

マウスを用いた経時的エンハンサー RNA 解析による卵巣体細胞特異的エンハンサー候補領域および疾患感受性領域の同定

中川 竜一, 鹿島田健一

東京医科歯科大学発生発達病態学

緒言

卵巣の遺伝学的分化制御メカニズムは、精巣と比較して未知な部分が多く、46, XY 性分化疾患 (differences of sex development; DSD) と比較し、46, XX DSD では遺伝学的原因が不明なものが多い¹⁾。また早発卵巣機能不全 (以下 POI) の疾患に関わる多型の多くはエンハンサーなどの非翻訳領域に存在するとされる²⁾。卵巣特異的遺伝子のエンハンサー同定は卵巣分化に関わる疾患の病態解明に有用と考えられる。一方、エンハンサーは標的遺伝子から遠く離れた位置にあることが多く³⁾、その同定は容易ではない。

以前よりエンハンサーの位置を推定する方法として DNase I hyper sensitivity 領域に注目する DNase-seq 法⁴⁾や、遺伝子とエンハンサーの空間的近接性に注目する chromosome conformation capture 法^{5,6)}等が用いられ、多くの成果を挙げてきた。一方、同定した領域とそのエンハンサー活性が必ずしも一致しない、解析に多量のサンプル(細胞数)を必要とする、などの制約がある。そのため既存の手法では、胎生期の臓器発生に関わるエンハンサーなど、解析のために採取できる細胞数が限られている場合、特に標的遺伝子から離れた場所に位置するエンハンサーを効率的に同定することは困難であった。

近年、活性化したエンハンサーから双方向に転写される enhancer RNA (以下 eRNA) が注目されるようになった⁷⁻⁹⁾。この双方向に転写される eRNA の転写量は、エンハンサーの活性と相関する可能性が示唆されている¹⁰⁾。われわれは、この eRNA と標的遺伝子の発現量を定量的・経時的に比較し、エンハンサーを同定することを試みた。この手法であれば、標的遺伝子から離れた位置に存在し、胎生期卵巣発生など、採取細胞数が限られた臓器において重要な役割を果たす遺伝子のエンハン

サーも同定可能であると考えたためである。

eRNA を解析するにあたり、本研究では cap analysis of gene expression (以下 CAGE) 法を用いた。CAGE 法は理化学研究所が開発したトランスクリプトーム解析手法で、メッセンジャー RNA (以下 mRNA) と同様に eRNA にも存在する 5' cap をキャプチャーしてシーケンスを行う手法である。このような原理から正確な転写開始点を決定でき、ひいては双方向に転写される eRNA を効率的に同定でき⁸⁾、eRNA の定量的・経時的な解析に有用と考えた。

目的

胎生期の卵巣分化に重要な既知の 5 遺伝子 (Wnt4, Rspo1, Foxl2, Bmp2, Fst) の mRNA および周辺領域の eRNA 発現プロファイルを、経時的、網羅的に比較し、新規エンハンサー候補領域を同定する。

方法

性腺支持細胞の回収には、Wt1 の周辺領域 150 kb、および Wt1 コード領域を mCherry-2A-EGFP に置き換えた配列をもつ Wt1-mCherry-2A-EGFP (以下 PBWt1-RG)¹¹⁾を用いた。このマウスでは Wt1 を発現する性腺支持細胞特異的に EGFP, mCherry を発現する。卵巣分化に関わる遺伝子の発現変化を捉えるために E13.5, E16.5, P0 の 3 ステージを選択し、各ステージにおいて PBWt1-RG マウスから性腺を摘出し、fluorescence-activated cell sorting を用いて EGFP/mCherry 陽性細胞を分取した。分取した細胞から total RNA を抽出し、1 ライブラリあたり 100 ng の RNA サンプルを用いて cDNA ライブラリを作成した。従来の CAGE 法では、1-2ug の RNA を必要とし、胎生期性腺支持細胞から得られる RNA サンプルで解析可能な量を直接回収することは困難であるため、PCR によって cDNA を増幅し、ステージ毎に、n = 3 でライブラリを作成した。またコントロールとして、E13.5 の精巣からも同様に支持細胞を回収し、

連絡先: 鹿島田健一 東京医科歯科大学発生発達病態学
〒113-8510 東京都文京区湯島 1 丁目 5-4
TEL: 03-5803-5249
FAX: 03-5803-5246
E-mail: kkashimada.ped@tmd.ac.jp

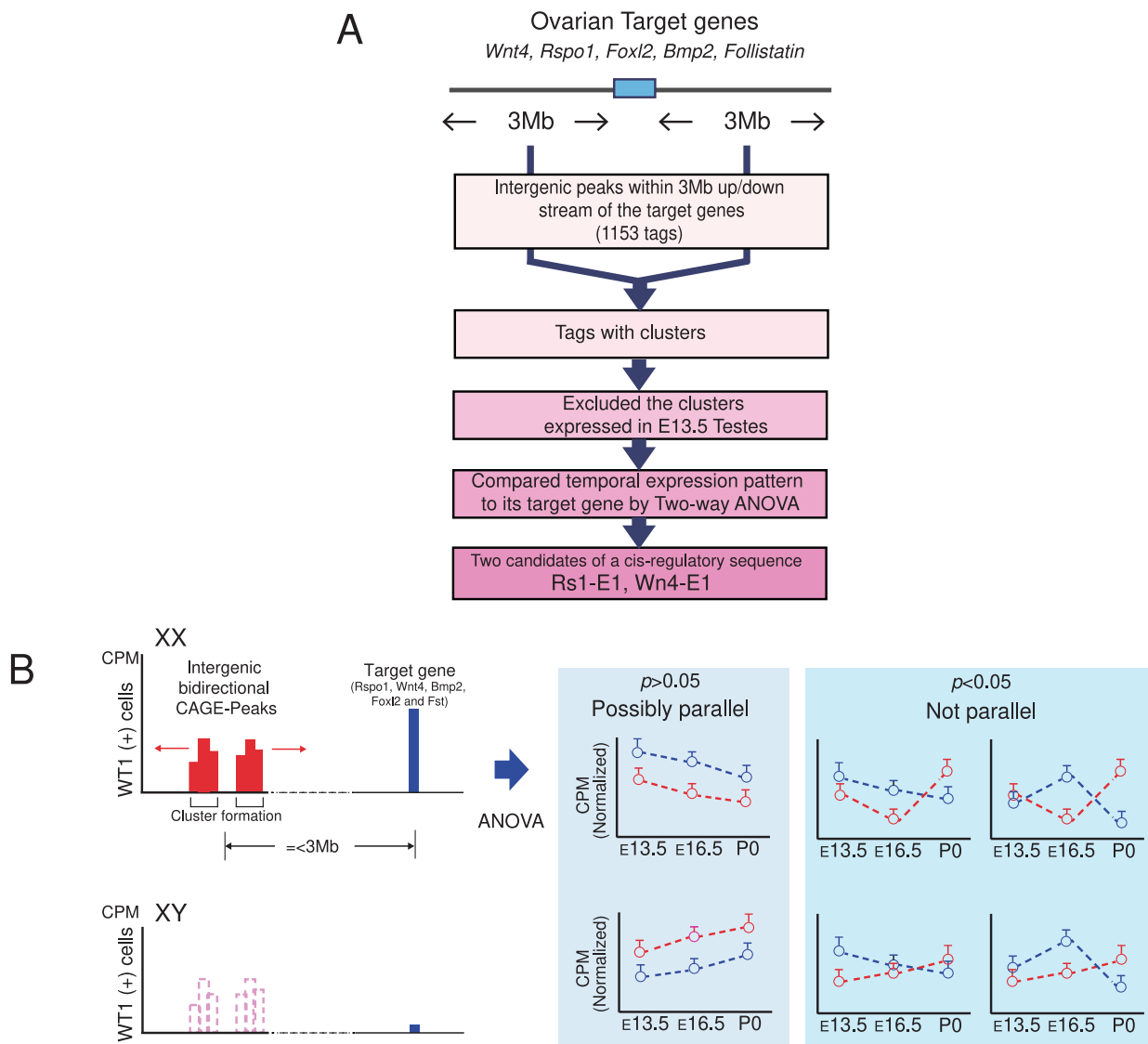


図1 エンハンサー候補領域の抽出 (Nakagawa et al HMG 2022, Figure 2を一部改変)

(A) 経時的・定量的にエンハンサー候補領域を抽出するためのフローチャート。

(B) Two-way ANOVA による検定の概要。eRNA の発現量/標的遺伝子の発現量を従属変数、各ステージを独立変数として設定した。eRNA の発現量/標的遺伝子の発現量は、それぞれの全ステージにおける発現量の平均で標準化した。CPM: count per million.

CAGE 法による解析を行った。

次に卵巣の分化に重要な 5 遺伝子 (*Wnt4, Rspo1, Foxl2, Bmp2, Fst*) のエンハンサーを同定するため、CAGE 法でプロファイルした eRNA と該当遺伝子 mRNA の経時的発現量変化を解析した。エンハンサーから標的遺伝子までの距離は 1Mb 程度以内が多く¹²⁾、余裕をもって、5 遺伝子の上下流 3 Mb に存在する CAGE tag に注目した。続いて、転写開始点の特徴であるクラスターを形成し、かつ卵巣特異的に発現をしている (=E13.5 の精巣では発現を認めない) のものを選択した。このうち、経時的発現量変化が標的遺伝子 mRNA と相似する tag の領域を、Two-way ANOVA を用いて検定、抽出した (図 1

A)。Two-way ANOVA では、帰無仮説を「エンハンサーの経時的発現変化が標的遺伝子と相似する」とし、従属変数を eRNA/標的遺伝子の発現量、独立変数を各ステージとして設定した。Two-way ANOVA で $p < 0.05$ となる tag は、標的遺伝子との発現量の比がステージ毎に変化する、すなわち eRNA の発現が標的遺伝子とは相似しないと判断されるため除外した。一方、 $p \geq 0.05$ の tag は標的遺伝子との発現量の比がステージ間で変化しない、すなわち eRNA の発現が標的遺伝子と相似する可能性があるとして判断し、エンハンサー候補として選択した (図 1 B)。

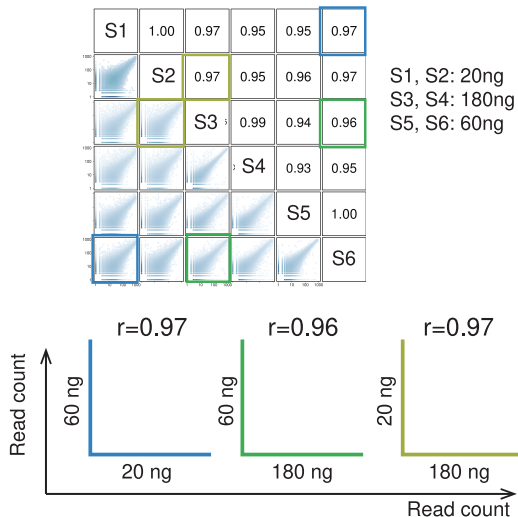


図2 PCRを併用したCAGEライブラリの定量性 (Nakagawa et al HMG 2022, Figure 1を一部改変)
 RNA input=20ng, 60ng, 180ngでcDNAライブラリに対しそれぞれ異なるサイクルのPCRによる増幅を行い、得られたライブラリ間の発現プロファイルと比較した。上段は各ライブラリ間の相関係数を表で表している。下段は、各サンプル間の発現プロファイルの一部を拡大したもの。

結果

PCRを併用したCAGEでは定量性が担保されていた

CAGE法のライブラリ作成において、PCRによる増幅の有用性を確認するため、20ng/60ng/180ngのRNAからcDNAを作成し、それぞれ11/8/6サイクル数でPCRによる増幅を行い、得られたライブラリ間の発現プロファイルの相関を検証した。結果、ライブラリ間で高い相関 ($r=0.96\sim0.97$) を示し、PCRで増幅したCAGE法でも発現プロファイルが変化しないことを確認した(図2)。

Rspo1とWnt4のエンハンサー候補領域を同定した

eRNA/標的遺伝子 mRNAの発現量変化を経時的に解析し、最終的にRspo1の上流1 Mb(mm9 chr4: 123667210-123669170, 以下Rsl-E1)およびWnt4上流100 kb(mm9 chr4: 136712850-136713270, 以下Wn4-E1)にそれぞれエンハンサー候補領域を同定した(図3)。これらの領域は assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) でいずれも open chromatin 構造をもつことを確認し、エンハンサー領域として矛盾しない結果であった。一方、Rsl-E1/Wn4-E1の周辺に存在するRspo1/Wnt4以外の遺伝子がRsl

-E1/Wn4-E1の標的となっている可能性を除外するため、Rsl-E1/Wn4-E1の周辺3Mbに存在する遺伝子のmRNA発現量が、Rsl-E1/Wn4-E1の発現量と相似する可能性の有無を検討した。結果、Rsl-E1/Wn4-E1周辺3MbにはRspo1やWnt4以外に該当領域と発現量変化が相似する遺伝子は存在しなかった。このことは、Rspo1/Wnt4がRsl-E1/Wn4-E1による調節を受けている可能性が高いことを示唆した。

Rsl-E1, Wn4-E1はIn vitroで活性を示した

Rsl-E1, Wn4-E1をluciferase vectorにクローニングし、KK-1細胞(マウス顆粒膜細胞由来)にLipofection法を用いて導入し転写活性を評価した。結果、いずれも有意な転写活性を示し(図3)、2つの候補領域のエンハンサー活性をin vitroで証明した。

エンハンサー候補領域周辺にヒトPOI患者で有意なバリエントを同定した

同定されたエンハンサー候補領域はヒトとマウス間で比較的高い保存性を認めた(図4)。そこで、同定したエンハンサー候補領域が臨床的にもつ意義を検討した。卵巣分化が障害されていると考えられる日本人女性の患者(46, XX DSD 5例, 性腺異形成3例, POI 16例, 計24例)について、血液から抽出されたgenomic DNAを用い、ヒトにおけるエンハンサー候補領域(以下RS1-E1, WN4-E1)周辺のcopy number variation(以下CNV)およびsingle nucleotide variant(以下SNV)を検証した。結果、CNVは認められなかったが、POI患者において稀なSNVをRS1-E1内(NC_000001.10:g.39249689C>G(rs975488659), Minor Allele Frequency 0.0208)とWN4-E1周辺(NC_000001.10:g.22615333G>A(rs141561868), Minor Allele Frequency 0.0208)に1つずつ同定した。日本人健常人のバリエントデータベースであるToMMo(Tohoku Medical Megabank Organization database)を参照して関連解析を行い算出したOdds Ratioはそれぞれ27.15(3.48-211.73), 353.17(21.77-5730)と有意であった。

考察

本研究の特筆すべき点は主に以下の3つである。

第一に、PCRを併用することで、少量のRNAを用いたCAGE法の開発に成功したこと、第二に、eRNAを

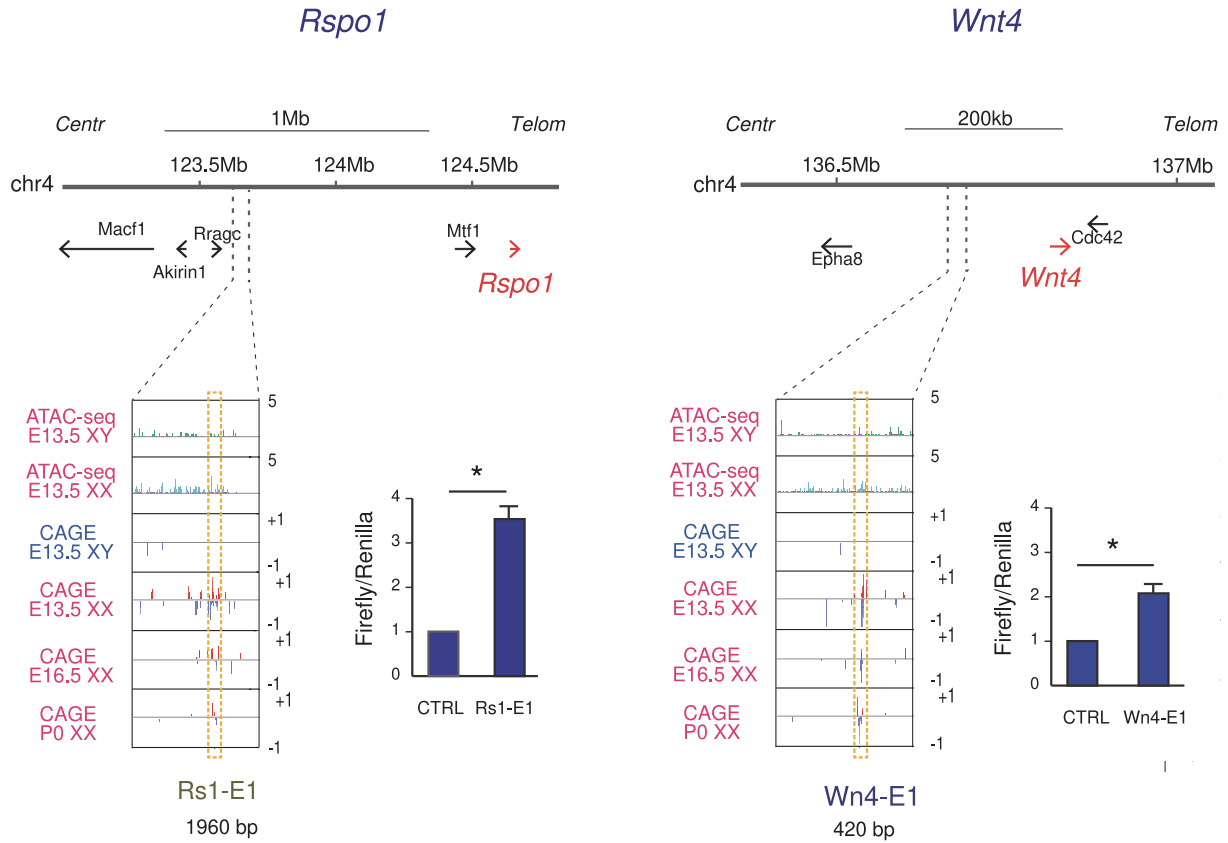


図3 同定されたエンハンサー候補領域の特徴 (Nakagawa et al HMG 2022, Figure 3を一部改変)

E13, 5/E16. 5/P0卵巣支持細胞の CAGE データ, E13. 5精巣支持細胞の CAGE データおよび CAGE と同様の方法で分取した性腺支持細胞を使った ATAC-seq のデータ (E13. 5精巣支持細胞, E13. 5卵巣支持細胞) をゲノムブラウザ上に表示している。縦軸は CPM (count per million) を示し, CAGE では eRNA が転写される方向が 5' → 3' の場合に + (正), 3' → 5' の場合に - (負) とした。Rs1-E1 と Wn4-E1 の in vitro における転写活性の評価は Renilla を内因性コントロールとし, dual luciferase assay system で解析した。CTRL は Firefly luciferase ベクター単体。Mean ± SEM, n = 3, * : p < 0.05.

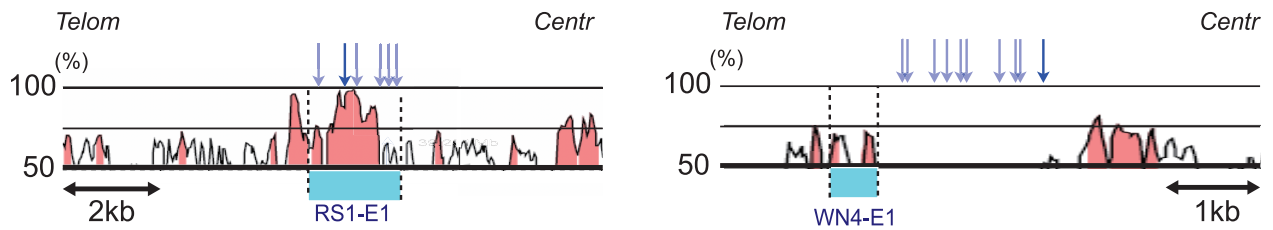


図4 エンハンサー候補領域のマウス/ヒト間における保存性 (Nakagawa et al HMG 2022, Figure 4を一部改変)

マウスとヒトにおけるエンハンサー候補領域の配列の保存性を VISTA (visualization tool for alignments) を用いて表示した。赤で示された部分は, 保存性が70%を超える領域を示している。青矢印は患者で同定されたバリエントを示しており, このうち濃い青の矢印は関連解析で Odds ratio が有意に高かったバリエントを意味する。

定量することで, エンハンサー活性をもつエンハンサー候補領域を同定することができる新たな手法を見出したこと, 第三に, ヒトにおける本法で同定された候補領域から, ヒトにおける卵巣発生障害をもつ患者からまれな多型を同定したことである。

CAGE の原法は解析に必要な RNA 量が多く大量の細胞を必要とする技術的な限界があるため⁸⁾, 過去にもいくつかの方法が試みられてきた^{13, 14)}。本研究では, ライ

ブラリ作成前に PCR により増幅するという比較的簡便な方法を試みたが, この方法でも, 特に遺伝子プロファイルに大きな変化を与えることなく増幅できることが明らかとなった。これにより, 少ない RNA を試料とする解析, とくに胎仔マウスの臓器や特定の細胞など in vivo における希少な細胞の解析実験への応用可能と考えられる。

また, 本研究で最も大きな問題となったのは, 同定し

たエンハンサー候補領域と候補となる対象遺伝子との関連づけである。エンハンサーは標的遺伝子から遠く離れた位置に存在することが多く、実際に候補となるエンハンサーが対象遺伝子の制御に関わることを証明することは容易ではない。本研究ではeRNAの定量性に着目し^{7,8)}、絞り込んだ卵巣特異的なエンハンサー候補領域に対して、さらにエンハンサー候補領域周辺の別の遺伝子が逆にエンハンサー候補領域のターゲットとなる可能性を除外することで、同定したエンハンサー候補領域と候補となる対象遺伝子との関連づけを行った。今後はより直接的な手法を用いた解析法の開発、例えば少量の試料で可能なHiC法などが望まれる。

最後に、本来 genome wide association study (以下 GWAS) のような統計的手法で疾患と関連しうるバリエントを同定する場合、多くの患者数を含むコホートが必要となる。その点、POIに対してGWASで解析を行った既報¹⁵⁻¹⁸⁾と比較すると、本研究はより少ない症例数から有意なバリエントを同定できた。マウスを用いた機能解析結果をもとにヒトのエンハンサーを同定する手法は、以前より用いられているが、そうした手法ではより簡便にかつ網羅的にマウスのサンプルを用いる技術が望まれる。本研究により開発した手法はそうした点を満たすと考えられ、今後さらなる疾患に関連するバリエントの探索に有用であると考えられた。

今後、本研究で示したアプローチを発展させることで、卵巣以外の組織/臓器においてもエンハンサーを効率的にとらえられるようになる可能性がある。これにより将来的には、様々な組織/臓器の発分化のメカニズムが解明され、ひいては多くの疾患の病態解明につながる事が期待される。(なお本研究は、Human Molecular Genetics (2022) 3に掲載された。)

引用文献

1. Knarston I, Ayers K, Sinclair A (2016) Molecular mechanisms associated with 46, XX disorders of sex development. *Clin Sci (Lond)*, 130, 421-432.
2. Rossetti R, Ferrari I, Bonomi M, Persani L (2017) Genetics of primary ovarian insufficiency. *Clin Genet* 91, 183-198.
3. Chandra V, Bhattacharyya S, Schmiedel BJ, Madrigal A, Gonzalez-Colin C, Fotsing S, Crinklaw A, Seumo G, Mohammadi P, Kronenberg M, Peters B, Ay F, Vijayanand P (2021) Promoter-interacting expression quantitative trait loci are enriched for functional genetic variants. *Nat. Genet* 53, 110-119.
4. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, Giste E, Johnson AK, Johnson EM, Kutuyavin T, Lajoie B, Lee BK, Lee K, London D, Lotakis D, Neph S, Neri F, Nguyen ED, Qu H, Reynolds AP, Roach V, Safi A, Sanchez ME, Sanyal A, Shafer A, Simon JM, Song L, Vong S, Weaver M, Yan Y, Zhang Z, Zhang Z, Lenhard B, Tewari M, Dorschner MO, Hansen RS, Navas PA, Stamatoyannopoulos G, Iyer VR, Lieb JD, Sunyaev SR, Akey JM, Sabo PJ, Kaul R, Furey TS, Dekker J, Crawford GE, Stamatoyannopoulos JA (2012) The accessible chromatin landscape of the human genome. *Nature* 489, 75-82.
5. Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. *Science* 295, 1306-1311.
6. Sanyal A, Lajoie BR, Jain G, Dekker J (2012) The long-range interaction landscape of gene promoters. *Nature* 489, 109-113.
7. Melgar MF, Collins FS, Sethupathy P (2011) Discovery of active enhancers through bidirectional expression of short transcripts. *Genome Biol* 12, R113.
8. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfischer L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, Jørgensen M, Andersen PR, Bertin N, Rackham O, Burroughs AM, Bailie JK, Ishizu Y, Shimizu Y, Furuhashi E, Maeda S, Negishi Y, Mungall CJ, Meehan TF, Lassmann T, Itoh M, Kawaji H, Kondo N, Kawai J, Lennartsson A, Daub CO, Heutink P, Hume DA, Jensen TH, Suzuki H, Hayashizaki Y, Müller F, Forrest ARR, Carninci P, Rehli M, Sandelin A (2014) An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455-461.
9. Arnold PR, Wells AD, Li XC (2019) Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. *Front Cell Dev Biol* 7, 377.
10. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182-187.
11. Zhao L, Ng ET, Koopman P (2014) A piggyBac transposon- and gateway-enhanced system for efficient BAC transgenesis. *Dev Dyn* 243, 1086-1094.
12. Lettice LA, Heaney SJH, Purdie LA, Li L, de Beer P, Oostra BA, Goode D, Elgar G, Hill RE, de Graaff E (2003) A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* 12, 1725-1735.
13. Cvetesic N, Leitch HG, Borkowska M, Müller F, Carninci P, Hajkova P, Lenhard B (2018) SLIC-CAGE: high-resolution transcription start site mapping using nanogram-levels of total RNA. *Genome Res* 28, 1943-1956.
14. Plessy C, Bertin N, Takahashi H, Simone R, Salimullah M, Lassmann T, Vitezic M, Severin J, Olivarius S, Lazarevic D, Hornig N, Orlando V, Bell I, Gao H, Dumais J, Kapranov P, Wang H, Davis CA, Gingeras TR, Kawai J, Daub CO, Hayashizaki Y, Gustincich S, Carninci P (2010) Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nat. Methods* 7,

- 528-534.
15. Kang H, Lee SK, Kim MH, Song J, Bae SJ, Kim NK, Lee SH, Kwack K (2008) Parathyroid hormone-responsive B1 gene is associated with premature ovarian failure. *Hum Reprod* 23, 1457-1465.
 16. Knauff EAH, Franke L, van Es MA, van den Berg LH, van der Schouw YT, Laven JSE, Lambalk CB, Hoek A, Goverde AJ, Christin-Maitre S, Hsueh AJ, Wijmenga C, Fauser BCJM, Consortium DPOF (2009) Genome-wide association study in premature ovarian failure patients suggests ADAMTS19 as a possible candidate gene. *Hum. Reprod* 24, 2372-2378.
 17. Pyun JA, Cha DH, Kwack K (2012) LAMC1 gene is associated with premature ovarian failure. *Maturitas* 71, 402-406.
 18. Qin Y, Zhao H, Xu J, Shi Y, Li Z, Qiao J, Liu J, Qin C, Ren C, Li J, Qin C, Ren C, Li J, Chen S, Cao Y, China POF Study Group; Joe Leigh Simpson, Chen ZJ (2012) Association of 8q22.3 locus in Chinese Han with idiopathic premature ovarian failure (POF). *Hum Mol Genet* 21, 430-436.