提案する。なお、ツメガエル属の異種交配の解析は、“結果と個々の考察” の第 1 章に、雑種ヨーロッパトノサマガエルの解析は、“結果と個々の考察” の第 2 章にそれぞれ記した。

3. 材料と方法

【ツメガエル (*Xenopus*) 属について】

使用データ

アフリカツメガエル (*X. laevis*) およびネッタイツメガエル (*X. tropicalis*) のゲノム情報は Xenbase (<http://www.xenbase.org/entry/>) から取得し、アフリカツメガエルは v10.01, ネッタイツメガエルは v10.0 を使用した。キタアフリカツメガエル (*X. borealis*) ゲノムアセンブリは Genbank に登録されているもの (以下はサイト) を使用した:

(https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/024/363/595/GCA_024363595.1_UCB_Xborealis1/) 。

キタアフリカツメガエルの遺伝子アノテーションは後述のトランスポゾンアノテーションでマスキングしたゲノムアセンブリ、オタマジャクシ幼生の中腎 (SRR11844034), 成体の肝臓の 2 種の RNA-seq から、braker2 (<https://github.com/Gaius-Augustus/BRAKER>) を用いて作成した。また、ツメガエルのトランスポゾンライブラリは Repbase

(<https://www.girinst.org/replib/>) に登録されているアフリカツメガエルとネッタイツメガエルのデータを組み合わせて使用した。

トランスポゾン同定

ゲノム内のトランスポゾンの同定は RepeatMasker (<http://www.repeatmasker.org/>) を使用した。同定したトランスポゾン配列は Repbase から取得したコンセンサス配列に MAFFT (<https://mafft.cbrc.jp/alignment/software/>) の --retree2 --reorder オプションを使用してアライメントした。更に作成した Python スクリプト(補足資料 1)を用いて CG 配列を全アライメントから削除し、FastTree (<http://www.microbesonline.org/fasttree/>) -nt でゲノムトランスポゾンのコンセンサス配列からの塩基置換率を再計算した。その後、別に作成した Python スクリプト(補足資料 2)を用いて、塩基置換率の低さを基準にして、位置が重複しているゲノムトランスポゾンを再配置した。トランスポゾンランドスケープは上記の再計算したゲノムトランスポゾンアノテーションから RepeatLandscape (github.com/caballero/RepeatLandscape)を用いて作成した(補足資料 3)。

各種ゲノムにおけるオーソログの同定

ネッタイツメガエル、アフリカツメガエル、キタアフリカツメガエルのゲノム間のオーソログあるいは L/S 間のホメオログ (同祖遺伝子) を同定するために、ネッタイツメガエルのゲノムデータベース (上記参照) から多くがアノテーションされ、相同性解析やシンテニー解析が容易なタンパク質コード遺伝子をピックアップした。その際、タンパク質コード遺伝子と予想されていても、タンデム型重複型遺伝子や、機能が未知の遺伝子は除外した。ピックアップしたネッタイツメガエル遺伝子の各エクソンをアフリカツメガエルあるいはキタアフリカツメガエルの L/S ゲノムにおいて blastn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn>) を用いた相同性解析

を行い、ネットアイツメガエル遺伝子のオーソログ候補を、ビットスコア (≥ 300) で候補遺伝子として選択した (補足資料 4)。候補遺伝子は、その周囲 1 Mbp ゲノム配列の範囲内に、3 つ以上の遺伝子が保存されている場合、シンテニーがあると判断し、オーソログ遺伝子とした (補足資料 5)。

染色体逆位

ネットアイツメガエルとアフリカツメガエルの 2 種間の祖先の遺伝子の並びが、正確には推定できない。そのため、本研究では、ネットアイツメガエルでの染色体上の遺伝子の並び順を祖先系として、これに対して、アフリカツメガエルの L/S サブゲノムの各染色体で、並び順が逆なものを逆位とした。逆位を同定するために、上記したそれぞれのゲノムのオーソログ遺伝子を利用し、上流・下流 20Mbp の合計 40Mbp の範囲で一単位とし、逆位を探索した (補足資料 6)。これらの逆位を異なるサブゲノムの逆位と比較し、重複していないものをサブゲノム固有の逆位として扱った。

各染色体上の DNA 欠失量とトランスポゾン量との相関

ネットアイツメガエル各染色体に対して、アフリカツメガエル L あるいは S サブゲノムの対応する染色体における DNA の欠失量を、シンテニー間距離を計測するスクリプトを Python によって作成し、計測した (補足資料 7)。得られた各シンテニー間距離からトランスポゾンを除いた距離を計測し、この値を用いて変動率 (アフリカツメガエル L or S / ネットアイツメガエル) を算出し、各染色体の変動率の中央値を染色体欠失率とした。一方、異種交配後のトランスポゾンの増加量を評価するために、塩基置換率が 0.035 ~ 0.05 の各トランスポゾンサブファミリー断片に着目した。染色体ごとに、これらのトランスポゾン断片を各サブファミリーのコンセンサス配列にアライメントして、カバレッジを求めた。次に、カバレッジの最高値を、トランスポゾンを除いた各染色体の長さで標準化し、この標準化カバレッジの値を各トランスポゾンサブファミリーの増加量とした (補足資料 8)。得られた各染色体の欠失率とトランスポゾン増加量から、トランスポゾンサブファミリーごとの相関係数と近似直線の傾きを算出した。この解析では、塩基置換率が 0.035 ~ 0.05 の範囲において、合計が 30,000bp 以上のトランスポゾンのみを使用した。

【トノサマガエル (*Pelophylax*) 属】

カエルの供与

ワライガエル (*P. ridibundus*), コガタガエル (*P. lessonae*), これら 2 種を親種とする雑種ヨーロッパトノサマガエル (*P. esculentus*) は、ロシア西部に生息する個体を Dr. Vladimir Vershinin (Ural Federal University, Russia) より供与された。

ゲノムの単離と全ゲノムシーケンス

ワライガエルとコガタガエルの高分子量ゲノムは成体の血液より、それぞれ MagAttract HMW DNA kit (Qiagen, Venio, Netherlands) と NucleoBond HMW DNA kit (Macherey-Nagel, Duren, Germany) を使用して精製した。ゲノムシーケンスは Takara (Kusatsu, Japan) に委託した。

ゲノムの再アセンブリ

得られたワライガエル、コガタガエルのゲノムシーケンスはヨーロッパアカガエル(*Rana temporaria*)のゲノム配列(https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_905171775.1/)に、Ragtag (<https://github.com/malonge/RagTag>) を用いてアライメントし、再アセンブリした。

遺伝子アノテーション

再アセンブリしたワライガエル・コガタガエルのゲノム配列、精巣・卵巣の mRNA-seq 情報から braker (<https://github.com/Gaius-Augustus/BRAKER#braker-with-mrna-seq-data>) を用いて遺伝子を予測した。遺伝子群のタンパク配列は、ヨーロッパアカガエル、ネッタイツメガエルのタンパク配列に MMseqs2 (<https://github.com/soedinglab/MMseqs2>) を使用してアライメントした。アライメント結果から、両方のタンパク質にアライメントされ、identity ≥ 0.4 , alignment length ≥ 0.5 を満たすものを遺伝子とした。

ゲノム上におけるトランスポゾンの同定

再アセンブリしたワライガエル・コガタガエルのゲノムから RepeatScout (<https://github.com/mmcco/RepeatScout>)および RepeatModeler (<https://www.repeatmasker.org/RepeatModeler/>) を用いて、それぞれのゲノムのトランスポゾンライブラリを作成した。RepeatScout は build_lmer_table -l 14 で実行後、filter-stage-1.prl でフィルタリングをし、filter-stage-2.prl でゲノム内に 20 個以上存在するトランスポゾンを抽出後、1,000bp 以上の長さのトランスポゾンを選択した。RepeatModeler は -LTRStruct で実行した。2 種のトランスポゾンライブラリは、fasta-trf-filter.pl を用いて、タンデムリピートのフィルタリング後、megablast -max_target_seqs 25 -culling_limit 2 -evaluate 1e-10 により遺伝子アノテーションから作成したタンパク質コード配列にアライメントし、assemblage (<https://github.com/sujaikumar/assemblage>)の fastaqual_select.pl で、タンパク質コード配列と相同配列を除外した。これらのフィルタリングした RepeatScout, RepeatModeler 製のトランスポゾンライブラリを、CD-HIT (Weizhong et al., 2006) -c 0.9 -d 0 -T 12 -aS 0.8 を用いて部分一致する配列を探索し、作成した python スクリプトでトランスポゾンのコンセンサス配列を作成した(補足資料 8, 9, 10)。コンセンサス配列は、Repbase に登録されている既知のトランスポゾンと相同な配列を除いた後、DeepTE (<https://github.com/LiLabAtVT/DeepTE>) によって分類した。トランスポゾンランドスケープはツメガエルと同様の方法で作成した。

Total RNA の調整 および mRNA-seq と small RNA-seq

各種の無尾両生類の成体精巣は液体窒素で瞬間凍結、成体卵巣は分割後瞬間凍結したのち、凍結状態で粉碎を行い、直ちに ISOGEN II (ニッポンジーン, 東京) を用いて Total RNA を単離した。Total RNA は、ダナフォーム (横浜) に委託して、mRNA と small RNA を分離され、網羅的な RNA 配列 (mRNA-seq および small RNA-seq) の決定が行われた

piRNA ライブラリ

small RNA-seq によって得られた塩基配列情報は、アダプターを cutadapt (Martin Marcel. 2011) でトリミングした。その後、miRNA、rRNA、tRNA、snoRNA のデータベース/ファイル (それぞれ mirDB・mirgeneDB/mature.fa・ALL.fas、Silva / SILVA_138.1_SSURef_tax_silva.fasta、Genomic tRNA/eukaryotic-tRNAs.fa、snoDB/snoDB_FASTA.fa) を使 い、bowtie (<https://github.com/BenLangmead/bowtie>)-v 3 -seed 0 で相当配列を検出し、piRNA 以外の配列を除いた。残った配列の中で 24~32bp の長さの配列を piRNA とした。

塩基置換速度の推定

両生類の無尾目 7 種 (ワライガエル、コガタガエル、アメリカウシガエル [*Aquarana catesbeianus*]、ヨーロッパアカガエル [*Rana temporaria*]、ヒマラヤガエル [*Nanorana parkeri*]、ネッタイツメガエル [*Xenopus tropicalis*] および アフリカツメガエル [*Xenopus laevis*] L/S サブゲノムと両生類の無尾目 1 種 (アシナシイモリ [*Rhinatrema bivittatum*] の計 8 種のタンパク質コード遺伝子を、相互に diamond (Buchfinc et al., 2015) --eval 1e-5 --max-target-seqs 10000 --query-cover 50 --subject-cover 50 でアライメントした。その結果を、作成した python スクリプトで、identity >= 60, bits score >= 300 を満たす最もビットスコアの高いペアをオーソログとした(補足資料 11)。両生類 8 種で共通のオーソログ 597 個をピックアップし、それぞれのオーソログのタンパク質配列を MACSE (Ranwez et al., 2011) でアライメントし、Gblocks (Castresana J. 2000) を用いて 7 種間で保存されているコドンのみ抽出した。その後、作成したアライメントから、作成した python スクリプト (補足資料 12) で終始コドンを削除し、PAML (Yang Z. 1997, 2007) の yn00 で塩基置換率を算出した。

トランスポゾンの mRNA の定量

ワライガエル・コガタガエルの再アセンブリしたゲノム DNA とトランスポゾンアノテーションから、BEDTools (Quinlan AR and Hall IM. 2010) の getfasta オプションで、それぞれの種のゲノム内のトランスポゾンライブラリの fasta ファイルを作成した。次に、そのトランスポゾンライブラリに対して、精巣の mRNA-seq ライブラリを hisat2 (Sirén J et al., 2014) でアライメントし、SAMtools (Petr et al., 2021) によって、ペアエンドマッピングされたアライメントを抽出した。次に、作成した python スクリプトでゲノム内の各トランスポゾン断片にアライメントされたリードを集計して、RPKM を計算した (補足資料 13)。トランスポゾン発現量は、サブ

ファミリーごとの合計値のサンプル間の平均とした。雑種ヨーロップパトノサマガエルに関しては、その精巢 mRNA-seq ライブラリをワライガエル・コガタガエル両方のゲノムトランスポゾンライブラリにアライメントし、作成した python スクリプトにより、各ペアエンドリードで最もミスマッチの少ないアライメントを選択し(補足資料 14)、RPKM を計算した。

piRNA 発現量の定量

strobealign (Sahlin K. 2022)を用いて、piRNA ライブラリを対応するゲノムにアライメントし、作成したスクリプト(補足資料 15)で、ミスマッチが最も少なかったアライメントを抽出した。piRNA 遺伝子座は個体間で変わらないと想定し、一部でも場所が重なるアライメントを結合して piRNA 前駆体を疑似的に再構成し、個体間で場所が重なる(全体長の 1/5) piRNA 前駆体に属する piRNA のみ後の解析に用いた。strobealign -U -B 0 -v -N 1000 -M 1000 -S 0 を用いて、得られた piRNA ライブラリを、ゲノムトランスポゾンライブラリにアライメントし、作成したスクリプトで各トランスポゾンサブファミリーの CPM を計算した。

4. 結果 および 個々の考察

【第1章】 異質四倍体ツメガエルにおける非対称サブゲノム進化

1-1. 異質四倍体化後のゲノム進化では 約 15,000 個のオーソログセットを保持する遺伝子欠失が進行してきた

ツメガエル属のゲノムデータベースを用いて、2 倍体ネッタイツメガエルと 2 種の異質四倍体アフリカツメガエル及びキタアフリカツメガエルのゲノムを公表された最新のデータベース (ネッタイツメガエル ver10.0 およびアフリカツメガエル 10.01, <http://www.xenbase.org/entry/;キタアフリカツメガエル,https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/024/363/595/GCA02436359>

5.1_UCB_Xborealis1/) を用いて、各染色体の長さや遺伝子 (エクソンとイントロン)・トランスポゾンの解析を行った (図 1; 表 1-4)。その結果、初期のデータベース (ネッタイツメガエル ver9.0・アフリカツメガエル 9.2: <http://www.xenbase.org/entry/>)を用いた解析 (Session et al., 2016) より、精度のより高い解析結果が得られた。染色体の長さに関しては、以前のゲノム解析での報告 (Session et al., 2016) と同じく、L サブゲノムの各染色体の長さは、ネッタイツメガエルの対応する進化的起源が同じ染色体と比較してほぼ変わらないのに対して、S サブゲノムでは 10%以上短くなっていた。重要点の 1 つは、二倍体のネッタイツメガエルでは、ゲノムにおける遺伝子 (エクソン+イントロン) の占める割合は 34.6%であったのに対し、アフリカツメガエルの L あるいは S サブゲノムにおける遺伝子の占める割合は、それぞれ 30.2% あるいは 26.4%で、ネッタイツメガエルのそれよりも低いことがわかった (図 1; 表 2; 表 3)。また、DNA トランスポゾンあるいはレトロトランスポゾンに対応する DNA のゲノムあるいはサブゲノムに対する割合を調べたところ、二倍体のネッタイツメガエルでそれぞれ 24.9%あるいは 10.5%であった (図 1; 表 2)。一方、アフリカツメガエルの L サブゲノムでは、それぞれ 29.0%あるいは 11.0%、S サブゲノムではそれぞれ 29.8%あるいは 12.5%で、ネッタイツメガエルのそれ들에比べ、いずれもより高い値を示した (図 1; 表 3)。一方、キタアフリカツメガエルはゲノムデータベースが不十分であるため、L/S サブゲノム共におよそ半分が未分類な DNA で占められて (図 1; 表 4)、ネッタイツメガエルやアフリカツメガエルと比較することができなかった。

次に、アフリカツメガエルとキタアフリカツメガエルの L あるいは S サブゲノムの遺伝子欠失を詳細に調べるために、各サブゲノムで二倍体ネッタイツメガエルのオーソログ関係にある遺伝子を相同性およびシンテニー解析によって同定を行った (表 5)。ネッタイツメガエルのゲノムには、タンDEM重複遺伝子や LOC 登録された遺伝子を除き、15,022 個のタンパク質コード遺伝子が同定された。一方、アフリカツメガエルの L あるいは S サブゲノムには、ネッタイツメガエルの遺伝子とオーソログ関係にある遺伝子が、それぞれ 13,883 個あるいは 11,712 個、キタアフリカツメガエルの L あるいは S サブゲノムには、それぞれ 11,046 個あるいは 9,264 個が存在していた (表 5)。このことから、これらの異質四

倍体の両者は共に、L あるいは S サブゲノム両方で遺伝子欠失が起こり、さらに、遺伝子の欠失数が L サブゲノムに比べ S サブゲノムに多いことがわかった。

ネットアイツメガエルのタンパク質コード遺伝子の 15,022 個のうち、アフリカツメガエルの L サブゲノムと S サブゲノムの両方でオーソログ関係にある遺伝子として存在していたのは、9,947 個 (66.2%) であった。言い換えれば、ネットアイツメガエルに対応する 66.2% の遺伝子はホメオログとして、アフリカツメガエルゲノムに 2 個存在するということになる。次に、アフリカツメガエルのゲノムには、ネットアイツメガエルのオーソログセットをどの程度保持しているか、L サブゲノムおよび S サブゲノムに単独に存在する遺伝子数に、両サブゲノムに存在するホメオログペアを 1 種類の遺伝子として換算し、総和した。その結果、アフリカツメガエルのゲノムには、ネットアイツメガエルの遺伝子の 99.7% にあたる 14,982 種類のオーソログセットが存在することが推察された (表 5)。この結果から、ツメガエル属のカエル種では、タンパク質コード遺伝子のオーソログセットとして、約 15,000 の遺伝子が集団維持に必要である可能性が示唆される。また、異種交配前の L あるいは S ゲノムをもつ 2 種は、オーソログセットとして約 15,000 個の遺伝子を保持しており、異種交配後では、約 15,000 個のオーソログセットを保持可能な状態での遺伝子欠失であったことが予想される。

1-2. 異質四倍体ツメガエルの遺伝子欠失は異種交配の直後から両 L/S サブゲノム共に継続的に、かつ、S サブゲノムでより多く起ってきた

異種交配後の L/S サブゲノムにおける遺伝子欠失は、いつ、どのように起きたのかの概要を把握するために、二倍体ネットアイツメガエルのタンパク質コード遺伝子 15,022 個に対し、2 種の異質四倍体アフリカツメガエルとキタアフリカツメガエルの L/S ゲノムで比較解析を行った。2 種の異質四倍体ゲノムで共通に欠失したオーソログ関係にある遺伝子を調べたところ、L サブゲノムでは 581 個、S サブゲノムでは 2,546 個が検出され (図 2; 表 5)、約 4.4 倍、S サブゲノムの方が多いことがわかった。しかし、S サブゲノムに非常に大きく偏っているわけではなく、L/S サブゲノム両方で 500 個以上の遺伝子が欠失したと考えることもできる。L ゲノムと S ゲノムをもつツメガエルの異種交配は約 1,800~1,700 万年前、その後、アフリカツメガエルの祖先とキタアフリカツメガエルの祖先が種分化したのは 1,700 万年前に起ったと推定されている (Session et al., 2016) こと、[1-1] から異種交配前には約 15,000 個の遺伝子を保持していたと考えられるため、2 種で共通に検出された L サブゲノムでの 581 個、S サブゲノムでの 2,546 個の遺伝子の欠失の多くは、異種交配から 2 種が種分化するまでの短い期間で起ったと推定される (図 2)。

異質四倍体の 2 種で共通でない、即ち 2 種でオーソログ関係にはない遺伝子の欠失数は、アフリカツメガエル L あるいは S サブゲノムでそれぞれ 558 個あるいは 1,430 個、キタアフリカツメガエル L あるいは S サブゲノムでそれぞれ 2,729 個あるいは 3,212 個であり、種分化後の 1700 万年間で、分岐後に独立に L/S サブゲノムの遺伝子欠失が続いてきた

ことが推定された (図 2; 表 5)。

異質四倍体 2 種において、分岐前と分岐後に欠失したと予想される遺伝子をオントロジー解析したところ、全体だけでなく、時期、種、サブゲノム別でも欠失遺伝子のオントロジー的特徴は見出せなかった (データ示さず)。このことは、L/S サブゲノムにおける遺伝子欠失は、ツメガエル属が必須なオーソログセットの遺伝子のどちらかが残存するという制約以外は、ランダムであった可能性を強く示唆する。

1-3. 異種交配によって S サブゲノムに偏った逆位が起った可能性がある

染色体内の再編成に関して S サブゲノムが L サブゲノムよりも高頻度であることは報告されており、アフリカツメガエルの 2S、3S、4S、5S、8S 染色体では大規模な逆位が起きている (Session et al. 2016)。今回、私は、より新しいデータベースのアフリカツメガエルゲノムのだけでなく、キタアフリカツメガエルのゲノムの逆位の位置を調べた。ネットアイツメガエルの染色体との比較解析を行った結果、2S、3S、4S、8S 染色体の逆位はアフリカツメガエルとキタアフリカツメガエルに共通していることが明らかになった (図 3A, B)。すなわち、両種の分岐前に起こったと考えられる。L サブゲノム間、S サブゲノム間、あるいは L-S サブゲノム間で共通の逆位が多数観察された (図 3A) が、大規模な染色体逆位は、S サブゲノム間と L-S 間の共通の逆位で、L サブゲノム間の共通の逆位は少なかった (図 3B)。L-S 間の共通の逆位の多くは、ネットアイツメガエルの祖先から分岐した約 4,800 万年前から L ゲノムと S ゲノムのもつ祖先種が分岐する約 3,400 万年前に起こったと推定される。これらのことから、S サブゲノム共通の逆位の多くは、異種交配後のアフリカツメガエルとキタアフリカツメガエル共通祖先で起きたことが予想され、S サブゲノムに偏った染色体逆位は異種交配が要因である可能性が示唆された。

1-3. アフリカツメガエルの遺伝子間領域の DNA 欠失量は L よりも S サブゲノムが多い

S サブゲノムにある程度偏った遺伝子欠失 (表 5; 図 2) や染色体の逆位 (図 3) は、DNA の二重鎖切断が契機である可能性があるが、この二重鎖切断が S サブゲノムに偏ってある程度ランダムで起きるのであれば、遺伝子間領域も同様な結果になると考えられる。そこで、S あるいは L サブゲノムにおける DNA 欠失と遺伝子間領域との関連性を調べるために、ネットアイツメガエルゲノムにおける 2 つの遺伝子間距離を指標として、アフリカツメガエル L あるいは S サブゲノムにおける対応する 2 つのオーソログ遺伝子間の距離を比較した。なお、キタアフリカツメガエルはコンティグが不十分であるため、この解析には使用しなかった。トランスポゾンを除いた遺伝子間距離の中央値をネットアイツメガエルとアフリカツメガエルの L あるいは S サブゲノムの各染色体で計算し、対応するネットアイツメガエルの染色体に対して減少率を求めた (表 6)。その結果、S サブゲノム由来の染色体では、全ての染色体で減少しており、一番少ない 1 番染色体で 5.6%、一番多い 8 番染色体では 23% の減少率が検出された。一方、L サブゲノム由来の染色体では、逆に 2 番と 6 番の染色

体は少なからず増加 (1.7%) しており、また、他の染色体の減少率は 0.9-4.6%の間と S サブゲノムの対応する染色体に比べてどれもかなり低かった (表 6)。これらの結果および遺伝子欠失がランダムであることを想定する (1-2) と、DNA 欠失の始まりの二重鎖切断のターゲットは、S サブゲノムで比較的多かったある特定の DNA 配列の可能性を想定することができる。

1-4. 異種交配から現在まで続く S サブゲノムに偏った DNA トランスポゾン活性化

我々は以前、異質四倍体のアフリカツメガエルとキタアフリカツメガエルのゲノム情報 (Xenbase, *X. laevis* v9.2 genome assembly; SRR6357673 and SRR6357672) を用い、塩基置換率を基盤としたトランスポゾンのランドスケープから、約 1,800~1,700 万年前のツメガエル異種交配を介した DNA トランスポゾンの活性化を報告した (Suda et al., 2022)。本研究では、これら異質四倍体 2 種と二倍体ネッタイツメガエルの新規ゲノムアセンブリ (“材料と方法” 参照) を用いて、トランスポゾンのランドスケープを再構築した (図 4)。図 4 の横軸は、コンセンサスのトランスポゾンからの突然変異率を、縦軸は、(サブ) ゲノム長に対するある突然変異率の DNA あるいはレトロトランスポゾンの総和長の % である。その結果、異質四倍体 2 種の L および S サブゲノム両方で、塩基置換率が 0.05 付近 (1,700 万年前付近に相当) で DNA トランスポゾン活性化のピークが見られたが (図 4B, 4C)、二倍体のネッタイツメガエルでは類似のピークは見られなかった (図 4A)。一方、レトロトランスポゾンは、いずれも塩基置換率が 0.05 以下の現代に近い時期に活性化のピークが観察されたが、種によってピークの時期が異なっていた (図 4A - 4C)。これらツメガエル属のトランスポゾンのランドスケープにおける山の形や活性ピーク期は以前の報告 (Suda et al., 2022) とほぼ類似していたが、新規ゲノムアセンブリにより情報量が増加され、各トランスポゾンの活性化時期と割合はより正確になったと考えられる。

次に、S サブゲノム選択的 DNA 欠失とトランスポゾン活性化の関連性を評価するために、異質四倍体 2 種 (アフリカツメガエルとキタアフリカツメガエル) の L/S サブゲノムのトランスポゾン活性化量を異種交配前後で比較した。図 5 の横軸は、図 4 の横軸と同様に、コンセンサスのトランスポゾンからの突然変異率を、縦軸は、ある突然変異率の S サブゲノムのトランスポゾン総和長から L サブゲノムのトランスポゾン総和長を差し引いた値を示したものである。興味深いことに、両方のゲノムで DNA トランスポゾン (青線) のみが異種交配の前後で急激に傾向が変わっていた。両 2 種ゲノム共に、異種交配前では S サブゲノムの DNA トランスポゾン長が短い (図 5 ではマイナス値) が、異種交配後では S サブゲノムが L サブゲノムを上回る (図 5 ではプラス値)。一方、レトロトランスポゾン長は、両ゲノム共に、異種交配とは関係なく、S サブゲノム長が僅かに L サブゲノム長を上回っていた。更に、異種交配後、DNA トランスポゾン長の L/S サブゲノム間の差は、キタアフリカツメガエルの祖先と比較して、アフリカツメガエルの祖先は顕著であった (図 4 左側青線)。また、Session et al. (2016) が報告した異種交配直前に活性化していた S サブゲノムに

特有の DNA トランスポゾン *Tc1-14_Xt* サブファミリーのランドスケープを作製したところ、このサブファミリーの活性化は、異種交配 (0.05) 以降も継続していることがわかった (図 6)。興味深いことに、異種交配後では同じ細胞に L/S サブゲノムが存在するにも関わらず、L サブゲノムには殆ど存在しないことから、この DNA トランスポゾンは同じ染色体に転移する可能性が極めて高いと考えられる。この結果は、植物や昆虫で見られる DNA トランスポゾンが近位に優先的に転移する特性 (Local Hopping) が、両生類の DNA トランスポゾンも同様であることを示唆している (Tower et al., 1993; Dooner et al., 1994)。更に、このサブファミリーが、異種交配直後の S サブゲノムの偏った DNA 欠失に関与していた可能性が強いと考えられる。

以上、図 4-6 の結果から、S サブゲノム特異的な *Tc1-14_Xt* サブファミリーを含め、L/S サブゲノムの DNA トランスポゾンは共に、異種交配を介して活性化されるが、異種交配後は、徐々に S サブゲノム由来の DNA トランスポゾンの活動がより活発化し、アフリカツメガエルとキタアフリカツメガエル共通祖先が分岐したあとも、その傾向が続くことが考えられた。また、分岐後、DNA トランスポゾンはそれぞれの系統で独立に活性化・化石化 (盛衰) してきたと考えられる。

1-5. アフリカツメガエルの S サブゲノム選択的 DNA 欠失と DNA トランスポゾン活性化は共相関している

L サブゲノムに比べ S サブゲノムの DNA トランスポゾンの活性化 (図 5, 6) が見られたので、次に、私はトランスポゾンの種類と DNA 欠失の連関性があるのではないかと考え、それぞれの染色体の退縮量とそれぞれのトランスポゾンのサブファミリーの活性化の相関性を調べた。まず、欠失 DNA 量に対する各サブファミリーの DNA の増加量を、L あるいは S サブゲノム由来の 9 個の染色体でそれぞれ数値化し、9 点の散布図を作製した。次に、その散布図から、近似直線を作製し、相関係数を求めた後、横軸を傾き、縦軸に相関係数とした散布図を調べた全てのトランスポゾンのサブファミリーに対する散布図を作製した (図 7)。キタアフリカツメガエルはコンティグが不十分であるため、アフリカツメガエルのゲノム情報で調べたところ、DNA トランスポゾンの分布 (図 7 の青点) は L/S サブゲノム間で明確に異なった傾向を示した。即ち、L サブゲノムでは 7 割程のサブファミリーが負の傾きを示していたが、S サブゲノムでは逆に 7 割程のサブファミリーが正の傾きを示していた。また、0.75 以上の相関係数を示すサブファミリーは、L サブゲノムでは 8 個、S サブゲノムは 15 個あり、前者の 8 個のサブファミリーは全て負の傾きを示し、そのうち 6 個が DNA トランスポゾンであったのに対し、後者の 15 個のサブファミリーは全て正の傾きを示し、そのうち 13 個が DNA トランスポゾンであった (表 7, 8)。これらの結果から、DNA トランスポゾンの活性化が異種交配後の S サブゲノムに偏った DNA 欠失に関与していることが推察された。

【第2章】“雑種ヨーロッパトノサマガエル”における片親種ゲノム排除

2-1. ワライガエルとコガタガエルのゲノム配列の決定と塩基置換率の決定

雑種ヨーロッパトノサマガエルの親2種であるワライガエルとコガタガエルの雄の血液から高分子量ゲノムDNAを精製し、ゲノムDNA配列の決定とアセンブリを外部注文（宝バイオ）した。次に、得られたゲノム情報のコンティグを、ヨーロッパアカガエルのゲノムアセンブリ (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_905171775.1/) にアノテーションし、再アセンブリを行ったのち、これを用いて、遺伝子のアノテーションを行った。次に、ワライガエル・コガタガエルの精巣及び卵巣の mRNA-seq 情報から、ネッタイツメガエルに対応するオーソログを抽出したところ、ワライガエルで 9,763 個、コガタガエルで 11,219 個の遺伝子が同定された。これらのトノサマガエル属の親2種とゲノム解析されている無尾両生類6種のゲノムにおける遺伝子アノテーションより、597 個のコーディング配列（CDS [coding sequence]）を用いて同義置換率（dS）の値を計算した（表 9）。その結果、ワライガエル・コガタガエル間の dS は 0.0519 であり、この6種間では他のどの組合せよりも、低い値であった。Petros and Nikos (2010) は、ワライガエルとコガタガエルの種分化は約 800 万年前と推定している。これを用いると、ワライガエルとコガタガエルの塩基置換率/年は 1.17×10^{-9} と推定された。

2-2. ワライガエルとコガタガエルはトランスポゾンがそれぞれ独自に進化したゲノムをもつ

2-1 でリアセンブリしたゲノム情報からトランスポゾン配列の同定を行い、トランスポゾンライブラリを作成した。その結果、DNA トランスポゾンが 594 個、レトロトランスポゾンが 846 個、Helitron が 3 個、分類不明のトランスポゾンが 76 個の計 1,519 個がトノサマガエル属固有のトランスポゾンサブファミリーとして同定された。これらの遺伝子及びトランスポゾンのワライガエル・コガタガエルのゲノムに対する割合を調べた（表 10）。その結果、ツエガエル属とは異なり、DNA トランスポゾンとレトロトランスポゾンは同程度の割合だった。このトランスポゾンライブラリより、第1章で記したツメガエルゲノム・サブゲノムで作製したトランスポゾンのランドスケイプの作製と同様の方法を用いて、[2-1]で記述したワライガエルとコガタガエルのゲノムアセンブリから、これら2種のゲノム情報におけるトランスポゾンのランドスケイプを作成した（図 8）。2種が種分化した約 800 万年前（図 6 の横軸の 0.0262 に相当/塩基置換率は $0.0262 = 0.0519/2$ ）以前は、予想されるように、DNA トランスポゾン（図 8 青線）、レトロトランスポゾン（図 8 赤線）共に、ランドスケイプの波形が両種のゲノムで類似していた。一方、種分化期の約 800 万年前から現在に至るまで、DNA トランスポゾンもレトロトランスポゾンもランドスケイプの波形が少し異なっていた。そこで、種分岐後に2種のゲノム間で異なる活動をしてきたトランスポゾンを探索した。ゲノム内の存在量が 1.5 倍以上の相違が示されたトランスポゾンサブファミリーは、それぞれ数は同じで 52 個であったが、お互いを比較した場合、ワライガエル

では DNA トランスポゾンが、コガタガエルではレトロトランスポゾンがより活発化し増殖してきた傾向が見られた(表 11)。このことから、共通祖先から生殖隔離後、両種ゲノムでは生殖系列細胞内でトランスポゾンが、それぞれ独自に変異を伴いながら、活動(増殖・転移)および化石化してきたと予想される。言い換えれば、トランスポゾンの活動を介して、親ガエル 2 種は独立にゲノム進化してきたと言える。

2-3. 雑種の精巣ではコガタガエル由来のトランスポゾンの発現がワライガエルに比べ大きく変動している

ショウジョウバエの雑種の生殖細胞では、piRNA の不和合性により片親ゲノム由来のトランスポゾンが脱抑制する (Khurana et al., 2012)。そこで、コガタガエルゲノムが選択的に排除される雑種ヨーロッパトノサマガエルの生殖系列細胞で、コガタガエルゲノム由来のトランスポゾンの発現に、脱抑制が起きている可能性を検討した。まず、雑種(ヨーロッパトノサマガエル)および親 2 種(ワライガエル、コガタガエル)の各精巣の全 RNA を抽出した。[材料と方法] で記述したように、全 RNA を、外部委託(ダナフォーム; 横浜)により、mRNA と small RNA に分別し、それぞれ RNA-seq を行った (small RNA-seq に関しては、下記 2-4 で記述)。得られた mRNA-seq 情報から、トランスポゾン由来の配列を抽出したのち、発現量を調べた。[2-2]で上述したように、サブファミリーは 1,519 個あったが、そのうち 10 個は発現が確認されなかったため、1,509 個のトランスポゾンサブファミリーについて解析し、2 群間比較として MA plot を作成した。横軸は、それぞれのトランスポゾンサブファミリーについて、片親種と雑種での平均発現量を示し、縦軸は片親種の発現に対する雑種の発現の比を対数で表した (図 9)。発現量は正規化するために、[reads per kilobase of exon per million mapped reads (RPKM)+1]を対数変換 (+1 は対数変換のため)させた。この 1,509 個のサブファミリーの MA plot の散布図から、雑種の精巣でのトランスポゾンサブファミリーの発現は、ワライガエル (図 9A)もコガタガエル (図 9B)の精巣での発現に比べ、同程度のものや、発現誘導(〜30 倍)あるいは発現抑制(〜1/30)されるものなど、サブファミリーの種類によって様々であった。しかし、図 9A と図 9B を俯瞰すると、ワライガエルガエル由来のトランスポゾンが、コガタガエルのそれに比べ、雑種で発現抑制されるものが、より多い傾向が見られた。そこで、発現量比を、1/2 倍以下、1/2-1 倍、1-2 倍、2 倍以上でサブファミリーの数を調べたところ、雑種で発現量が下がったのは、ワライガエルで 81.3%、コガタガエルで 62.9%であり、ワライガエルガエル由来のトランスポゾンがより抑制される傾向が示された (表 12)。一方、雑種で発現誘導されたサブファミリーは、ワライガエルで 17.8%、コガタガエルで 35.3%であり、2 倍以上の発現誘導のサブファミリーは、ワライガエルでは 1.5%であったのに対し、コガタガエルでは 9.3%であった (表 12)。さらに、4 倍以上のものは、ワライガエルゲノムで 0.53% (8 個)、コガタガエルゲノム由来が 2.65% (40 個)であった。これらの結果から、雑種では、コガタガエルゲノム由来のトランスポゾンがより活発化していることがわかった。

次に、この発現量の変動のばらつきにゲノム間で有意な差があるかを、雑種あるいは親種の精巣で発現が見られる両ゲノムに共通の 1,470 個のトランスポゾンサブファミリーに対して、発現比のばらつきをボックスプロットし、F テストを行った (図 9C)。その結果、コガタガエルゲノム由来のトランスポゾンは、ワライガエルのそれと比較して、有為に発現変動のばらつきがより大きいことが示された (図 9C 左)。DNA トランスポゾンとレトロトランスポゾンをピックアップした解析でも、同様であった (図 9C 中央・右)。これらの知見から、雑種の精巣では、コガタガエルゲノムに若干偏ったトランスポゾン制御の混乱が起き、トランスポゾンの抑制 (不活性化) あるいは脱抑制 (活性化) が、コガタガエルゲノム由来のトランスポゾンにより多く起きた可能性も示唆された。

2-4. 雑種の精巣ではワライガエルに比べコガタガエルゲノム由来のトランスポゾンに対応する piRNA 発現量がより低下する傾向がある

雑種の精巣において、トランスポゾンの発現変動が、ワライガエルに比べ、コガタガエルゲノム由来のトランスポゾンの活性化が示された (図 9; 表 12) ので、次にトランスポゾンの発現を負に制御する piRNA 発現の比較解析を行うため、2-3 で記述したように、雑種および親ガエル 2 種の精巣の small RNA-seq を行った。ちなみに、mRNA-seq と small RNA-seq は同じ精巣由来の RNA の情報である。得られた small RNA-seq 配列から、miRNA、rRNA、tRNA、snoRNA、tRNA 由来の配列を除き、更に 24-32 塩基長の配列を選択し、piRNA 配列として解析に用いた。次に、piRNA 配列をワライガエルおよびコガタガエルのゲノムにマッピングし、標的のトランスポゾンを同定した。次に、それぞれのトランスポゾンに対する piRNA の発現量を、 $\text{Log}_2(\text{Counts Per Million mapped reads [CPM]}+1)$ として、片親種と雑種での平均発現量を横軸に、片親種の発現に対する雑種の発現の比を対数で表したものを縦軸とした MA plot を作製した (図 10: ワライガエルと雑種のプロットが A、コガタガエルと雑種のプロットが B)。多くのトランスポゾンサブファミリー対応する piRNA 発現が、ワライガエル、コガタガエル共に雑種で大きな増減が見られ、更に、プロットの概観においては、増減の幅がワライガエルに比べコガタガエルでより大きいように観察された。そこで、1,509 個のサブファミリーに関して、対応する piRNA の発現量比を、1/2 倍以下、1/2-1 倍、1-2 倍、2 倍以上に分け、サブファミリー数を調べたところ、雑種で発現量が 2 倍以上減ったのは、ワライガエルゲノム由来が 10.9%、コガタガエルゲノム由来が 21.3%、2 倍以上増えたのは、ワライガエルで 6.2%、コガタガエルで 14.8%もあった (表 13)。これらを合わせると、2 倍以上の増減は、ワライガエルで 17.1%、コガタガエルで 36.1%であり、ワライガエルに比べ、コガタガエルゲノム由来のトランスポゾンに対する piRNA の発現の増減変動がより大きいことが示唆された。次に、このゲノム間の piRNA 発現量の変動のばらつきが、各ゲノム由来のトランスポゾンの発現に関連性があるかを調べた。雑種あるいは親種において、トランスポゾンの発現とそれに対応する piRNA の発現が検出される 879 種の共通のトランスポゾンサブファミリーに対して、発現比のばらつきを

ボックスプロットし、Fテストを行った（図 10C）。その結果、コガタガエルゲノム由来の各トランスポゾンに対応する piRNA は、ワライガエルのそれと比較して、有為に発現変動のばらつきが、より大きいことが示された（図 10C 左）。DNA トランスポゾンとレトロトランスポゾンを別々に解析した結果も同様であった（図 10C 中央・右）。

ショウジョウバエの雑種では、piRNA 量の低下がトランスポゾン脱抑制に繋がり、次世代産生が不能になる例がある（Khurana JS et al., 2012）。このことと、ツメガエル異種交配での DNA トランスポゾンの活性化という私の研究知見（Suda et al., 2022; 図 3, 4）から、雑種ヨーロッパトノサマガエルでは、トランスポゾンの脱抑制が、生殖系列細胞でのコガタガエルゲノムの欠失に関与するという作業仮説を考えた。そこで、雑種の精巣で発現するトランスポゾンで、piRNA の発現量が大きく低下したトランスポゾンサブファミリーに注目した。対応する piRNA 発現量が 4 倍以上低下したサブファミリーは、ワライガエルゲノム由来で 34 個、コガタガエルゲノム由来で 98 個であり、約 2.9 倍、コガタガエルゲノム由来のサブファミリーが多かった（表 14）。発現低下したサブファミリーを DNA トランスポゾンとレトロトランスポゾンに分類したところ、両ゲノム共に、若干 DNA トランスポゾンが少ない傾向が見られた。さらに、雑種の精巣で、親種に比べ、piRNA 発現量が 4 倍以上低下し、対応するトランスポゾンの発現量が 2 倍以上に上昇したサブファミリー数は、ワライガエルゲノム由来のサブファミリーの 1 個に対し、コガタガエルゲノム由来は 17 個で、コガタガエルゲノム由来が圧倒的に多かった（表 14）。この結果から、雑種ヨーロッパトノサマガエルの精巣では、piRNA の発現低下によりコガタガエルゲノムのトランスポゾンが脱抑制されている可能性が考えられる。

5. 全体の考察 および 仮説提唱

【ツメガエル異種交配後の L/S サブゲノム非対称 DNA 欠失における分子進化モデル】

第1章で記述したツメガエル属の異質四倍体2種と二倍体1種のゲノム解析から、私は、ツメガエル異種交配を介した異質四倍体化から現在までのサブゲノム非対称進化のモデルを以下に提案・考察する（図 11）。アフリカツメガエルとキタアフリカツメガエルの3,400 万年以上前の二倍体の共通祖先は、約 3,400 万年前から約 1,800~1,700 万年前の間に、生殖隔離を介して二倍体の L ゲノムをもつ種と二倍体の S ゲノムをもつ種に種分化した (Session et al., 2016)。即ち、L ゲノムあるいは S ゲノム特異的な [トランスポゾン-piRNA システム] がそれぞれのゲノムで構築された（図 3 - 6）。そして、種分化した L ゲノムと S ゲノムをもつツメガエル 2 種の異種交配が、約 1,800~1,700 万年前に起こった (Session et al., 2016)。異種交配個体は異なるトランスポゾン進化を介したゲノム進化により、異種ゲノム間の対合が不可能で通常は次世代ができなかったと予想されるが、ある個体で、異質四倍化した生殖幹細胞が減数分裂して、異質二倍化した生殖細胞が形成された。異種ゲノムが生殖系列細胞内で混合され、それぞれの [トランスポゾン-piRNA システム] の不和合性（異種ゲノム間闘争）からか、トランスポゾン抑制システムの変調（一部崩壊）が起こったと予想される。これにより、DNA トランスポゾンの活性化（脱抑制）が L/S サブゲノム両者で起きた (Suda et al., 2022; 図 4)。DNA トランスポゾンのカットアンドペースト形式の増殖方法は、染色体の二重鎖切断の頻度が増え、DNA トランスポゾンの隣接した遺伝子間領域や遺伝子の欠失や染色体逆位が引き起こされた（表 5, 6; 図 2, 3, 4）。L サブゲノムに比べ、S サブゲノムにおける DNA トランスポゾン脱抑制がより多かったと予想される。異種交配後に活性化された DNA トランスポゾンは、性決定機構が不安定な異種交配後において、性決定遺伝子 *dm-W* の誕生に直接関与した (Mawaribuchi et al., 2017; Hayashi et al., 2022; Hayashi et al., 2023)。

S/L サブゲノムに起きた遺伝子の大規模な欠失は、異種交配個体およびその次世代種の交配後の集団で、短期間に起きた（図 2）。これは、異種交配してから 100 万年たらずで種分岐したアフリカツメガエルとキタアフリカツメガエルにおいて、2 種のゲノム解析より、欠失遺伝子は各サブゲノム間でのオーソログで多数共通していた点から推定される。また、種分岐後にもそれぞれの種の祖先で独立に多くの遺伝子欠失が起こった。欠失遺伝子の特徴は、欠失時期や L/S サブゲノムを含めたジーンオントロジーあるいはパスウエー解析からは検出されなかった（データ示さず）。一方、アフリカツメガエルゲノムで、二倍体のネッタイツメガエルゲノムのタンパク質コード遺伝子のオーソログの維持率が 99.7%であること（表 5）から、個体・集団維持に必要な遺伝子には、どちらかのサブゲノムには少なくとも 1 つ残存する選択圧があったと考えられる。すなわち、異種交配ツメガエルゲノムの DNA 欠失は、遺伝子を標的にしていたのではない、と考えられる。後述するトランスポゾンを介した遺伝子欠失モデルは、この遺伝子欠失のランダム性に合致する。

L サブゲノムよりも S サブゲノムで活発な DNA トランスポゾンの活動 (図 5、6) により、S サブゲノムにより多い遺伝子欠失が、異種交配直後だけでなく、その後も現在まで継続されてきた (図 2)。このようなサブゲノム非対称進化は、これからも緩やかに継続していくと予想される。

一般的に、遺伝子重複では、一方の遺伝子が機能維持し、もう一方が機能分化、新機能獲得、偽遺伝子化あるいは遺伝子欠失という運命を辿ることが知られている (Lynch M and Force A., 2000; Lynch et al., 2001)。ツメガエルの異種交配を介した全ゲノム重複 (異質四倍体化) では、ほぼ全遺伝子が重複したと考えられるが、異種交配後に、トランスポゾンを通じて、多くが遺伝子欠失の運命となったと考えられる。一方、新機能獲得遺伝子は、性決定遺伝子 (Yoshimoto et al., 2008; Hayashi et al., 2022) 以外はほぼないと考えられるが、機能分化している遺伝子はいくつか報告されている (Session et al., 2016)。これは、異質四倍体化後、種分化の過程でホメオログが機能分担されてきたと考えられ、その場合は、遺伝子欠失の選択圧を免れたと考えることもできる。S サブゲノム由来の遺伝子の機能低下 (相補性の低下) や遺伝子発現の減少が一部報告されている (Session et al., 2016) が、今後、異種交配後に間もなく種分岐したアフリカツメガエルとキタアフリカツメガエルとの間で機能分化したホメオログの共通性や相違性を解析することは、非対称サブゲノム進化、遺伝子の分子進化・機能分化、種分化、および環境適応との連関を考察する上で、極めて興味深いことと考えられる。

【雑種ヨーロッパトノサマガエルにおける片親ゲノム選択的排除の分子機構モデル】

第二章の研究では、私は、トノサマガエル属のワライガエルとコガタガエルのゲノムおよび、これら 2 種とその雑種のヨーロッパトノサマガエルの RNA-seq 解析を行った。この雑種の配偶子でのコガタガエルゲノム排除 (片親ゲノム選択的排除) は、雑種生成 (hybridogenesis) と呼ばれ、ヨーロッパの研究者が何十年も前に報告してきた (Uzzell et al., 1980) が、現在も未解明の古典的命題となっている。私は、この未だ謎とされる古典的命題に対して、本研究結果から、その分子機構のモデルを以下に提案する (図 12)。

まず、親ガエル 2 種のゲノム解析とゲノム内トランスポゾンの進化 (増殖・活性化・変異および化石化) の解析 (図 8; 表 10, 11) から、約 800 万年に種分岐したと考えられている親ガエル 2 種は、生殖隔離により、それぞれ独自のトランスポゾンの進化が起り、新たに活性化したトランスポゾンを抑制する [トランスポゾン-piRNA システム] を、それぞれのゲノム内で再構築してきた。この親 2 種の雑種であるヨーロッパトノサマガエルでは、生殖系列細胞内で、2 種の異なるゲノムが混在する状況となり、[トランスポゾン-piRNA システム] に不和合性が生じたと考えることができる。この不和合性は、トランスポゾンを抑制する piRNA 生成・発現の混乱 (図 10) と考えると、piRNA の生成・発現の違いがワライガエルとコガタガエルのトランスポゾン発現の違いを引き起こすことになる (図 9)。特に、コガタガエルゲノム由来のトランスポゾンに対する piRNA 発現量の低下が対応するト

ランスポゾンの脱抑制を引き起こす (図 10; 表 12, 14)。異質四倍体ツメガエルで触れたように DNA トランスポゾンの活性化は二重鎖切断を伴うため、コガタガエルゲノムに選択的にゲノム損傷が起き、ゲノム排除に繋がった (図 12)。

【無尾両生類の 2 つのタイプの異種交配における考察：ゲノム不和合性】

私は、2 種類のタイプの異種交配における片親ゲノム選択的 DNA 欠失の研究の考察から、近縁種間の異種交配における DNA 欠失の原因は、近縁種ゲノム間における“ゲノム不和合性”であると提案する。これは、“ゲノム間自己・非自己認識、ゲノム間アロ認識、あるいはゲノム間闘争”とも言い換えられるような概念で、動物界ではほとんど言及されていないものである。本研究における無尾両生類の 2 種類の異種交配系で観察される一方のサブゲノムの選択的 DNA 欠失と一方のサブゲノムの排除という一見異なった現象が、どちらも異種交配時のトランスポゾン制御システムの変容あるいは崩壊による、という考えに基づいて、“ゲノム不和合性”の実体を、それぞれの種のゲノムがもつ [トランスポゾン-piRNA システム] の相違と想定する。節足動物であるショウジョウバエ (*Drosophila*) 属では、トランスポゾンを抑制する piRNA 産生に、Rhino-Deadlock 複合体が関与しており、*D. melanogaster* と *D. simulans* の不妊性交雑種での次世代の非産生は、*D. simulans* の Deadlock と *D. melanogaster* の Rhino の組み合わせを介した不和合性によると報告されている (Parhad et al., 2017)。即ち、“ゲノム不和合性”は、旧口動物 (前口動物) を含む生殖細胞の生産器官である生殖巣をもつ真生後生動物 (海綿動物を除く後生動物) に適応できる可能性がある。

このゲノム不和合性による DNA 欠失の分子機構については、主に DNA トランスポゾンの活性化の際の二重鎖切断を想定している。これは、CRISPR-Cas9 による二重鎖切断において、小核形成を伴った染色体欠失を引き起こすことが報告 (Papathanasiou et al. 2021) から、機序として妥当性があると考えられる。

次に、異種交配における DNA 部分欠失とゲノム排除の違いを [トランスポゾン-piRNA システム] という観点から、考察したいと思う。今回解析したツメガエル属あるいはトノサマガエル属の種分岐から交雑までの時間は、それぞれ約 1,800~1,700 万年あるいは約 800 万年であり、ゲノム欠失が一部であるツメガエル属の方が異種交配までの時間が、むしろ長い。しかし、種分岐から異種交配までのトランスポゾン活性の違いは、ツメガエル属 L/S サブゲノム間 (図 4B: 0.1 ~ 0.05) とワライガエル/コガタガエル間 (図 8 A, B: 0.026 ~ 0) の方がランドスケープの形態が異なる。このことから、ツメガエル属と比較してトノサマガエル属の異種交配は種特有トランスポゾンの活性化が多い種の交雑であると予想される。したがって、非対称的なサブゲノム退縮あるいは排除という、程度の異なる 2 種のサブゲノム進化には、トランスポゾン脱抑制量や種類が関与している可能性が考えられる。piRNA 産生機構がよく研究されているショウジョウバエでは、トランスポゾン脱抑制機構が 2 種間で異なり、その piRNA 産生機構が 2 種間の適応性の違いに関わることが報告されている

(Parhad et al., 2017)。したがって、2 種類の無尾両生類の異種交配系でも、piRNA 産生機構の相違が DNA 部分欠失とゲノム排除との相違に関与する可能性も考えられる。

ツメガエル属の異質四倍体のアフリカツメガエルと二倍体のネッタイツメガエルとの実験的な異種交配では、雌雄（卵子と精子）の組み合わせが、卵子がアフリカツメガエル、精子がネッタイツメガエルのとき正常発生するのに対し、その逆では、胚発生過程で、アフリカツメガエルゲノムに DNA 欠失が起こり、致死に至ることが報告されている (Gibeaux et al., 2018)。この現象を、[トランスポゾン-piRNA システム] の関与があったとして考察すると、卵子がネッタイツメガエルのときは、ネッタイツメガエルの卵母細胞に存在した piRNA 種では、アフリカツメガエルのゲノム由来のトランスポゾンを抑制しきれなかったと解釈できるかもしれない。この解釈の検証を行うことは、本研究で提案する [トランスポゾン-piRNA システム] を基盤とする“ゲノム不和合性”の検証となるだけでなく、生殖細胞だけでなく、初期発生に [トランスポゾン-piRNA システム] が重要な機能を担うことを示唆する可能性もあり、興味深く思われる。

【今後の展望】

本研究では、[トランスポゾン・piRNA システム]の不和合性がサブゲノム非対称進化の原因であると予想したが、piRNA 発現の変動が片親に偏る分子機構は未解明である。ショウジョウバエの不妊性交雑種は母系の piRNA 発現が子孫に引き継がれることで片親ゲノムのトランスポゾンに対する不和合性が起こることが報告されている。しかし、今回用いた雑種ヨーロッパトノサマガエルはワライガエルとコガタガエルの雌雄の組み合わせによらずゲノム排除が起こる。そのため、片親の piRNA 発現を引き継いだ故の不和合性ではなく、両方のサブゲノムの piRNA 発現によってトランスポゾンを抑制する閾値に達していないことが原因であると予想する。つまり、ゲノム排除は、ワライガエルとコガタガエルのゲノムが 1:1 であり、コガタガエル由来の piRNA 不足が原因であると考えられる。実際、ヨーロッパトノサマガエルは 3 倍体も存在するが、コガタガエルのサブゲノムが 2 倍である RLL のヨーロッパトノサマガエルではワライガエルの不完全なゲノム排除も観察されるようになる (Dedukh et al., 2020)。この仮説を証明するために、今後はヨーロッパトノサマガエルの piRNA 発現量が親の雌雄の組み合わせで変わらないことを検証していきたい。

【最後に：有性生殖における“種”を考える】

地球上では、現在、少なくとも何百万の種が存在する。絶滅した種を考えると、生命進化でさらに多くの種が誕生してきたことになる。種の定義・概念は、分類群や観点によって様々であり、“交配可能かどうか”という指標に普遍性はなく、生物界で共通した概念はない。しかし、本研究で言及してきた [トランスポゾン-piRNA システム] を基盤とする“ゲノム不和合性”は、有性生殖する真生後生動物の種の違いの定義・概念に、ある程度、分子レベルで合致するものではないかと思われる。

6. 参考文献

- 1 Alger, E. I. & Edger, P. P. One subgenome to rule them all: underlying mechanisms of subgenome dominance. *Curr Opin Plant Biol* 54, 108-113 (2020).
<https://doi.org/10.1016/j.pbi.2020.03.004>
- 2 Alonge, M. et al. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol* 23, 258 (2022).
<https://doi.org/10.1186/s13059-022-02823-7>
- 3 Belancio, V. P., Hedges, D. J. & Deininger, P. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome Res* 18, 343-358 (2008).
<https://doi.org/10.1101/gr.5558208>
- 4 Berger, L. Some characteristics of the crosses within *Rana esculenta* complex in postlarval development. Vol. 27 (Państwowe Wydawnictwo Naukowe, 1969).
- 5 Bouchard-Bourelle, P. et al. snoDB: an interactive database of human snoRNA sequences, abundance and interactions. *Nucleic Acids Res* 48, D220-D225 (2020).
<https://doi.org/10.1093/nar/gkz884>
- 6 Buchfink, B., Reuter, K. & Drost, H. G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 18, 366-368 (2021). <https://doi.org/10.1038/s41592-021-01101-x>
- 7 Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421 (2009). <https://doi.org/10.1186/1471-2105-10-421>
- 8 Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17, 540-552 (2000).
<https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- 9 Chan, P. P. & Lowe, T. M. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res* 37, D93-97 (2009). <https://doi.org/10.1093/nar/gkn787>
- 10 Cheng, F. et al. Gene retention, fractionation and subgenome differences in polyploid plants. *Nat Plants* 4, 258-268 (2018). <https://doi.org/10.1038/s41477-018-0136-7>
- 11 Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* 10 (2021).
<https://doi.org/10.1093/gigascience/giab008>
- 12 Dedukh, D. et al. Variation in hybridogenetic hybrid emergence between populations of water frogs from the *Pelophylax esculentus* complex. *PLoS One* 14, e0224759 (2019).
<https://doi.org/10.1371/journal.pone.0224759>
- 13 Dooner, H. K. et al. Distribution of unlinked receptor sites for transposed Ac elements from the bz-m2(Ac) allele in maize. *Genetics* 136, 261-279 (1994).
<https://doi.org/10.1093/genetics/136.1.261>

- 14 Flynn, J. M. et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A* 117, 9451-9457 (2020).
<https://doi.org/10.1073/pnas.1921046117>
- 15 Fromm, B. et al. MirGeneDB 2.0: the metazoan microRNA complement. *Nucleic Acids Res* 48, D132-D141 (2020). <https://doi.org/10.1093/nar/gkz885>
- 16 Gibeaux, R. et al. Paternal chromosome loss and metabolic crisis contribute to hybrid inviability in *Xenopus*. *Nature* 553, 337-341 (2018). <https://doi.org/10.1038/nature25188>
- 17 Hayashi, S. et al. Neofunctionalization of a Noncoding Portion of a DNA Transposon in the Coding Region of the Chimerical Sex-Determining Gene dm-W in *Xenopus* Frogs. *Mol Biol Evol* 39 (2022). <https://doi.org/10.1093/molbev/msac138>
- 18 Hayashi, S. et al. Promoter generation for the chimeric sex-determining gene dm-W in *Xenopus* frogs. *Genes Genet Syst* 98, 53-60 (2023). <https://doi.org/10.1266/ggs.22-00137>
- 19 Hoff, K. J., Lomsadze, A., Borodovsky, M. & Stanke, M. Whole-Genome Annotation with BRAKER. *Methods Mol Biol* 1962, 65-95 (2019). https://doi.org/10.1007/978-1-4939-9173-0_5
- 20 Holland, P. W., Garcia-Fernández, J., Williams, N. A. & Sidow, A. Gene duplications and the origins of vertebrate development. *Dev Suppl*, 125-133 (1994).
- 21 Iwasaki, Y. W., Siomi, M. C. & Siomi, H. PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu Rev Biochem* 84, 405-433 (2015). <https://doi.org/10.1146/annurev-biochem-060614-034258>
- 22 Jurka, J. et al. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110, 462-467 (2005). <https://doi.org/10.1159/000084979>
- 23 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30, 3059-3066 (2002).
<https://doi.org/10.1093/nar/gkf436>
- 24 Khurana, J. S. et al. Adaptation to P element transposon invasion in *Drosophila melanogaster*. *Cell* 147, 1551-1563 (2011). <https://doi.org/10.1016/j.cell.2011.11.042>
- 25 Kotov, A. A. et al. piRNA silencing contributes to interspecies hybrid sterility and reproductive isolation in *Drosophila melanogaster*. *Nucleic Acids Res* 47, 4255-4271 (2019).
<https://doi.org/10.1093/nar/gkz130>
- 26 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10, R25 (2009).
<https://doi.org/10.1186/gb-2009-10-3-r25>
- 27 Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079 (2009). <https://doi.org/10.1093/bioinformatics/btp352>

- 28 Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658-1659 (2006).
<https://doi.org/10.1093/bioinformatics/btl158>
- 29 Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151-1155 (2000). <https://doi.org/10.1126/science.290.5494.1151>
- 30 Lynch, M. & Force, A. The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154, 459-473 (2000). <https://doi.org/10.1093/genetics/154.1.459>
- 31 Lynch, M., O'Hely, M., Walsh, B. & Force, A. The probability of preservation of a newly arisen gene duplicate. *Genetics* 159, 1789-1804 (2001).
<https://doi.org/10.1093/genetics/159.4.1789>
- 32 Mawaribuchi, S. et al. Sex chromosome differentiation and the W- and Z-specific loci in *Xenopus laevis*. *Dev Biol* 426, 393-400 (2017). <https://doi.org/10.1016/j.ydbio.2016.06.015>
- 33 Miura, I. et al. Hybridogenesis in the Water Frogs from Western Russian Territory: Intrapopulation Variation in Genome Elimination. *Genes (Basel)* 12 (2021).
<https://doi.org/10.3390/genes12020244>
- 34 Nakatani, Y. et al. Reconstruction of proto-vertebrate, proto-cyclostome and proto-gnathostome genomes provides new insights into early vertebrate evolution. *Nat Commun* 12, 4489 (2021). <https://doi.org/10.1038/s41467-021-24573-z>
- 35 Ohno, S. *Evolution by gene duplication*. (Springer Berlin, Heidelberg, 1970).
- 36 Papathanasiou, S. et al. Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nat Commun* 12, 5855 (2021).
<https://doi.org/10.1038/s41467-021-26097-y>
- 37 Parhad, S. S., Tu, S., Weng, Z. & Theurkauf, W. E. Adaptive Evolution Leads to Cross-Species Incompatibility in the piRNA Transposon Silencing Machinery. *Dev Cell* 43, 60-70.e65 (2017). <https://doi.org/10.1016/j.devcel.2017.08.012>
- 38 Price, A. L., Jones, N. C. & Pevzner, P. A. De novo identification of repeat families in large genomes. *Bioinformatics* 21 Suppl 1, i351-358 (2005).
<https://doi.org/10.1093/bioinformatics/bti1018>
- 39 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26, 1641-1650 (2009).
<https://doi.org/10.1093/molbev/msp077>
- 40 Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41, D590-596 (2013).
<https://doi.org/10.1093/nar/gks1219>
- 41 Quinlan, A. R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr Protoc Bioinformatics* 47, 11.12.11-34 (2014). <https://doi.org/10.1002/0471250953.bi1112s47>

- 42 Ranwez, V., Harispe, S., Delsuc, F. & Douzery, E. J. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PLoS One* 6, e22594 (2011). <https://doi.org/10.1371/journal.pone.0022594>
- 43 Rodriguez, F. & Arkhipova, I. R. Transposable elements and polyploid evolution in animals. *Curr Opin Genet Dev* 49, 115-123 (2018). <https://doi.org/10.1016/j.gde.2018.04.003>
- 44 Sahlin, K. Strobealign: flexible seed size enables ultra-fast and accurate read alignment. *Genome Biol* 23, 260 (2022). <https://doi.org/10.1186/s13059-022-02831-7>
- 45 Session, A. M. et al. Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* 538, 336-343 (2016). <https://doi.org/10.1038/nature19840>
- 46 Simakov, O. et al. Deeply conserved synteny resolves early events in vertebrate evolution. *Nat Ecol Evol* 4, 820-830 (2020). <https://doi.org/10.1038/s41559-020-1156-z>
- 47 Spring, J. Vertebrate evolution by interspecific hybridisation—are we polyploid? *FEBS Lett* 400, 2-8 (1997). [https://doi.org/10.1016/s0014-5793\(96\)01351-8](https://doi.org/10.1016/s0014-5793(96)01351-8)
- 48 Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 35, 1026-1028 (2017). <https://doi.org/10.1038/nbt.3988>
- 49 Suda, K., Hayashi, S. R., Tamura, K., Takamatsu, N. & Ito, M. Activation of DNA Transposons and Evolution of piRNA Genes Through Interspecific Hybridization in. *Front Genet* 13, 766424 (2022). <https://doi.org/10.3389/fgene.2022.766424>
- 50 Tower, J., Karpen, G. H., Craig, N. & Spradling, A. C. Preferential transposition of *Drosophila* P elements to nearby chromosomal sites. *Genetics* 133, 347-359 (1993). <https://doi.org/10.1093/genetics/133.2.347>
- 51 Uzzell, T., Hotz, H. & Berger, L. Genome exclusion in gametogenesis by an interspecific *Rana* hybrid: Evidence from electrophoresis of individual oocytes. *Journal of Experimental Zoology* 214, 251-259 (1980). <https://doi.org/https://doi.org/10.1002/jez.1402140303>
- 52 Wang, X. miRDB: a microRNA target prediction and functional annotation database with a wiki interface. *RNA* 14, 1012-1017 (2008). <https://doi.org/10.1261/ma.965408>
- 53 Wang, Z. et al. Subgenome dominance and its evolutionary implications in crop domestication and breeding. *Hortic Res* 9, uhac090 (2022). <https://doi.org/10.1093/hr/uhac090>
- 54 Yan, H., Bombarely, A. & Li, S. DeepTE: a computational method for de novo classification of transposons with convolutional neural network. *Bioinformatics* 36, 4269-4275 (2020). <https://doi.org/10.1093/bioinformatics/btaa519>
- 55 Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13, 555-556 (1997). <https://doi.org/10.1093/bioinformatics/13.5.555>
- 56 Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24, 1586-1591 (2007). <https://doi.org/10.1093/molbev/msm088>

57 Yoshimoto, S. et al. A W-linked DM-domain gene, DM-W, participates in primary ovary development in *Xenopus laevis*. *Proc Natl Acad Sci U S A* 105, 2469-2474 (2008).
<https://doi.org/10.1073/pnas.0712244105>

7. 謝辞

本研究を行うにあたり、指導教授として多岐にわたる御指導、御鞭撻を賜りました北里大学理学部分子生物学講座の 松尾拓哉 教授 及び 高松信彦 前教授に深く御礼申し上げます。研究テーマや研究方針等、指導教官として、終始熱心な研究指導、御教示を賜りました 伊藤道彦 准教授に厚く御礼申し上げます。また、本研究の完成に必要な御助言を賜りました 田村啓 講師および 塚本大輔 助教に厚く御礼申し上げます。本研究のヨーロッパノサマガエルの研究に関しましては、サンプルの供与や御指導を賜りました広島大学両生類研究センターの 三浦郁夫 教授とウラル連邦大学の Vladimir Vershinin 教授に深く感謝申し上げます。また、本博士論文の審査に際して、熱心な御指導、御批評を賜りました北里大学医学部の 勝村啓史 准教授、北里大学理学部免疫学講座の 江島耕二 教授、広島大学両生類研究センターの 三浦郁夫 教授に深く御礼申し上げます。

8. 表

表 1. ツメガエル (*Xenopus*) 属の 3 種 (*X. tropicalis*、*X. laevis*、および *X. borealis*) における各染色体の DNA 長 (塩基対)

表 2: ネットアイツメガエル (*X. tropicalis*) ゲノムデータベース上の各染色体における遺伝子と転移因子の含有割合 (%)

表 3: アフリカツメガエル (*X. laevis*) ゲノムデータベース上の L/S サブゲノム各染色体における遺伝子と転移因子の含有割合 (%)

表 4: キタアフリカツメガエル (*X. borealis*) ゲノムデータベース上の L/S サブゲノム各染色体における遺伝子と転移因子の含有割合 (%)

表 5: ネットアイツメガエル (*X. tropicalis*) ゲノムと比較したアフリカツメガエル (*X. laevis*) およびキタアフリカツメガエル (*X. borealis*) ゲノムに保持あるいは欠失されたオーソログ数

*タンDEM重複遺伝子および"LOC"として登録されていない遺伝子を除いたタンパク質コーディング遺伝子数

**ネットアイツメガエルの 15,022 個のタンパク質コーディング遺伝子に対するオーソログ数

表 6. ネットアイツメガエルの遺伝子間の DNA 長に対するアフリカツメガエルの L あるいは S サブゲノムの遺伝子間の DNA 長の減少率

* ネットアイツメガエルゲノムとアフリカツメガエルの L あるいは S サブゲノムにおける対応するオーソログ間の DNA 長を染色体ごとに比較し、L あるいは S サブゲノムの DNA 長の減少率を示した。

表 7. 図 5 において相関係数が 0.75 以上の L サブゲノム由来のトランスポゾンサブファミリーの相関係数と近似直線の傾き

* 対応する S サブゲノム由来のトランスポゾンサブファミリーの相関係数と近似直線の傾きを右側に示す。

表 8. 図 5 において相関係数が 0.75 以上の S サブゲノム由来のトランスポゾンサブファミリーの相関係数と近似直線の傾き

*対応する L サブゲノム由来のトランスポゾンサブファミリーの相関係数と近似直線の傾きを右側に示す。

表 9. 無尾両生類 8 種の 597 個のオーソログのタンパク質コード配列 (CDS) を用いて計算された同義置換率 (dS)

Acat, *Aquarana catesbeianus*; Rtem, *Rana temporaria*; Prid, *Pelophylax. ridibundus*; Ples, *Pelophylax lessonae*; Npar, *Nanorana parkeri*; Xtro, *X. tropicalis*; XlaeL, *X. laevis* L subgenome; XlaeS, *X. laevis* S subgenome

表 10. トノサマガエル属 2 種の遺伝子及びトランスポゾンの割合 (%)

P. ridibundus, ワライガエル; *P. lessonae*, コガタガエル

表 11. ワライガエルとコガタガエルの種分岐後に互いに 1.5 倍以上ゲノム内存在量の差があったトランスポゾンサブファミリーの数

表 12. 雑種精巢のトランスポゾン発現量の親種（ワライガエルあるいはコガタガエル）との比を基にしたサブファミリー数と%

表 13. トランスポゾン対応 piRNA の雑種精巢の発現量の親種（ワライガエルあるいはコガタガエル）との比を基にしたサブファミリー数と%

表 14. piRNA 発現量が親種に比べ雑種で 4 倍以下に減少したトランスポゾンサブファミリーの数

※ 括弧内の数字は piRNA 発現量が 4 倍以下かつ図 8 にて発現量が 2 倍以上になったトランスポゾンサブファミリーの数

表 15. 雑種精巣における piRNA 発現量が 1/4 かつトランスポゾン発現量が 2 倍以上のコガタガエルゲノム由来トランスポゾンサブファミリー3 種と各染色体内での発現箇所数

rid, ワライガエル ; les, コガタガエル : *esculentus*, 雑種ヨーロッパトノサマガエル

9. 図

図 1. ツメガエル (*Xenopus*) 属の 3 種における (サブ) ゲノム内の DNA 成分の割合
X. tropicalis, ネットイツメガエル; *X. laevis*, アフリカツメガエル ; *X. borealis*, キタア
フリカツメガエル)

図 2. アフリカツメガエルおよびキタアフリカツメガエルの L/S サブゲノムにおけるネッタイツメガエルのオーソログ遺伝子に対応する失われた遺伝子の予測数

予測された失われた遺伝子数は、Venn 図 (A) およびのツメガエル属 3 種の進化系統樹 (B) で、L サブゲノム由来は赤、S サブゲノム由来は青で示す。

A

B

図3. ネットアイツメガエル (*X. tropicalis*) ゲノム、アフリカツメガエル (*X. laevis*) あるいはキタアフリカツメガエル (*X. borealis*) のL/Sサブゲノムにおける染色体逆位の模式図

4つの組合せ (*X. laevis* S - *X. tropicalis* - *X. borealis* S ; *X. laevis* L - *X. tropicalis* - *X. borealis* L; *X. laevis* L - *X. tropicalis* - *X. borealis* S; *X. laevis* S - *X. tropicalis* - *X. borealis* L) において遺伝子シンテニーを基盤として、各染色体の並びを調べ、一直線に並べて図示した (A)。L/S 共通 (緑色)、S 固有 (青色)、および L 固有 (赤色) をネットアイツメガエルの染色体に表示した (B)。

図 4. ツメガエル属 3 種ゲノムにおける DNA トランスポゾンとレトロトランスポゾンのランドスケープ.

二倍体のネッタイツメガエル (*X. tropicalis*) はゲノム (A)、異質四倍体のアフリカツメガエル (*X. laevis*) あるいはキタアフリカツメガエル (*X. borealis*) は L/S サブゲノム(それぞれ、B あるいは C) での結果を示す。横軸は DNA トランスポゾン (青) とレトロトランスポゾン (赤) の置換比率を、縦軸は。各 (サブ) ゲノムの DNA に対するトランスポゾンの割合 (%) を示す。17 Mya (million years ago) は、置換比率 0.05 に対応し、予測される異質四倍体化 (異種交配) 時期を表す。

図 5. アフリカツメガエル S と L サブゲノムにおける置換率ごとの転移因子 (TE; transposable element) の総長の比較

* 横軸の値: コンセンサス配列に対する DNA 塩基置換率

* 縦軸の値 (bp) = [(S サブゲノム TE 長/ TE 削除 S サブゲノム DNA 長)]

- [(L サブゲノム TE 長/ TE 削除 L サブゲノム DNA 長) x 1 G bp.

* 青と赤はそれぞれ DNA トランスポゾンとレトロトランスポゾンを表す。

図 6. アフリカツメガエルの L/S サブゲノムにおける S サブゲノム特異的 DNA トランスポゾン *Tc1-14_Xt* のランドスケープ

横軸はコンセンサス配列に対する置換率を、縦軸は bp を示す。青は S サブゲノム、赤は L サブゲノムを示す。

* この *Tc1/mariner* のサブファミリーは、異種交配期に活性化のピークをもつ。置換率が高くなるほど、L サブゲノムで検出されているのは、置換率が高いゆえに、同起源あるいは極近縁の *Tc1/mariner* が若干検出されていると考えられる。異種交配後、現代までほぼ L サブゲノムにほとんど検出されていないことは、注目に値する。

図 7. アフリカツメガエルの L/S サブゲノムの各染色体における DNA 欠失とトランスポゾンサブファミリーの長さの関係を示す散布図

*それぞれのトランスポゾンサブファミリーごとの近似直線の作製

横軸：L/S 各染色体に対応するネッタイツメガエルの染色体に対する減少率

($1 - \text{L/S 各染色体の長さ} / \text{対応するネッタイツメガエルの染色体の長さ}$)

縦軸：L/S 各染色体上の各トランスポゾンサブファミリーの長さ / L/S サブゲノム上の各トランスポゾンサブファミリーの長さ

*図 6：それぞれのトランスポゾンファミリーにおいて、上記のように近似直線を作製し、その傾き（横軸）と相関係数（縦軸）を求め、散布図を作製した。青は DNA トランスポゾン、赤はレトロトランスポゾンを表す。

図 8. ヨーロッパトノサマガエルの親 2 種であるワライガエルとコガタガエルゲノムにおけるトランスポゾンのランドスケープ

ワライガエル(A)とコガタガエル(B)のゲノムに存在する DNA トランスポゾン(青線)とレトロトランスポゾン(赤線)の塩基置換率を x 軸に、それぞれのゲノムに対する塩基置換率に対応したトランスポゾンの割合を y 軸に示す。下部の“8”は、塩基置換率 0.0262 に対応し、2 種が種分岐した 800 万年前を示す。

C

図 9. 雑種ヨーロッパトノサマガエルと親種（ワライガエルあるいはコガタガエル）の精巢における各トランスポゾンサブファミリーの発現量の比較解析

- (A) ワライガエルゲノム由来トランスポゾンサブファミリーにおけるワライガエル精巢に対する雑種精巢での発現量の MA プロットによる比較
- (B) コガタガエルゲノム由来トランスポゾンサブファミリーにおけるコガタガエル精巢に対する雑種精巢での発現量の MA プロットによる比較。
- (C) (A)と(B)の縦軸におけるボックスプロットと F テストによる分散解析

* ワライガエルおよびコガタガエルゲノム由来トランスポゾンサブファミリーの中で、雑種あるいは親種で発現する共通のトランスポゾンサブファミリーに関して解析を行った。

DNA TE は DNA トランスポゾン、Retro TE は、レトロトランスポズンを意味する。

図 10. 雑種ヨーロッパトノサマガエルと親種（ワライガエルあるいはコガタガエル）の精巣における各トランスポゾンサブファミリーに対応する piRNA 発現量の比較解析

(A) ワライガエルゲノムのトランスポゾンサブファミリーに対応する piRNA におけるワライガエル精巣に対する雑種精巣での発現量の MA プロットによる比較

(B) コガタガエルゲノムのトランスポゾンサブファミリーに対応する piRNA におけるコガタガエル精巣に対する雑種精巣での発現量の MA プロットによる比較

* A および B の MA プロットの横軸は 1,509 個のトランスポゾンサブファミリーに対応する piRNA のそれぞれについての片親種と雑種での平均発現量、縦軸は片親種の piRNA 発現に対する雑種の発現の比を対数で表した。

* A および B の黒枠のある各トランスポゾンサブファミリーの点は、雑種精巣においてサブファミリーの発現量($\log_2(\text{RPKM}+1)$)が 1 以上のものを示す。

(C) (A)と(B)の縦軸におけるボックスプロットと F テストによる分散解析

* ワライガエルおよびコガタガエルゲノム由来トランスポゾンサブファミリーの中で、DNA トランスポゾンとレトロトランスポゾンをピックアップした解析でも、同様であった（図 7C 中央・右）。解析を行った。DNA TE あるいは Retro TE は、それぞれ DNA トランスポゾンあるいはレトロトランスポゾンを意味する。

図 11. ツメガエル異質四倍体の異種交配を介した S サブゲノムに偏った DNA 欠失の分子機構モデルの提案

約 3,400 万年前に種分岐した二倍体の L あるいは S サブゲノムをもった 2 種は、生殖隔離により、それぞれ独自の「トランスポゾン-piRNA システム」を構築した。約 1,800 から 1,700 万年前に異種として出会い、交配し、異質四倍体化した個体および、その後の集団が誕生した。異質四倍体化個体の生殖系列細胞は、異種ゲノムが混在し、トランスポゾンの抑制システムの変調から、特に DNA トランスポゾンが脱抑制され、活動に転じた。これにより、二重鎖切断が頻発し、すぐに修復できない場合は、DNA 欠失が生じた（CRISP-cas9 系のように）。DNA トランスポゾンの活性化は L/S サブゲノム両方で起きたが、S サブゲノムゲノムに存在する DNA トランスポゾンの活動がより頻度が高かったため、S サブゲノムに偏った DNA 欠失が起こった。

図 12. 雑種ヨーロッパトノサマガエルにおける生殖系列細胞における片親種コガタガエルのゲノム全欠失の分子機構モデルの提案

トノサマガエル (*Peleophylax*) 属のワライガエル (*P. ridibundus*) とコガタガエル (*P. lessonae*) の交配雑種のヨーロッパトノサマガエル (*P. esculentus*) の生殖系列細胞では、ワライガエルとコガタガエルの異種ゲノム混在状態において、それぞれの種で独自の進化した「トランスポゾン-piRNA システム」の不和合性が起る。不和合性の分子機構は、今後の課題であるが、ワライガエルよりもコガタガエルの「トランスポゾン-piRNA システム」がより変調を起こし、多くの piRNA に発現の促進や抑制が起る。コガタガエルのゲノムに存在するいくつかの DNA トランスポゾンに対する piRNA の発現抑制により、そのトランスポゾンの脱抑制が起こり、活動が活発化する。これにより、二重鎖切断が頻発し、コガタガエルゲノムで特異的な DNA 欠失、およびそれに引き続く染色体欠失に至る。

