

## 博士論文

論文題目 歯状核赤核・淡蒼球レイ体萎縮症タンパク質  
(DRPLA protein) の転写調節標的遺伝子の探索

氏 名 波多野敬子

要旨.....	2
略語集.....	3
URL 集.....	5
序章.....	6
1. DRPLA とは.....	6
2. DRPLA 遺伝子・DRPLAp について.....	9
3. DRPLA の病態機序に関する先行研究.....	11
4. 本研究の目的.....	17
第 1 章 DRPLAp の転写 co-regulator としての標的遺伝子の RNA-seq による解析.....	18
材料と方法.....	21
1. 実験系の構築.....	21
2. RNA-seq.....	28
3. RNA-seq の解析.....	32
結果.....	38
1. 実験系の構築.....	38
2. RNA-seq.....	43
3. RNA-seq の解析の全体像.....	48
4. Q19 Dox (+) vs Dox (-) の比較.....	57
5. Q88 Dox (+) vs Dox (-) の比較.....	63
第 2 章 考察.....	72
第 3 章 結語と今後の課題.....	79
謝辞.....	80
参考文献.....	81

## 要旨

歯状核赤核・淡蒼球レイ体萎縮症（DRPLA）は、*DRPLA* 遺伝子の CAG リピートの異常伸長による常染色体優性遺伝の脊髄小脳変性症である。DRPLA タンパク質は転写 co-regulator として働くが、その標的遺伝子は未同定であった。今回 DRPLA タンパク質を定常発現する培養細胞を用いた RNA sequence により、DRPLA タンパク質の強制発現により、qRT-PCR で validate された発現変動遺伝子を 11 個、mutant DRPLA タンパク質の強制発現により、発現変動を認める遺伝子 10 個を同定した。標的遺伝子の有力な候補であり病態解明・治療法開発にあたり有用な知見と考えた。

## 略語集

bp; base pair

BSA; bovine serum albumin

CPM; count per million

CTP; cytidine triphosphate

DEGs; differentially expressed genes

D-MEM; Dulbecco's Modified Eagle Medium

Dox; doxycycline

DRPLA; Dentatorubral-Pallidoluysian Atrophy

DRPLAp; Dentatorubral-Pallidoluysian Atrophy protein

ds cDNA; double strand cDNA

FC; fold change

FPKM; Fragments Per Kilobase of exon per Million mapped fragments

FRDA; Freidreich's ataxia

FRT; Flp recombination target site

FU; arbitrary fluorescent unit

FXTAS; fragile X associated tremor/ataxia syndrome

gDNA; genomic DNA

GO; Gene Ontology

GTEEx; Genotype-Tissue Expression

HDL2; Huntington's disease-like 2

LDS; lithium dodecyl sulfate

MJD; Machado-Joseph disease

MMLV-RT; Moloney murine leukemia virus reverse transcriptase

nt; nucleotide

OMIM; Online Mendelian Inheritance in Man

PVDF; polyvinylidene difluoride

qRT-PCR; quantitative reverse transcription-PCR

RIN; RNA integrity number

RPKM; Reads Per Kilobase of transcript per Million mapped reads

RPM; read per million mapped read

rRNA; ribosomal RNA

SBMA; spinal and bulbar muscular atrophy

SCA; spinocerebellar ataxia

SDS; sodium dodecyl sulfate

TAF; TATA-binding protein-associated factor

TBS ; Tris buffered saline

Tet: tetracycline

## URL 集

OMIM; <http://www.omim.org>

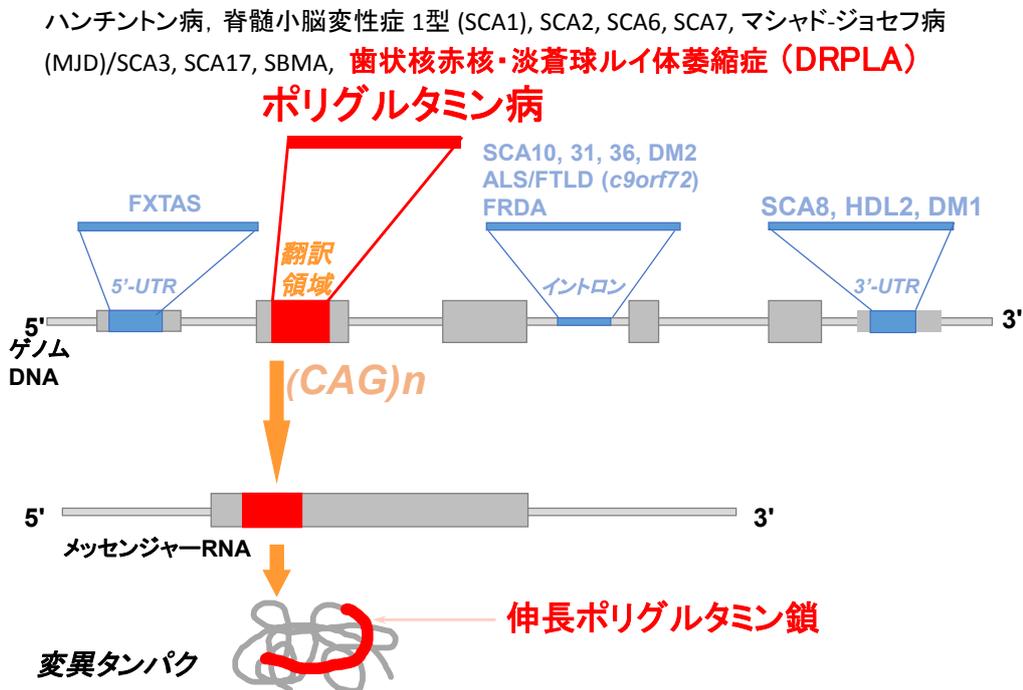
GTEEx; <http://www.gtexportal.org/home/>

## 序章

### 1. DRPLA とは

DRPLA (Dentatorubral-Pallidoluysian Atrophy, 齒状核赤核・淡蒼球レイ体萎縮症、OMIM (Online Mendelian Inheritance in Man) #125370) は常染色体優性遺伝性の脊髄小脳変性症であり、12p13.31 における *DRPLA* 遺伝子 (*ATN1*) exon5 の CAG リピート数の異常伸長を病因とする [1,2]。CAG リピートは DRPLAp (atrophin 1) のグルタミン鎖をコードしており (図 1)、正常では 3~36 回であるところ、DRPLA 患者においては 49~88 回と伸長を認める。CAG リピートの異常伸長は伸長ポリグルタミン鎖をコードすることから、DRPLA は spinal and bulbar muscular atrophy (SBMA) [3]、ハンチントン病 [4]、脊髄小脳変性症 1 型 (spinocerebellar ataxia type 1: SCA1) [5]、マシャド・ジョセフ病 (Machado-Joseph disease, MJD/SCA3) [6]、SCA2 [7-9]、SCA6 [10]、SCA7 [11]、SCA17 [12, 13] と共にポリグルタミン病に分類される。

図 1. リピート配列の異常伸長による神経変性疾患



DRPLA は翻訳領域における CAG リピート伸長によるポリグルタミン病である。

FXTAS; fragile X associated tremor/ataxia syndrome

DM; myotonic dystrophy

ALS/FTLD; amyotrophic lateral sclerosis/frontotemporal lobar degeneration

FRDA; Friedreich's ataxia

HDL2; Huntington's disease-like 2

DRPLA の発症の平均年齢は 47.3 歳で、世代を経る毎に発症年齢が早まるという特徴がある。臨床症状は、若年発症型（20 歳未満）では主に進行性ミオクローヌスてんかんの病像を呈することが多く、遅発成人発症型（40 歳以上）は小脳性運動失調に加えて舞踏運動などの不随意運動が主体である。若年成人型（20 歳以上 40 歳未満）は遅発成人発症型と若年成人型の間中型を示す [14]。

1994年にKoideら、Nagafuchiらによって、*DRPLA* 遺伝子のCAGリピート伸長がDRPLAの原因であることが報告され、CAGリピート長は発症年齢と逆相関を呈し、世代とともにCAGリピート長が伸長する現象が確かめられ、特に父系遺伝において顕著であることが判明した<sup>[1, 2]</sup>。つまり、DRPLAにおいて、世代を経る毎に発症年齢が早まる表現促進現象は、このリピート不安定性によるものであることが明らかとなっている

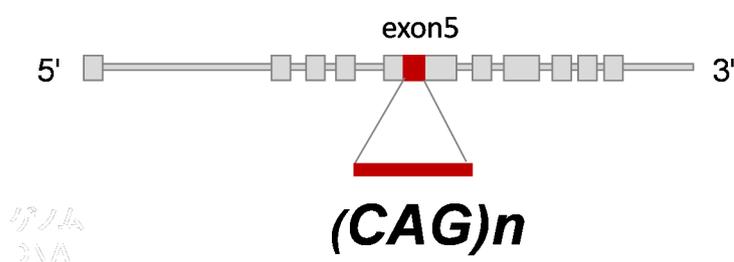
DRPLAは本邦において頻度の高い疾患であり、常染色体優性脊髄小脳変性症の中でMJD、SCA6、SCA31に次ぎ4番目に多い。実際、当教室において遺伝子診断を施行した507家系の常染色体優性脊髄小脳変性症の12.6%で診断されている。2008年の報告<sup>[15]</sup>より本邦では0.48人/100,000人の有病率と見積もられ、有病者は約600人と推定される。

アメリカでHaw River syndromeと呼ばれていた疾患<sup>[16]</sup>が、*DRPLA* 遺伝子の同定後に、*DRPLA* 遺伝子のCAGリピート伸長が認められ同一疾患であることが判明したという例<sup>[17]</sup>、イギリスにおける集積地の発見<sup>[18, 19]</sup>については有名であるが、基本的には世界的に非常に希な疾患である<sup>[20]</sup>。アメリカ以外には、マルタ島<sup>[18]</sup>、韓国<sup>[21, 22]</sup>、ポルトガル<sup>[23, 24]</sup>、デンマーク<sup>[25]</sup>、イタリア<sup>[26-28]</sup>、マケドニア<sup>[29]</sup>、ベネズエラ<sup>[30]</sup>、ハンガリー<sup>[31]</sup>などで少数例が報告されるにとどまり、一方中国<sup>[32]</sup>、タイ<sup>[33]</sup>、ブラジル<sup>[34]</sup>、フランス<sup>[35]</sup>からの報告ではDRPLAが見いだせなかったとされ、DRPLAは本邦において特異的に頻度が高い疾患であると言える。

## 2. DRPLA 遺伝子・DRPLAp について

DRPLA 遺伝子は約 20 kb で 10 個の exon からなる。DRPLA 遺伝子がコードする DRPLA protein (DRPLAp) は、1185 アミノ酸から成り、488 番目のアミノ酸残基からポリグルタミン鎖が存在する (図 2)。

図 2. DRPLA 遺伝子の模式図



DRPLAp は細胞内で、核と細胞質に存在することが知られている [36-40]。DRPLAp には、ポリグルタミン以外にも、serine-rich region, polyproline tract など反復配列を多く含むことが知られている [40]。また、N 末端側の 16-32 番目のアミノ酸と、C 末端側の 1033-1041 番目のアミノ酸からなる領域が、それぞれ nuclear localization signal (NLS) と nuclear export signal (NES) であると予測されており、Nucifora らは、Neuro2a を用いた検討で、これらが実際に機能的であることを示している [40]。

Onodera らは northern blot による分析で、wild type DRPLA mRNA、mutant DRPLA mRNA 共に、脳その他、心臓・肺・腎臓・胎盤・骨格筋・肝臓にも分布し、脳に

おいては、齒状核赤核・淡蒼球ルイ体系のみでなく、扁桃体、尾状核、脳梁、海馬、視床下部、黒質、視床下核、視床に幅広く分布すると報告している [41]。

### 3. DRPLA の病態機序に関する先行研究

DRPLA の病態機序の検討にあたり、本来、DRPLA<sub>p</sub> の細胞内の局在、processing、生理機能などを十分に明らかにすることが、分子病態機序研究、治療法開発でも必要である。ここでは、DRPLA<sub>p</sub> の機能に関する先行研究を述べる。

DRPLA の剖検脳における病理所見は、小脳歯状核遠心系、淡蒼球レイ体系の系統変性が主体であり、同部位の神経細胞脱落を呈する。歯状核赤核・淡蒼球レイ体において mutant DRPLA<sub>p</sub> を含む核内封入体が確認されている [42]。Sato らは全長 mutant DRPLA 遺伝子のトランスジェニックマウスを報告し Q129 マウスにおいては運動失調とてんかん、ミオクロヌスを呈し、DRPLA と類似した病態を示していることを明らかにした [43]。Q129 マウスにおいては、進行性の脳萎縮を認めるものの、明らかな神経細胞の消失は認めず、spine の数の減少・萎縮性変化、神経細胞の細胞体・神経突起の萎縮を認めた。神経細胞の核内には封入体形成が認められた。一方、DRPLA のノックアウトマウスでは、体重は野生型よりも小さいが、明らかな運動失調や痙攣などは認めず、脳の病理解学的検索においても明らかな神経細胞の脱落や変性は認めていない [44]。剖検脳、トランスジェニックマウスにおいて DRPLA<sub>p</sub> を含む核内封入体が認められること、トランスジェニックマウスにおいてヒトと類似した症状を認めたことから mutant DRPLA<sub>p</sub> が核において何らかの機能障害を呈する gain-of-toxic function 機序が病態に関与していることが推測されてきた。

培養細胞を用いて細胞死と封入体形成について検討した研究では、全長 DRPLA<sub>p</sub> を、

神経細胞に分化させた PC12 細胞・fibroblast<sup>[39]</sup>、Neuro2a 細胞<sup>[40]</sup> に発現させた系においては、正常ポリグルタミン鎖長の DRPLAp・伸長ポリグルタミン鎖長の DRPLAp の両方で、明らかな細胞死は誘導されなかった。凝集体については、正常ポリグルタミン鎖長または伸長ポリグルタミン鎖長の DRPLAp を発現させた、Neuro2a 細胞では観察されなかったが<sup>[40]</sup>、truncated DRPLA (Q82) 遺伝子 (伸長ポリグルタミン鎖 (Q82)) を発現させた際に、神経細胞に分化させた PC12 細胞では核に、fibroblast においては細胞質に凝集体が観察されている<sup>[39]</sup>。Sato らは、伸長ポリグルタミン鎖を有する truncated DRPLAp を発現させることで、細胞毒性が出現し、凝集体が多く観察されることを示している<sup>[39]</sup>。また Nucifora らも同様の結果を得ており、DRPLAp が細胞内でプロセッシングを受ける過程で、C 末端の NES が消失することで核への局在が増えることが細胞毒性の原因ではないかと推測している<sup>[40]</sup>。

DRPLAp の核における機能については、転写 co-regulator であることが複数の研究で示されている。*Drosophila* において DRPLA 遺伝子の相同体 Atrophin (Atro) は、Even-skipped (Eve) タンパク質と直接的に結合し、Atro が Eve の co-repressor であることが示されている<sup>[45]</sup>。他にも wild type mouse DRPLAp の GAL4 luciferase reporter assay を用いた報告では、HEK293 細胞では転写活性化<sup>[46]</sup>、Neuro2a 細胞においては用量依存性に転写抑制を行うこと<sup>[47]</sup> が示されている。

Shimohata T. らは、yeast two hybrid assay により、TAF II 130 (TATA-binding protein-associated factor) と polyQ 鎖 (Q19 or Q82) の結合を見出した<sup>[48]</sup>。COS-7 細

胞においても、polyQ 鎖 (Q57) が TAF II 130 と結合することを、polyQ 鎖の tag に対する抗体を用いた co-precipitation により示し、細胞内の共局在を、polyQ 鎖に結合した GFP の蛍光と TAF II 130 に対する免疫蛍光染色の overlay による観察により確認した [48]。DRPLA 剖検脳神経細胞核内封入体においても、wild type DRPLAp または mutant DRPLAp と CREB・TAF II 130 が共局在し、wild type DRPLAp と mutant DRPLAp の両者と結合していることが確認された [48]。TAF II 130 は CREB (cAMP response element binding protein) 依存性転写調節機構において転写共役因子として機能するタンパク質である [49, 50] ことから、DRPLAp が TAF II 130 を介して、CREB 依存性転写調節を行うとする仮説が立てられた (図 3)。実際、in vitro reporter assay を用いて伸長 polyQ 鎖のみを発現させたところ、luciferase の転写が抑制されるが、伸長 polyQ 鎖と TAF II 130 を同時に発現させると転写活性については rescue されることが示された。このことより、伸長 polyQ 鎖が、CREB 依存性転写活性化を阻害する可能性が示された。また、Shimohata M. らは、DRPLAp (Q0, Q19, Q57) を発現させた Neuro2a 細胞において、リン酸化 CREB 依存性の *c-fos* の転写が polyQ 鎖長依存性に抑制されると報告している [47]。これらの研究から、DRPLAp は少なくとも CREB に関連した系において転写の co-regulator として働いていると考えられた。

DRPLA における遺伝子発現プロファイル解析として、マウスモデルを用いた 2 つの研究がある。赤岩らは、DRPLA 遺伝子ノックアウトマウスと wild type マウスとの間で、同じ日齢において有意に発現変動する遺伝子をマイクロアレイを用いて解析した [44]。胎

生 14 日・生後 1 日、42 日の各日齢における、ノックアウトマウスと wild type マウスの全脳から抽出した total RNA を用いることで、それぞれ 63 遺伝子、5 遺伝子、3 遺伝子に変動が見られることを示している。Suzuki らは全長の *DRPLA* 遺伝子を導入したトランスジェニックマウス的大脑・小脳から抽出した RNA を用い、polyQ 鎖数・出生後週数依存性に発現が変動する遺伝子を探索した [44]。両研究とも神経細胞を用いた貴重な知見である一方で、個体の脳組織由来の RNA を用いていることによる検体間のばらつきや、混在した細胞種の homogenate を用いているという点は考慮する必要がある。

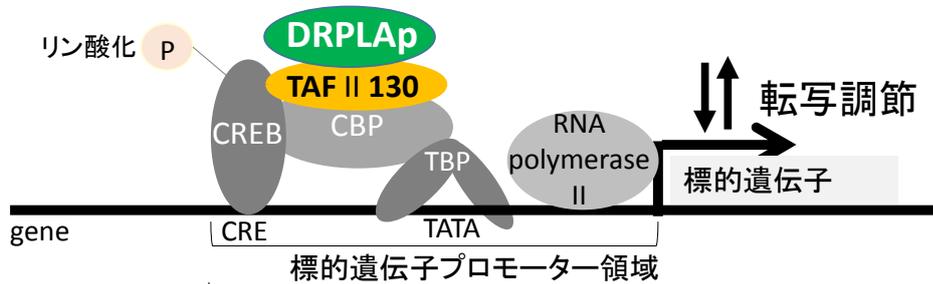
前述のように、DRPLAp の機能解析において用いられる *DRPLA* 遺伝子のコンストラクトは、全長 *DRPLA* 遺伝子ではなく truncated *DRPLA* 遺伝子や polyQ 鎖のみであることが多い [39, 40, 46, 48, 51]。Truncated DRPLAp の方が polyQ 鎖伸長による毒性を観察しやすい傾向があるため研究に用いられることが多いが、DRPLAp の機能を解明するためには、全長 DRPLAp を用いた研究を行うことが望ましいと考えられる。

同じくポリグルタミン病である SBMA は、AR (androgen receptor) 内の CAG リピート伸長が原因であるが、機能獲得によると思われる前角細胞におけるポリグルタミンの凝集が認められる一方、アンドロゲン不応に対応する女性化乳房、睾丸萎縮、乏精子症などの症状も呈することが知られている [3, 53]。

以上のことから、*DRPLA* の病態機序を理解するためには、全長の DRPLAp について、その生理的な役割を明らかにした上で、伸長ポリグルタミン鎖がその生理的機能にどのように影響を与えるのか、さらに、gain-of-toxic function 機序と loss-of-function 機序

の両方の観点から解明することが重要であると考えられ、本疾患の治療戦略を立てる上でも重要と考えられる。

図 3. TAF II 130 の CREB 介在性転写調節機構における転写共役因子としての機能概略

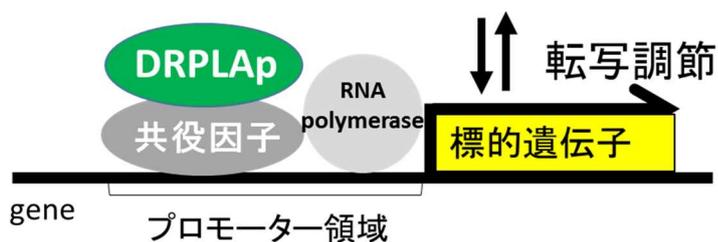


CREB 依存性転写調節機構では、標的遺伝子のプロモーター領域にある CRE (cyclic AMP responsive element) 応答配列に結合した CREB がリン酸化されると、CREB と CBP (CREB binding protein) によって TAF II 130 が活性化される。活性化された TAF II 130 により、TBP (TATA-binding protein) がプロモーター領域の TATA-box (図中 TATA) に誘導され、RNA polymerase II により標的遺伝子の転写調節がなされる。DRPLAp は TAF II 130 を介して、CREB 依存性転写調節を行うとする仮説が立てられた。

#### 4. 本研究の目的

本研究では、DRPLAp の転写の co-regulator としての機能に着目し、全長 DRPLAp をテトラサイクリンの添加による発現を誘導できる培養細胞系（HEK293 細胞）を用いて、RNA-seq 解析によりトランスクリプトーム解析を行い、wild type DRPLAp、mutant DRPLAp の発現により発現変動する遺伝子を網羅的に同定し、DRPLAp の転写の co-regulator としての生理的機能を明らかにすることを目的とした（図 4）。

図 4. 転写の co-regulator としての DRPLAp の機能



DRPLAp は共役因子と共に転写の co-regulator として働き、標的遺伝子の転写調節を行う。標的遺伝子を網羅的に探索することにより、DRPLAp の生理的機能を明らかにする。

## 第1章 DRPLAp の転写 co-regulator としての標的遺伝子の RNA-seq による解析

現在までの研究では、DRPLAp は転写の co-regulator として働くことが示唆されているため、本研究においては、DRPLAp の標的遺伝子を解明することを目的に、全長 *DRPLA* (Q19 or Q88) 遺伝子 を、テトラサイクリン依存性に発現誘導できる培養細胞を用いて RNA-seq<sup>[54]</sup> を行い、標的遺伝子を網羅的に探索することとした。

RNA-seq とは、total RNA から逆転写酵素により作成したシーケンスライブラリの塩基配列を次世代シーケンサーで解析し、塩基配列決定の後、得られた塩基配列をヒトゲノムの参照配列（リファレンスゲノム）に一致する部分に alignment（マップ）し、参照配列上の遺伝子に alignment された read の数の分布を元に、遺伝子の発現量を推定する方法である。RNA-seq を用いる利点として、既存の方法であるマイクロアレイによる解析と比較して、(a) 十分な全 read 数の確保により、より広いダイナミック・レンジ が得られ、マイクロアレイよりも少ない発現量の遺伝子の検出が可能である点、(b) 遺伝子発現量の相対的定量がより正確に可能である点、(c) 今回は直接の目的とはしていないが、新規遺伝子配列の予測、splice variant の探索が可能である点が挙げられる。

今回、TruSeq Stranded total RNA with RiboZero Gold のプロトコルを用いて RNA-seq 用のライブラリを用いた。このプロトコルを選択した理由は、total RNA の約 98% は rRNA であるが、rRNA の depletion により効率よくシーケンスを行えるようにするためである。また、このプロトコルでは、second strand cDNA 合成の際、dTTP

(deoxythymidine 5'-triphosphate) の代わりに dUTP (2'-deoxyuridine 5'-triphosphate) を組み込み、PCR 増幅前に uracil-DNA glycosylase で second strand cDNA 鎖からは増幅できないようにする。このことにより first strand cDNA 鎖と second strand cDNA 鎖とを区別してシーケンスすることが可能になり、転写の方向性についての情報が保持されるという利点がある。

本研究においては、Tet-On system (Tet; tetracycline) [55] 制御下で *GFP-full length DRPLA* 融合遺伝子を定常発現する HEK293 細胞株由来の Flp-In™ T-REx™ 293 細胞を用いた (Invitrogen® by Thermo Fisher Scientific, Inc.)。Tet-On system とは tetracycline (Tet) を添加することで細胞に目的遺伝子を発現させる系である。本研究では半減期が Tet より長い Tet の誘導体である doxycycline (Dox) を添加した (Dox (+))。Flp-In™ T-REx™ 293 細胞を用いた理由は、Flp-In™ T-REx™ 293 host cell line のゲノムには FRT (Flp recombination target site) が一箇所固定した位置に存在するため、マルチクローニングサイトに今回の目的遺伝子である *GFP-full length DRPLA* (Q19 or Q88) 融合遺伝子が挿入された pcDNA5™/FR/TO (Invitrogen® by Thermo Fisher Scientific, Inc., Waltham, MA USA) と Flp recombinase を発現する pOG44 (Invitrogen® by Thermo Fisher Scientific, Inc.) とを co-transfection することで、FRT との相同組み換えが生じ、ゲノムに一箇所固定した位置に目的とする *GFP-DRPLA* 融合遺伝子が組み込むことができ、結果として、Tet-On (Dox (+)) によりライン間で均一な発現環境で *GFP-full length DRPLA* (Q19 or Q88) 融合遺伝子を stable に発現させる

ことが可能になるからである。再現性の高い転写プロファイリングデータが得られることが期待されたため、この細胞を用いる方針とした。

## 材料と方法

### 1. 実験系の構築

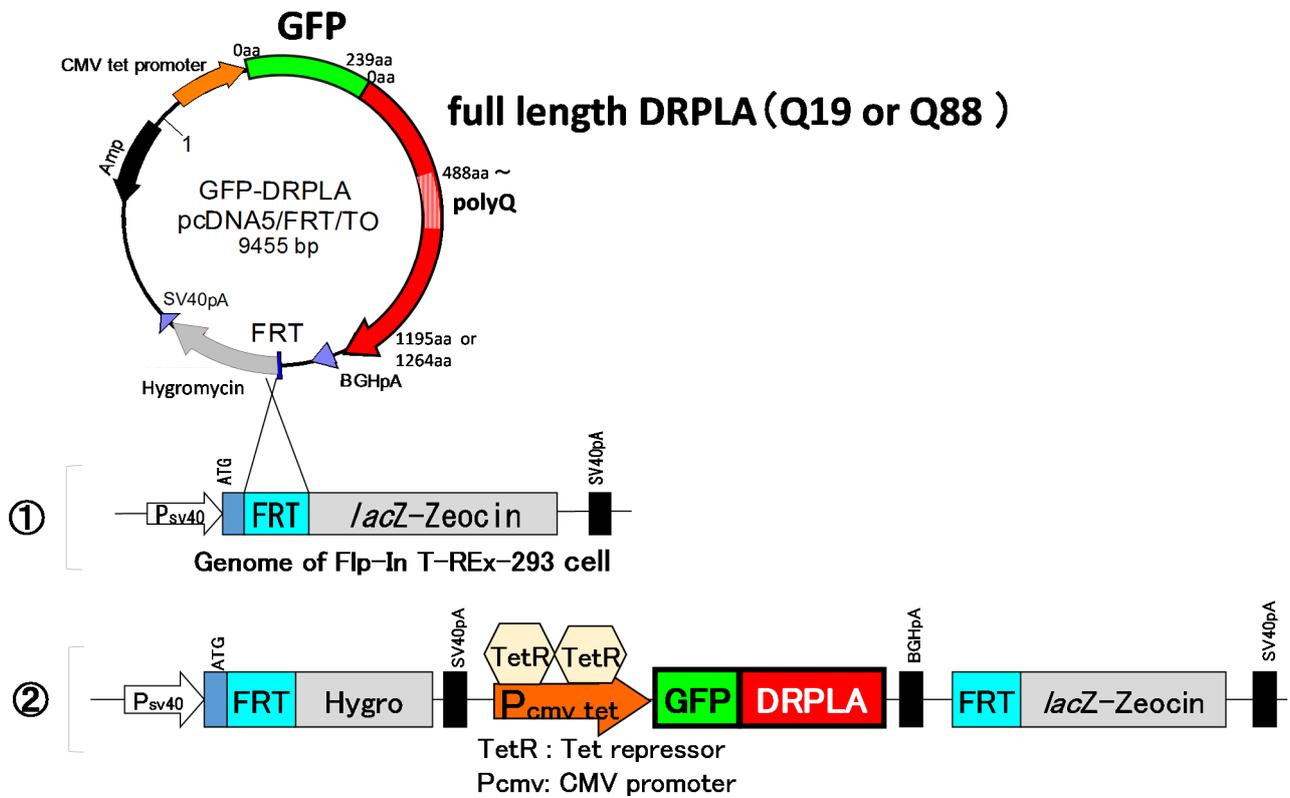
#### (1) Tet-On system 制御下 *GFP-full length DRPLA* 遺伝子定常発現株 (Flp-In™ T-REx™ 293 細胞) の作成

本研究においては、Tet-On system 制御下で *GFP-full length DRPLA* (Q19 or Q88) 融合遺伝子を定常発現する Flp-In™ T-REx™ 293 細胞を用いた (以下 Q19、Q88、共同研究者東京大学伊達英俊博士より供与) (図 5)。

具体的には、マルチクローニングサイトの BamHI と XhoI site に、今回の目的遺伝子である *GFP-full length DRPLA* (Q19 or Q88) 融合遺伝子が挿入された pcDNA5™/FR/TO (Invitrogen® by Thermo Fisher Scientific, Inc.) を準備し、Flp recombinase を発現する pOG44 とこのベクターを Flp-In™ T-REx™ 293 細胞に transfection し、hygromycin で selection を行うことで、定常発現細胞を得た。

作成された細胞の培養は、10% Fetal Bovine Serum (Gibco® by Thermo Fisher Scientific, Inc., 12483-020)、100 µg/ml Hygromycin B (Wako Pure Chemical Industries, Ltd, Osaka, Japan, 084-07681)、15 µg/ml Blasticidin (Thermo Fisher Scientific, Inc., R210-01) を含む Dulbecco's Modified Eagle Medium (D-MEM) 培地 (Wako Pure Chemical Industries, Ltd, 044-29765) を用い、37°C、5%CO<sub>2</sub> にて行った。

図 5. Tet-On system 制御下 *GFP-full length DRPLA* 遺伝子定常発現株 (Flp-In™ T-REx™ 293 細胞) の作成



Flp-In™ T-REx™ 293 細胞のゲノム (①) に、*GFP- full length DRPLA* 融合遺伝子が、一箇所固定した FRT (Flp recombination target site) に導入された状態 (②) を示す。Tet-On (Dox (+)) によりライン間で均一な発現環境で *GFP-DRPLA* 融合遺伝子を発現させることができる。

## (2) *GFP-DRPLA* 融合遺伝子の mRNA とタンパク質の発現量の測定

### 細胞培養と mRNA、タンパク質抽出

RNA-seq の対象とするタイムウィンドウの検討のため、quantitative reverse transcription-PCR (qRT-PCR) と Western blot を行った。qRT-PCR は Q19 と Q88 を用いて Tet-On 0・12・24・48・72 時間後のタイミングで行い、Western blot は Q19 を用いて Tet-On 24・48・72 時間後、Q88 を用いて Tet-On 24 時間後のタイミングで行った。細胞培養は total RNA 抽出時は 6 well plate を用いて約  $3.0 \times 10^6$  cells/well、前述と同様の 10% Fetal Bovine Serum を含む D-MEM 培地 (Wako Pure Chemical Industries, Ltd, 044-29765) 3 ml/well の条件下で行った。タンパク質抽出時は 6 cm dish を用いて約  $5.0 \times 10^6$  cells/dish、前述と同様の 10% Fetal Bovine Serum を含む D-MEM 培地 3 ml/dish の条件下で行った。発現誘導は Dox (LKT Laboratories, Inc., St. Paul, MN, USA, D5897) を  $1 \mu\text{g}/\mu\text{l}$  の stock solution を用いて  $1 \mu\text{g}/\text{培地 } 1 \text{ ml}$  の濃度で添加して行い、24 時間以上 Tet-On とする場合は 24 時間毎に培地交換と Dox 添加を行った。

### (i) *GFP-DRPLA* 融合遺伝子の mRNA 発現の qRT-PCR による定量

細胞からの total RNA 抽出は RNeasy Plus Mini kit (Qiagen, N.V., Venlo, Netherlands) により行い、Agilent RNA 6000 Nano kit (Agilent Technologies, Inc., Santa Clara, CA, USA) と Bioanalyzer (Agilent Technologies, Inc.) を用いて RNA の質の評価を行った。RNA integrity number (RIN)  $\geq 9$  (最高 10) であるサンプルのみ

を使用した。

100 ng - 500 ng の total RNA を鋳型とし、Rever Tra Ace® qRT-PCR RT Master Mix with gDNA Remover (TOYOBO CO., LTD, Osaka, Japan) の 4×DN Master Mix

(RNase inhibitor、反応 buffer、2 v/v% gDNA remover (DNase I)) 2 µl、nuclease-free water とが計 8 µl となるよう調製し、37°C 5 分間 incubate し gDNA を除去した。

次に 5×RT Master Mix II (逆転写酵素、random primer、oligo dT primer、反応バッファ、dNTPs) 2 µl を添加し計 10 µl とし、37°C 15 分間・50°C 5 分間の逆転写反応、98°C 5 分間の酵素失活反応を行い、gDNA 除去済 cDNA を生成した。

gDNA 除去済 cDNA を用い、target を *DRPLA* 遺伝子、endogenous control を *ACTB* 遺伝子として qRT-PCR を行った。Forward primer ・ reverse primer 入り配列特異的蛍光標識プローブを用い、target (Applied Biosystems® by Thermo Fisher Scientific, Inc., TaqMan® Gene Expression Assays, FAM-MGB (minor groove binder) , Hs00157312\_m1) と endogenous control (Applied Biosystems® by Thermo Fisher Scientific, Inc., VIC-MGB, 4326315E) とを各 1 µl ずつ、THUNDERBIRD® Probe qPCR Mix (TOYOBO CO., LTD, Japan) probe (Taq DNA polymerase) 10 µl、50X ROX 0.4 µl、distilled water 6.6 µl の組成で計 20 µl の反応液とした。PCR は (1) 95°C1 分間の DNA 変性、(2) 95°C15 秒間のアニーリング、(3) 65°C1 分間の伸長反応の (1) ~ (3) を 40 サイクル繰り返す、ABI PRISM 7900 HT (Applied Biosystems®

by Thermo Fisher Scientific, Inc.) を用いて比較定量法により行った。

比較定量法においては、内在性コントロール遺伝子と目的遺伝子の増幅効率が同じで、全てのサンプルの鋳型 DNA において、内在性コントロール遺伝子が一定の割合で発現し、PCR n cycle 後の DNA 量 = (初期鋳型量 I)  $\times 2^n$  が成り立つと仮定する。Ct とは一定の PCR 産物量 (DNA 量) に達するために必要な threshold cycle 数であり、以下の式から、endogenous control (EC) 遺伝子発現量で補正した target 遺伝子の Dox (+) 時と Dox (-) 時の初期鋳型量比 ( $I_{\text{target 遺伝子 (Dox (+))}} / I_{\text{target 遺伝子 (Dox (-))}}$ ) を算出した。

$$\begin{aligned} & \left( \frac{I_{\text{target 遺伝子 (Dox (+))}}}{I_{\text{EC 遺伝子 (Dox (+))}}} \right) / \left( \frac{I_{\text{target 遺伝子 (Dox (-))}}}{I_{\text{EC 遺伝子 (Dox (-))}}} \right) \\ &= 2^{-\Delta Ct (\text{Dox (+)})} / 2^{-\Delta Ct (\text{Dox (-)})} \\ &= 2^{-(\Delta Ct (\text{Dox (+)}) - \Delta Ct (\text{Dox (-)}))} \\ &= 2^{-\Delta \Delta Ct} \end{aligned}$$

一つの RNA サンプルから独立して 3 つの cDNA を作成し、一回の qRT-PCR により  $\Delta Ct$  値を 3 サンプルの平均値として算出した。一回の実験の  $\Delta Ct$  値については、標準偏差が 1 未満の時のみ、有効な結果として用いた。有効な実験の Dox (+) と Dox (-) の比較より、 $-\Delta \Delta Ct$  値を計算した。qRT-PCR は合計 9 回施行し、9 回の実験の  $-\Delta \Delta Ct$  の平均値、 $-\Delta \Delta Ct$  の標準誤差から、初期鋳型量比を算出した。

(ii) GFP-DRPLA 融合タンパク質の発現の Western blot による解析

GFP-DRPLA (Q19 or Q88) 融合タンパク質の GFP の蛍光を、Q19 は Tet-On 24・48・72 時間後、Q88 は Tet-On 24 時間後の細胞回収時に倒立蛍光顕微鏡により倍率 100 倍で観察した後、タンパク質を抽出した。

Total protein 抽出は、最初に細胞と lysis buffer (Tris HCl pH 7.6 50 mM、NaCl 150 mM、NP40 1.0%、DOC 0.5%、SDS 2%) を、 $6.0 \times 10^6$  cells/ml の濃度で 1.5 ml eppendorf tube 内において vortex により混和し、次に shearing を 1 ml シリンジと注射針 (18 G→22 G→25 G 各径 5 回ずつ) で行い、さらに sonication を SONIFIER250 (Branson Ultrasonics Corporation, Danbury, CT, USA) にて output 1 にて 10 秒間 /tube 行い、2 時間室温で静置して行った。5% 2-hydroxy-1-ethanethiol を加えた Laemmli sample buffer (Tris-HCl pH 6.8 277.8 mM, glycerol 44.4% (v/v), LDS (lithium dodecyl sulfate) 4.4%, bromophenol blue 0.02%; BioRad Laboratories, Inc. Hercules, California, USA, 1610737) と、細胞を溶解した lysis buffer 溶液 (Tris-HCl pH 7.6 50 mM, NaCl 150 mM, NP40 1.0%, DOC (sodium deoxycholate) 0.5%, SDS (sodium dodecyl sulfate) 2%) を等量混合し、80°C 1 分間で incubate の後、氷冷し、total protein 溶液とした。

10%ポリアクリルアミドゲル (ATTO Corporation, Tokyo, Japan, E-T1020L) を running buffer (Tris-HCl pH 7.4 0.83 M, glycine 6.4 M, SDS 0.12 M) 中にて、total protein 溶液を 40  $\mu$ l/well、マーカーとして Precision Plus Protein™ Standards Dual

Color (BioRad Laboratories, Inc., 161-0374) を 10  $\mu$ l/well 入れ、20 mA/ゲルで 80 分間電気泳動した。

Transfer buffer (Tris-HCl pH 7.4 0.1 M, glycine 1.9 M, methanol 20%) 内で、メタノールに 1 分間浸した PVDF (polyvinylidene difluoride) 膜 Immobilon-PS<sup>Q</sup> (Merck Millipore Corporation, Darmstadt, Germany, ISEQ00010) に 0.2 A、8 時間で転写した。PVDF 膜を TBS-T (Tris buffered saline; Tris-HCl pH 7.4 0.05 M, NaCl 0.138 M, KCl 0.0027 M, 0.05% Tween 20) に 1 回浸した後、EzBlock CAS (Tris-HCl, 牛由来カゼイン, Tween 20 (ATTO Corporation, Tokyo, Japan, AE-1477)) に PVDF 膜を 1 時間 30 分振盪しブロッキングを行った。

一次抗体反応は、別々の PVDF 膜に 1% BSA (bovine serum albumin) を用いて抗 GFP rabbit polyclonal 抗体 (MBL Co., Ltd., Aichi, Japan, GFP 598, 1: 500) と抗 $\beta$ -actin C4 mouse monoclonal 抗体 (Santa Cruz Biotechnology, Inc., Dallas, Texas, USA, sc-47778, 1:1000) を用いて 4°C overnight で行った。PVDF 膜は TBS-T に 1 回浸した後、20 分間 $\times$ 3 回洗浄し、二次抗体反応は 1% BSA を用いて anti rabbit IgG (GE Healthcare UK Ltd, Buckinghamshire, England, NA 934V, 1:5000) もしくは anti mouse IgG (Santa Cruz Biotechnology, Inc., sc-2005, 1:5000) を用いて 1 時間振盪して行った。PVDF 膜は TBS-T に 1 回浸した後、40 分間 $\times$ 3 回洗浄の後、WSE-7120S EzWest Lumi plus (ATTO Corporation) で化学発光を行い、LAS 3000 Mini (Fujifilm Corporation, Tokyo, Japan) で撮影した。

なお、positive control として、Flp-In™ T-REx™ 293 細胞に、*GFP-DRPLA* 融合遺伝子の代わりに *GFP-GR* (*glucocorticoid receptor*) 融合遺伝子が導入された GFP-GR 細胞を用いた。GFP-GR 細胞も、*GFP-DRPLA* 融合遺伝子を発現する細胞と同様の方法で、Tet-On 24・48 時間後に GFP の蛍光の確認をし、Western blot を行った。

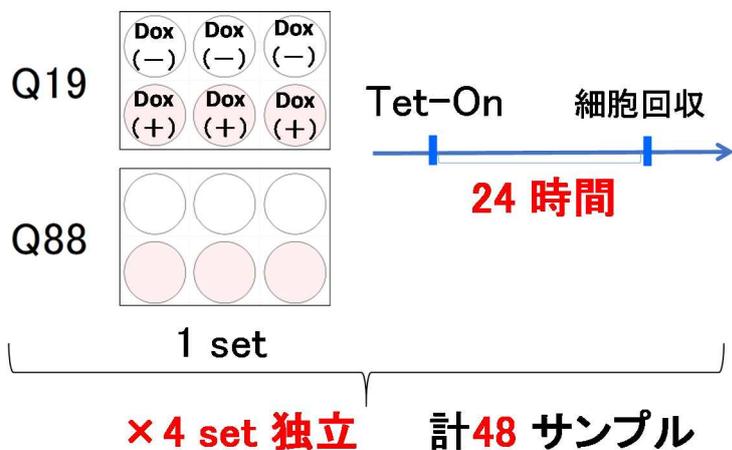
## 2. RNA-seq

### (1) RNA-seq に用いるための RNA サンプル抽出

Q19、Q88 の細胞は 6 well plate で約  $3.0 \times 10^6$  cells/well、培地 3 ml/well として前述のように培養した。Dox 添加する際の濃度は  $1 \mu\text{g}/\text{培地 } 1\text{ml}$  とした。Dox (+) 24 時間後に培養細胞から total RNA を抽出した。

Q19・Q88 の、各々 Dox (+) or Dox (-) の 2 条件について、同時並行に 3 つの well を用いて培養し、1 set で 12 サンプルとした。実験日と細胞の継代数を変えて合計 4 set、計 48 サンプルから total RNA を抽出した (図 6)。

図 6. RNA-seq に用いるための total RNA の抽出



今回の RNA-seq のためのサンプル処理の概略を示す。1 回のセットで、培養細胞を 3 well ずつ培養し、Dox (+) 24 時間後に RNA 抽出を行った。Q19・Q88 の細胞について、Dox (+) or Dox (-) の 2 条件について RNA 抽出を行うことを、合計 4 set 独立して行った。得られた 48 サンプルについて RNA-seq を行った。

## (2) シーケンスライブラリ調製

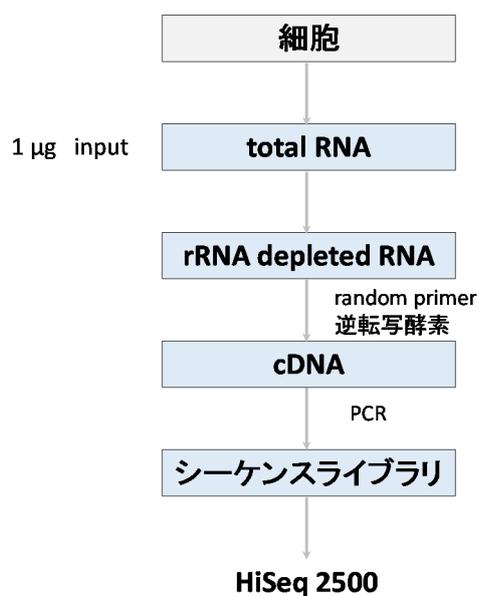
RNA-seqのためのシーケンスライブラリ調製は、細胞から RNeasy Plus Mini kit を用いて抽出した total RNA 1 $\mu$ g を用い、TruSeq Stranded Total RNA with Ribo Zero Gold Kit (Illumina, Inc., San Diego, CA, USA) のプロトコールにより行った (図 7)。具体的には、total RNA に対し種特異的 ribosomal RNA (rRNA) probe を加え 68°C 5 分間で total RNA を denature した後、probe と hybridization させ、マグネットビーズを加えて rRNA を除去した。Bioanalyzer と Agilent RNA 6000 pico kit (Agilent Technologies, Inc.) を用いて rRNA が除去されていることを確認した。2 価陽イオン (Elute, Prime, Fragment High Mix) と共に 94°C 8 分間 incubate し、insert length (adapter 間の DNA 長) が 120-210 bp (base pair) となるように断片化した。断片化 RNA に対し、random primer と逆転写酵素 (SuperScript® II Reverse Transcriptase (Invitrogen® by Thermo Fisher Scientific, Inc.)) を用いて double strand (ds) cDNA ライブラリを調製した。さらにシーケンサーのフローセル (スライドガラス) にハイブリダイゼーションするため、ds cDNA ライブラリ の両端に indexing adapter 配列 (ライブラリを識別するための 6 塩基のインデックスのついたアダプター) を ligate した。最後に PCR を 15 サイクル行い、シーケンスライブラリを作成した。Bioanalyzer と Agilent DNA1000 を用い、最終産物がおおよそ 260 bp となることを確認した。

## (3) RNA-seq

大規模シーケンスは、HiSeq 2500 を用いて、101 塩基長のペアエンド法を用いて行っ

た。各レーンに投入時、各サンプルは 2 nM の濃度とした。

図 7. Total RNA から index adapter 付加シーケンスライブラリの調製



培養細胞から抽出した total RNA 1 µg を input とし、strand specific シーケンスライブラリ調製を行った。total RNA より rRNA (ribosomal RNA) を除去し rRNA depleted RNA とした。random primer と逆転写酵素を用いてシーケンスライブラリを調製し HiSeq 2500 でペアエンド RNA-seq を行った。

### 3. RNA-seq の解析

#### (1) シーケンスデータの alignment と 2 つのソフトウェアを用いた遺伝子発現量算出

シーケンスデータ (FASTQ ファイル) のヒトリファレンスゲノム (UCSC human genome 19, hg19) に対する alignment は TopHat 2 (v2.0.8b<sup>[56]</sup>) - Bowtie 2 (v2.1.0<sup>[57]</sup>) を用いて行った。遺伝子発現量については、Cufflinks 2 (v2.1.1<sup>[58]</sup>) と edgeR (v3.6.8<sup>[59]</sup>) の 2 つのソフトウェアを用いて算出した。edgeR による処理の前に HTSeq (v0.6.1<sup>[60]</sup>) を使い、alignment データである bam ファイルから edgeR における遺伝子発現量算出に必要なカウントデータを算出した (図 8)。

なおソフトウェアの相違点は、Cufflinks 2 はカウントデータの単位が fragment (ペアエンド法で得られたリードを同一の DNA 分子由来のデータと考え、一つとしてカウントする単位) であり、alignment された 1 ライブラリ の全 fragment 数を 100 万 fragment とした場合の各遺伝子の相対的発現量を算出した後、遺伝子配列長で補正し、FPKM (Fragments Per Kilobase of exon per Million mapped fragments) という単位で各遺伝子の発現量を算出する。一方で edgeR はカウントデータの単位が read (ペアエンド法で得られたリードを独立にカウントする単位) であり、alignment された 1 ライブラリ の全 read 数を 100 万 read とした場合の各遺伝子の相対的発現量 (RPM : read per million mapped read) を算出した後、遺伝子長による補正は行わない代わりに、TMM 補正 (trimmed mean of M : RPM 値が突出して多い・または少ない遺伝子による、同一

サンプル内の他遺伝子のリード数の過小・または過大評価を補正することで発現変動遺伝子の偽陽性・偽陰性を減少する統計学的手法) を行い、各遺伝子の発現量として CPM (count per million) 値を算出した (表 1)。

表 1. Cufflinks 2 と edgeR の主な相違点

	Cufflinks 2	edgeR
遺伝子発現量の単位	FPKM	CPM
遺伝子発現量の元となるカウントデータ	fragment	read
遺伝子長による補正	有り	無し

なお、*DRPLA* 遺伝子の発現量を検討する際に、内在性の *DRPLA* と、導入コンストラクト由来の *GFP-DRPLA* 遺伝子の発現量について比較するため、以下の解析を行った。内在性 *DRPLA* 遺伝子の cDNA のスタートコドンの最初の A の上流・下流 100 塩基 (c.-100\_100 の 200 塩基) の配列及び、*GFP-DRPLA* 遺伝子コンストラクトのスタートコドンの最初の A の上流・下流 100 塩基からなる 200 塩基 (c.-100\_100) の 2 つの配列を準備した。BWA (v0.7.12<sup>[61]</sup>) を用いて得られた ショートリードをこの 2 つの配列に alignment し、中央 (c.-1) にマップされたリードの数を以て、内因性 *DRPLA* 遺伝子由来の発現と *GFP-DRPLA* 融合遺伝子由来の発現を検討した。

## (2) 発現変動遺伝子の抽出

各setにおいて、Dox (+) の3 サンプルとDox (-) の3サンプルを比較することで、発現変動遺伝子 (differentially expressed genes, DEGs) と遺伝子発現量比 (fold change) を算出した。統計学的な検定は、Cufflinks 2はt検定、edgeRはフィッシャーの正確検定により行い、false discovery rateにより多重比較の補正を行ったq値を算出した。有意水準は $q < 0.05$ とした。

## (3) 1<sup>st</sup> step-DEGs の抽出

可能な限り再現性のある結果を得るため、Cufflinks 2 によって解析した全 set で共通して有意に発現変動する遺伝子、edgeR によって解析した全 set で共通して有意に発現変動する遺伝子をそれぞれ 1<sup>st</sup> step-DEGs と定義した。遺伝子発現量の比である fold change が、1 以上の場合 up-regulated DEGs、1 未満の場合 down-regulated DEGs に分類した。

## (4) 1<sup>st</sup> step-DEGs を用いた Gene Ontology 解析

1<sup>st</sup> step-DEGs を対象に、各遺伝子に付与された Gene Ontology term<sup>[62]</sup> の特徴抽出を、DAVID Bioinformatics Resources 6.7<sup>[63]</sup> を用いて行った。

なお Gene Ontology (GO) とは生物学的概念を記述するために種を超えた共通の語彙の策定をすることを指し、biological process (BP)、cellular component (CC)、

molecular function (MF) につき、様々な具体性の程度の語彙が含まれている。具体性の程度については今回は FAT (具体性が低すぎる GO に対し filter をかけた後の適度に具体的な GO) を選択した。

DAVID は Fisher 正確検定により、帰無仮説「ある GO term が、入力した Gene list に付与されている頻度と、Genome 全体の遺伝子に付与されている頻度に差がない」の下に p 値を算出する。多重比較検定の補正值 q 値は Benjamini-Hochberg 法により求めた。Clustering については、1<sup>st</sup> step-DEGs の特徴を捉えるため、 $p < 0.05$  を満たす GO term に対して行った。なおこの GO 解析における clustering とは入力した遺伝子に付与される GO の中で、複数の遺伝子に共通して付与され、かつ意味が類似しているとされた GO term の群を指す。

## (5) 2<sup>nd</sup> step-DEGs の抽出

Cufflinks2、edgeR を用いて得られた 1<sup>st</sup> step-DEGs について、両ソフトウェアで共通して有意に変動する遺伝子を、2<sup>nd</sup> step-DEGs と定義した。

## (6) 2<sup>nd</sup> step-DEGs に対する qRT-PCR による validation

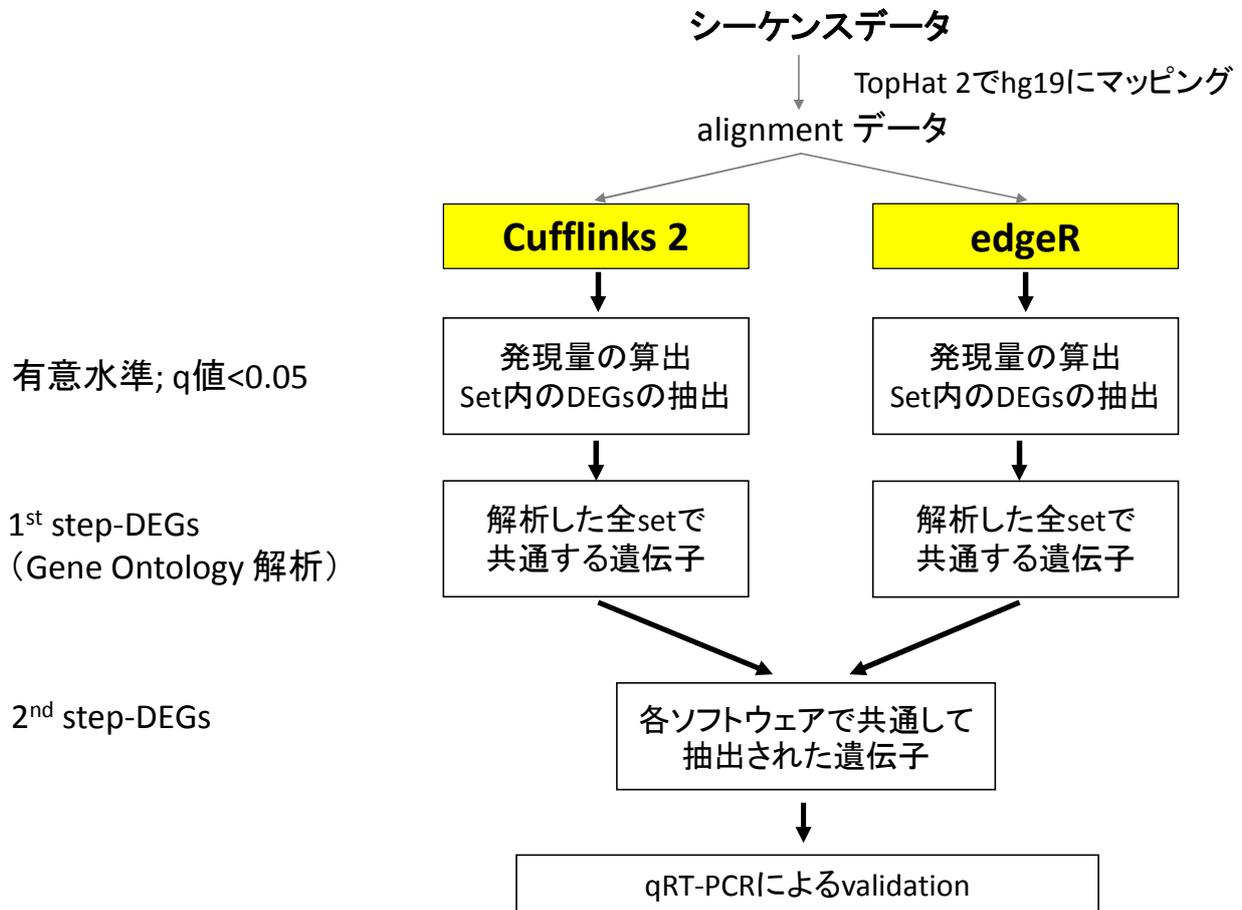
2<sup>nd</sup> step-DEGs の validation のため、qRT-PCR を行った。発現変動の方向性が同一であることを以て validate されたとした。

Q19 Dox (+) vs Dox (-) の 11 遺伝子、Q88 Dox (+) vs Dox (-) の遺伝子の 13 遺伝子

に対して、TaqMan® Gene Expression Assays (Applied Biosystems® by Thermo Fisher Scientific, Inc.) を用いて比較定量法により行った。qRT-PCR における相対的発現量比は、同一セットの中から Dox (+) vs Dox (-) の組を合計 3 組用い、計 3 回の実験の  $-\Delta\Delta Ct$  の平均から算出した。計 3 回の実験における標準誤差は、 $-\Delta\Delta Ct$  の標準誤差を用いた。

RNA-seq における相対的発現量比は、set 毎の FPKM 値の、(Dox (+) の発現量の平均) / (Dox (-) の発現量の平均) を用い、解析した全 set における平均値と標準誤差を算出した。DEGs 抽出から 2<sup>nd</sup> step-DEGs の validation までの解析概略を図 8 に示した。

図 8. シーケンスデータの取得から 2<sup>nd</sup> step-DEGs の validation までの解析概略



シーケンスデータを用いて TopHat 2 によりリファレンスゲノム (hg19) に alignment した。Cufflinks 2 と edgeR を用いて、遺伝子発現量の算出を行い、各 set において Dox (+) の 3 サンプルと Dox (-) の 3 サンプルの比較を行い、有意 ( $q < 0.05$ ) に発現変動する遺伝子 (DEGs) の抽出を行った。1<sup>st</sup> step-DEGs とは解析した全 set で共通して有意に変動する遺伝子と定義し、Gene Ontology 解析を行った。1<sup>st</sup> step-DEGs のうち両ソフトウェアで共通して有意に変動する遺伝子を 2<sup>nd</sup> step-DEGs と定義し、qRT-PCR により validation を行った。

## 結果

### 1. 実験系の構築

Tet-On による *GFP-DRPLA* 融合遺伝子の mRNA とタンパク質の発現の時間経過に関する検討

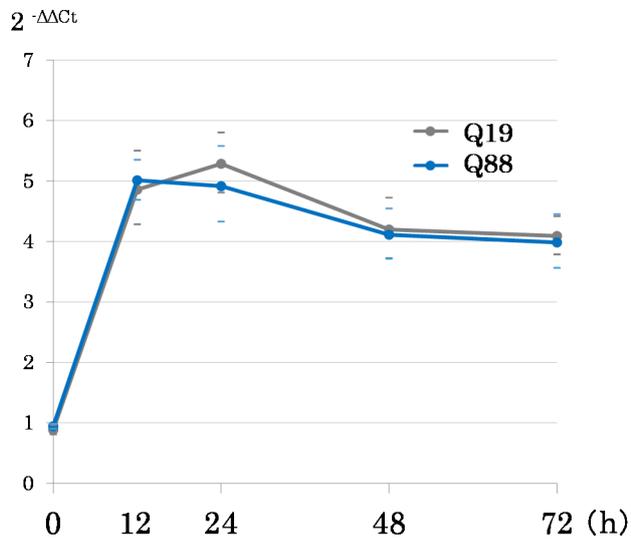
Tet-On による *GFP-DRPLA* 融合遺伝子の mRNA 発現の時間経過の観察を qRT-PCR により行った。qRT-PCR では、Tet-On 0・12・24・48・72 時間後までの観察を行った。

Q19・Q88 共に、*DRPLA* 遺伝子の相対的発現量は Tet-On 12・24 時間後にプラトーに達した (図 9)。

Tet-On 24 時間後における GFP-DRPLA (Q19 or Q88) 融合タンパク質の発現を、蛍光顕微鏡と (図 10)、Western blot により確認した (図 11)。また GFP-DRPLA (Q19) 融合タンパク質発現の時間経過を、Western blot により観察した (図 12)。Tet-On 24 時間後から 72 時間後までの観察において、GFP-DRPLA (Q19) 融合タンパク質の発現量は Tet-On 24 時間後に最大であった。

以上より、*DRPLA* 遺伝子の転写がプラトーに達し、GFP-DRPLA (Q19) 融合タンパク質の発現量が最大であった、Tet-On 24 時間後の total RNA を用いて RNAseq を行う方針とした。

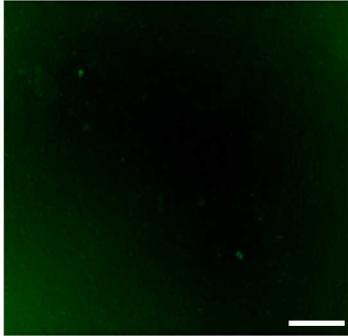
図 9. *GFP-DRPLA* (Q19 or Q88) 融合遺伝子の mRNA 発現の時間経過の qRT-PCR による観察



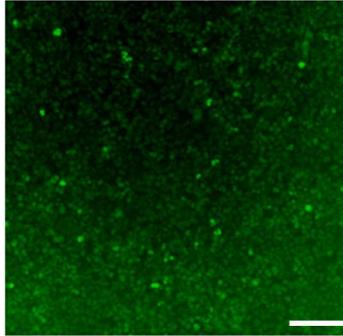
縦軸は endogenous control 遺伝子で補正した target 遺伝子の Dox (+) 時と Dox (-) 時の初期鋳型量比 ( $2^{-\Delta\Delta C_t}$ ) である。横軸は Tet-On 後の時間 (h) を示す。エラーバーは  $-\Delta\Delta C_t$  の標準誤差より算出した。Tet-On 0・12・24・48・72 時間後の観察を行い、Q19・Q88 共に、*DRPLA* 遺伝子の相対的発現量は Tet-On 12・24 時間後にプラトーに達した。実験回数は Q19・Q88 共に、各タイミングにつき 9 回である。

図 10. Tet-On 24 時間後の GFP- DRPLA (Q19 or Q88) 融合タンパク質の GFP の蛍光の観察

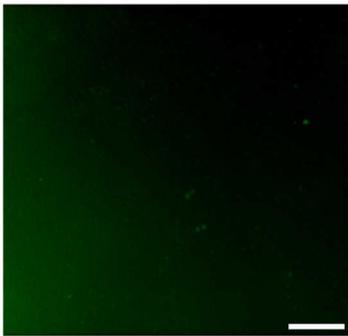
A. Q19 Dox (-)



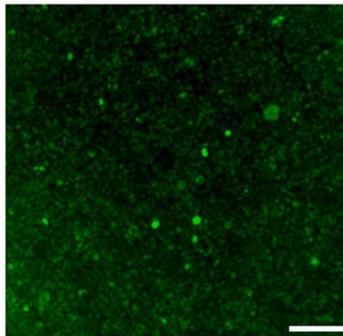
B. Q19 Dox (+)



C. Q88 Dox (-)

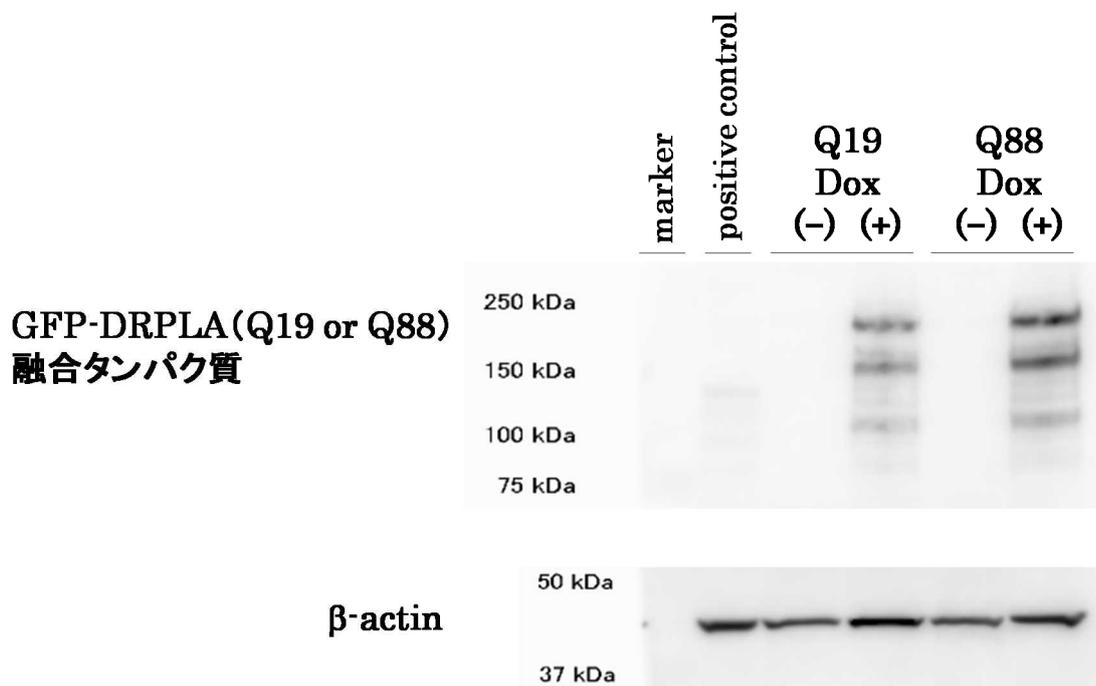


D. Q88 Dox (+)



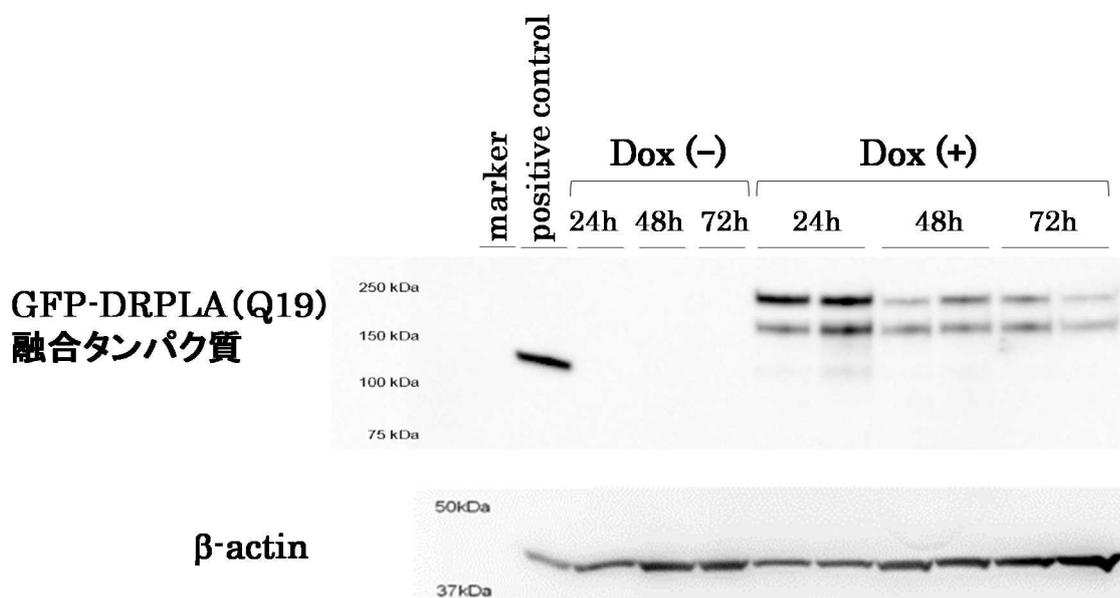
Tet-On 24 時間後の GFP-DRPLA (Q19 or Q88) 融合タンパク質の GFP の蛍光を、倒立蛍光顕微鏡により倍率 100 倍で確認した。A が Q19 Dox (-)、B が Q19 Dox (+)、C が Q88 Dox (-)、D が Q88 Dox (+) である。Scale bar は 100  $\mu\text{m}$  である。

図 11. Tet-On 24 時間後の GFP-DRPLA (Q19 or Q88) 融合タンパク質発現の Western blot による観察



Tet-On 24 時間後に Q19 または Q88 から抽出したタンパク質を用いた。GFP-DRPLA 融合タンパク質を抗 GFP 抗体を用いて、β-actin を抗β-actin 抗体を用いて標識した。Positive control は Tet-On 48 時間後の GFP-GR 細胞から抽出したタンパク質である。抗 GFP 抗体により 100 kDa から 250 kDa の間に 3 本のバンドが標識され、分子量が最大のバンドは GFP-full length DRPLA 融合タンパク質、その他のバンドは核移行の際に C 末端側が切断された GFP-truncated DRPLA 融合タンパク質であると考えられている。

図 12. Tet-On による GFP-DRPLA (Q19) 融合タンパク質発現の時間経過の Western blot による観察



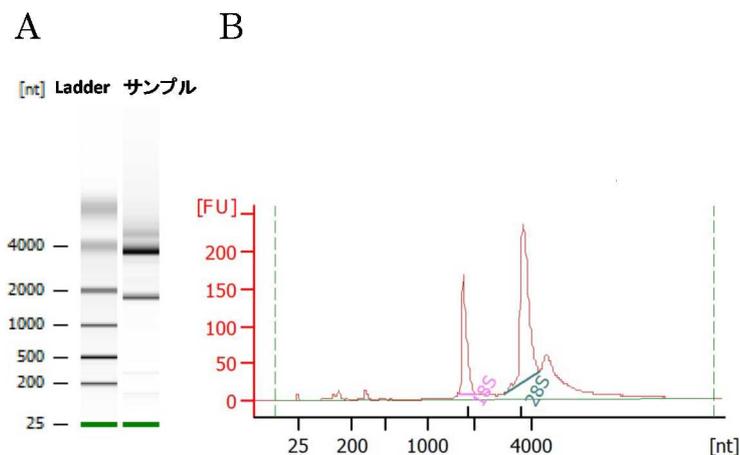
Tet-On 24・48・72 時間後に Q19 から抽出したタンパク質を用いた。GFP-DRPLA 融合タンパク質を抗 GFP 抗体を用いて、β-actin を抗β-actin 抗体を用いて標識した。Positive control は Tet-On 24 時間後の GFP-GR 細胞から抽出したタンパク質である。抗 GFP 抗体により 100 kDa から 250 kDa の間に 2 本または 3 本 (Dox (+) 24 時間後の右側のレーン) のバンドが標識され、分子量が最大のバンドは GFP-full length DRPLA 融合タンパク質、その他の分子量が小さいバンドは核移行の際に C 末端側が切断された GFP-truncated DRPLA 融合タンパク質と考えられる。GFP-DRPLA (Q19) 融合タンパク質の発現量は Tet-On 24 時間後が最大であった。

## 2. RNA-seq

### (1) RNA 抽出とシーケンスライブラリ調製

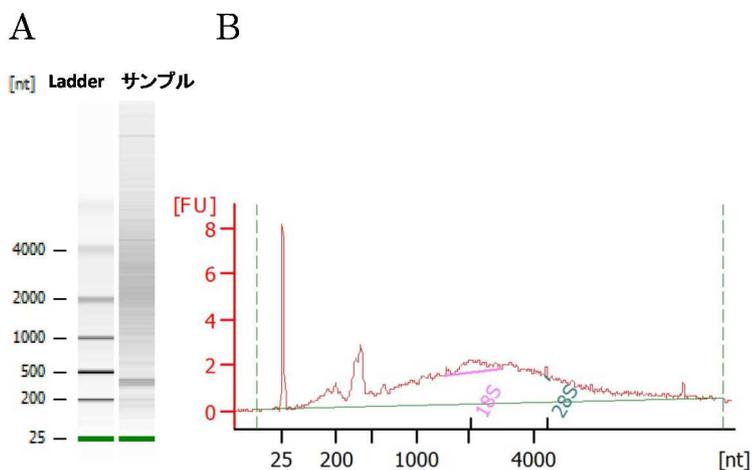
Total RNA からシーケンスライブラリを調製するため、約  $1.0 \times 10^6$  cell の培養細胞から約 15  $\mu\text{g}$  の total RNA を抽出し、Bioanalyzer と Agilent RNA 6000 Nano kit を用いて質の評価を行い RNA integrity number (RIN) が 9 以上であることを確認した (図 13)。さらに、total RNA に対し種特異的 ribosomal RNA probe とマグネットビーズにより ribosomal RNA を除去することで、total RNA から ribosomal RNA が除去されたことを、Bioanalyzer と Agilent RNA 6000 pico kit を用いて確認した (図 14)。Input の total RNA を 1  $\mu\text{g}$  とした場合、ribosomal RNA depleted RNA は 0.5-5 ng と、0.05~0.5%の量に減少した。シーケンスライブラリ調製後、Bioanalyzer と Agilent DNA1000 を用いて最終産物がおおよそ 260 bp となることを確認した (図 15)。Total RNA の質、ribosomal RNA の depletion の状況、シーケンスライブラリの長さの確認は、RNA-seq を行った 48 サンプル全てに対して行った。

図 13. Total RNA の質の評価



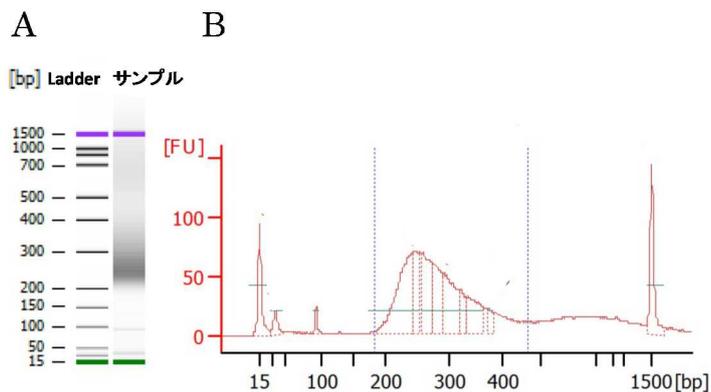
抽出した total RNA の質の評価の一例を示す、Bioanalyzer と Agilent RNA 6000 nano kit による結果である。A がバーチャル電気泳動、B は横軸を nt (nucleotide)、縦軸を FU (arbitrary fluorescent unit ; 蛍光色素で標識された total RNA 量) としてグラフ化したものである。概ね 2000 nt、4000 nt に、18S rRNA、28S rRNA に相当するピークが観察された。本サンプルにおいては RIN (RNA integrity number) は 9.8 であった。

図 14. Total RNA から ribosomal RNA が除去されたことの確認



Total RNA から ribosomal RNA が除去されたことを、Bioanalyzer と Agilent RNA6000 pico kit を用いて確認した 1 例を示した。A がバーチャル電気泳動、B は横軸を nt、縦軸を FU としてグラフ化したものである。Total RNA において、概ね 2000 nt、4000 nt に観察された、18S rRNA、28S rRNA に相当するピークの消失を確認した。

図 15. シーケンスライブラリの長さの確認



シーケンスライブラリが概ね 260 bp のバンドとなることを、Bioanalyzer と Agilent DNA1000 とを用いて確認した 1 例を示した。A がバーチャル電気泳動、B は A を横軸を bp、縦軸を FU としてグラフ化したものである。B における 150 bp-200 bp、400 bp-500 bp の間のグレーの 2 本の縦線は、バーチャル電気泳動におけるバンドの範囲を手動で設定したものである。このサンプルにおいては、グレーの 2 本の縦線と、赤のグラフの曲線と、FU=0 のベースラインで囲まれた範囲のシーケンスライブラリのサイズは、平均 290 bp (範囲; 184 bp-467 bp) と算出された。

## (2) RNA-seq

HiSeq2500 を用いて RNA-seq を行った。同一の set は同じフローセルでシーケンスした。Q88 の set 2 の Dox (+) の 3 サンプルは read 数のばらつきが大きかったため、Q88 の set 2 を解析から除外し、Q19 は 4 set、Q88 は 3 set 分の解析を行った。各サンプルにおいて平均 60 (範囲 : 33-93)  $\times 10^6$ /sample の read が得られ、TopHat 2 による alignment の結果、平均 30 (範囲 : 11-53)  $\times 10^6$ /sample の read が参照配列に alignment された (表 2)。

表 2. 計 42 サンプルの総 read 数と alignment された read 数

set	実験条件	sense鎖			antisense鎖		
		総read数	alignmentされたread数	alignment率	総read数	alignmentされたread数	alignment率
1	Q19 Off	44,735,211	28,393,438	63%	46,467,795	31,077,661	67%
		38,408,202	24,300,869	63%	39,916,769	26,696,335	67%
		58,070,735	34,615,965	60%	58,071,211	35,469,896	61%
	Q19 On	82,135,066	48,780,016	59%	82,188,757	48,869,435	59%
		32,780,710	15,518,388	47%	32,801,638	15,557,817	47%
		45,724,993	22,606,437	49%	45,754,044	22,684,855	50%
	Q88 Off	69,827,207	28,573,293	41%	69,826,763	29,145,691	42%
		48,583,626	26,084,549	54%	50,589,158	28,471,578	56%
		41,486,486	25,825,338	62%	43,083,691	28,038,866	65%
	Q88 On	66,090,183	38,021,682	58%	66,088,421	38,807,121	59%
		33,652,728	20,171,445	60%	34,943,093	22,094,518	63%
		72,179,979	42,485,136	59%	72,178,906	43,393,958	60%
平均	52,806,261	29,614,713	56%	53,492,521	30,858,978	58%	
2	Q19 Off	47,629,534	28,110,951	59%	47,625,158	28,403,644	60%
		65,873,370	35,380,587	54%	65,877,372	35,382,737	54%
		66,829,041	41,594,395	62%	66,834,803	41,604,665	62%
	Q19 On	65,663,946	37,953,761	58%	65,667,172	38,034,426	58%
		51,692,815	28,611,973	55%	51,696,107	28,624,134	55%
		93,201,580	53,367,225	57%	93,206,226	53,435,129	57%
	平均	65,148,381	37,503,149	58%	65,151,140	37,580,789	58%
3	Q19 Off	54,642,166	40,058,172	73%	54,645,506	40,120,731	73%
		66,307,265	42,509,588	64%	66,316,884	42,734,600	64%
		56,100,210	28,555,007	51%	56,111,276	29,015,141	52%
	Q19 On	52,882,634	21,010,270	40%	52,892,123	21,066,933	40%
		71,869,951	17,981,862	25%	71,882,980	18,085,758	25%
		54,541,125	32,048,365	59%	54,550,590	32,163,028	59%
	Q88 Off	62,340,264	41,942,530	67%	62,354,041	42,020,388	67%
		55,315,946	30,130,596	54%	55,328,473	30,159,551	55%
		67,745,860	46,128,156	68%	67,761,004	46,206,229	68%
	Q88 On	58,206,534	33,474,578	58%	58,222,564	33,588,597	58%
		55,320,966	24,999,545	45%	55,336,750	25,117,351	45%
		80,476,597	47,996,242	60%	80,506,447	49,865,693	62%
平均	61,312,460	33,902,909	55%	61,325,720	34,178,667	56%	
4	Q19 Off	58,419,174	24,366,637	42%	58,431,495	24,471,110	42%
		61,163,377	38,049,737	62%	61,163,377	38,208,762	62%
		52,882,318	22,929,773	43%	52,882,318	23,886,943	45%
	Q19 On	60,479,046	39,650,063	66%	60,479,046	39,891,979	66%
		74,613,466	23,771,850	32%	74,613,466	24,040,459	32%
		71,074,035	33,113,393	47%	71,074,035	33,419,011	47%
	Q88 Off	67,030,953	15,376,901	23%	67,030,953	15,484,150	23%
		69,456,211	13,995,427	20%	69,456,211	14,085,720	20%
		56,474,466	17,682,155	31%	56,474,466	17,766,867	31%
	Q88 On	76,697,281	11,427,895	15%	76,697,281	11,519,932	15%
		64,787,561	25,377,288	39%	64,787,561	25,455,033	39%
		53,783,642	21,104,701	39%	53,783,642	21,465,052	40%
平均	63,905,128	23,903,818	38%	63,906,154	24,141,251	39%	

### 3. RNA-seq の解析の全体像

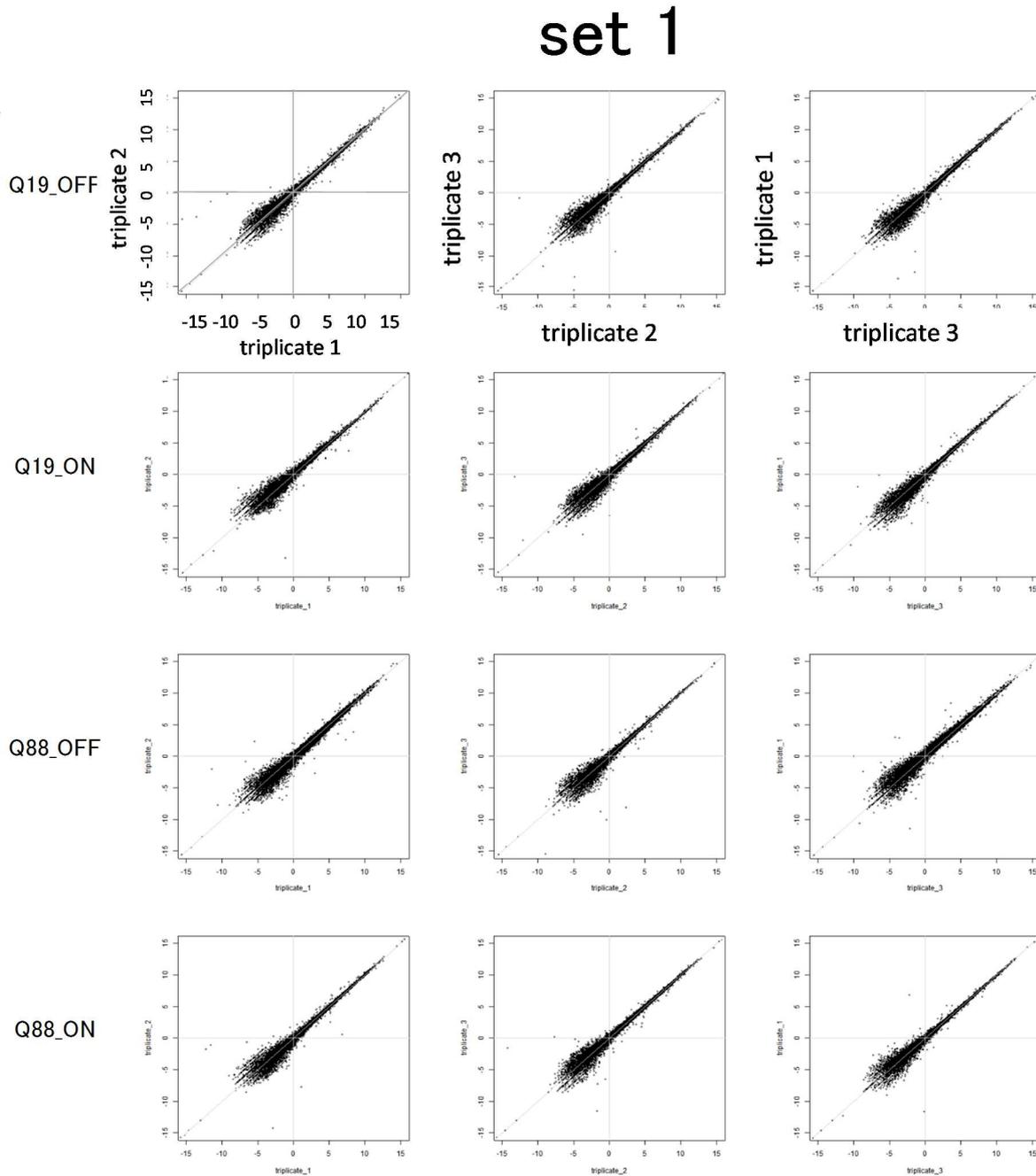
#### (1) 各 set の各条件毎の RNA-seq から得られた発現量の再現性

各 set の各条件の 3 サンプル同士の遺伝子発現量の再現性を、Cufflinks 2 によって算出した FPKM 値を用いて検討した。各 set の 3 サンプルの FPKM 値が全て 0 である遺伝子を除外したところ、観察対象となった遺伝子数は表 3 に示した通りであった。観察対象とした遺伝子に対して、縦軸・横軸を共に  $\log_2$  (FPKM) とする散布図 (図 16) を作成したところ、低発現量の遺伝子 ( $\log_2$  (FPKM)  $< 0$ ) においてばらつきが大きい傾向はあるものの、概ね  $y = x$  の直線上に分布した。各 set の各条件の中の 3 サンプルにつき、再現性を Spearman's correlation coefficient  $\rho$  (表 4) で検討したところ、全サンプルにおいて  $\rho > 0.98$  であった。

表 3. 各 set の 3 サンプル同士の FPKM 値の再現性の観察対象とした遺伝子数

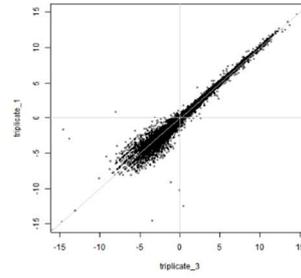
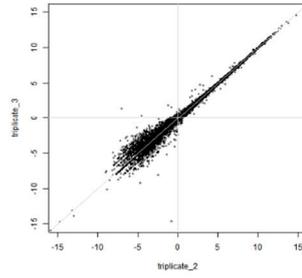
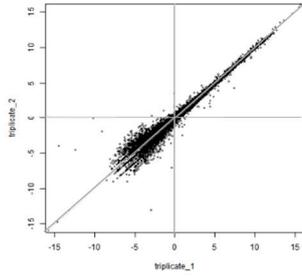
set	Q19	Q88
1	13544	13613
2	13971	—
3	13715	13952
4	13891	13244

図 16. 各 set の各条件における RNA-seq から得られた発現量の再現性

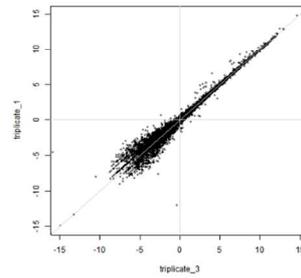
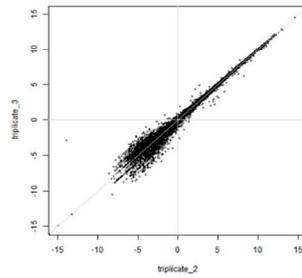
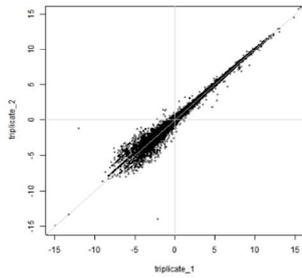


# set 2

Q19\_OFF

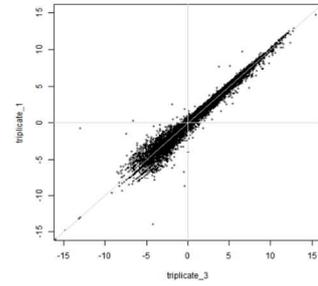
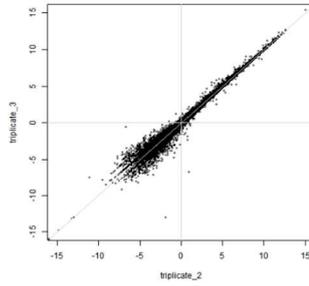
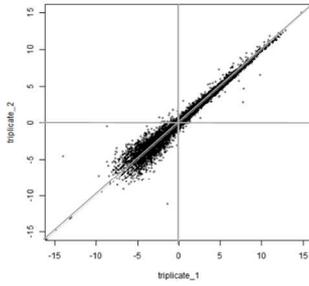


Q19\_ON

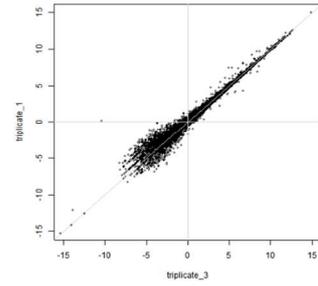
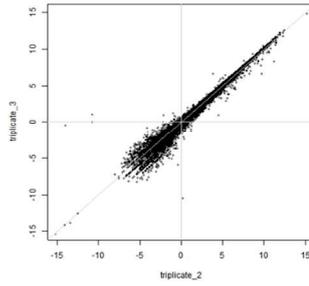
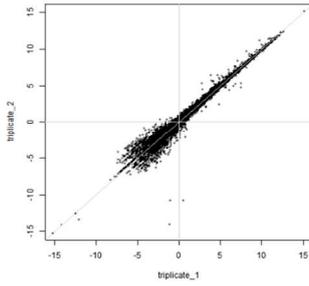


# set 3

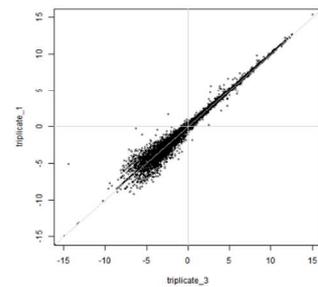
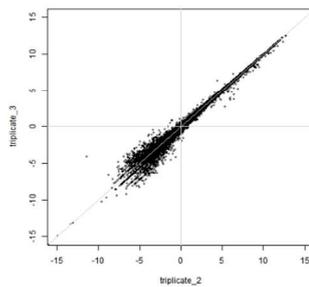
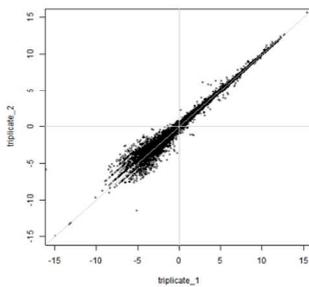
Q19\_OFF



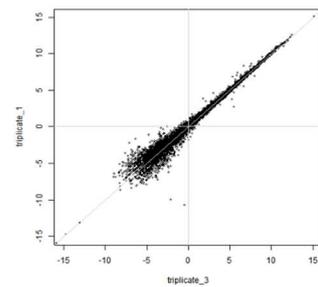
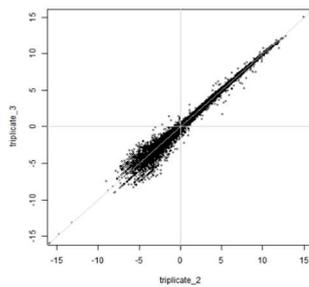
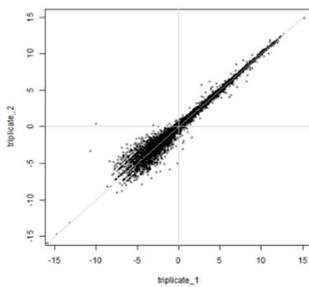
Q19\_ON



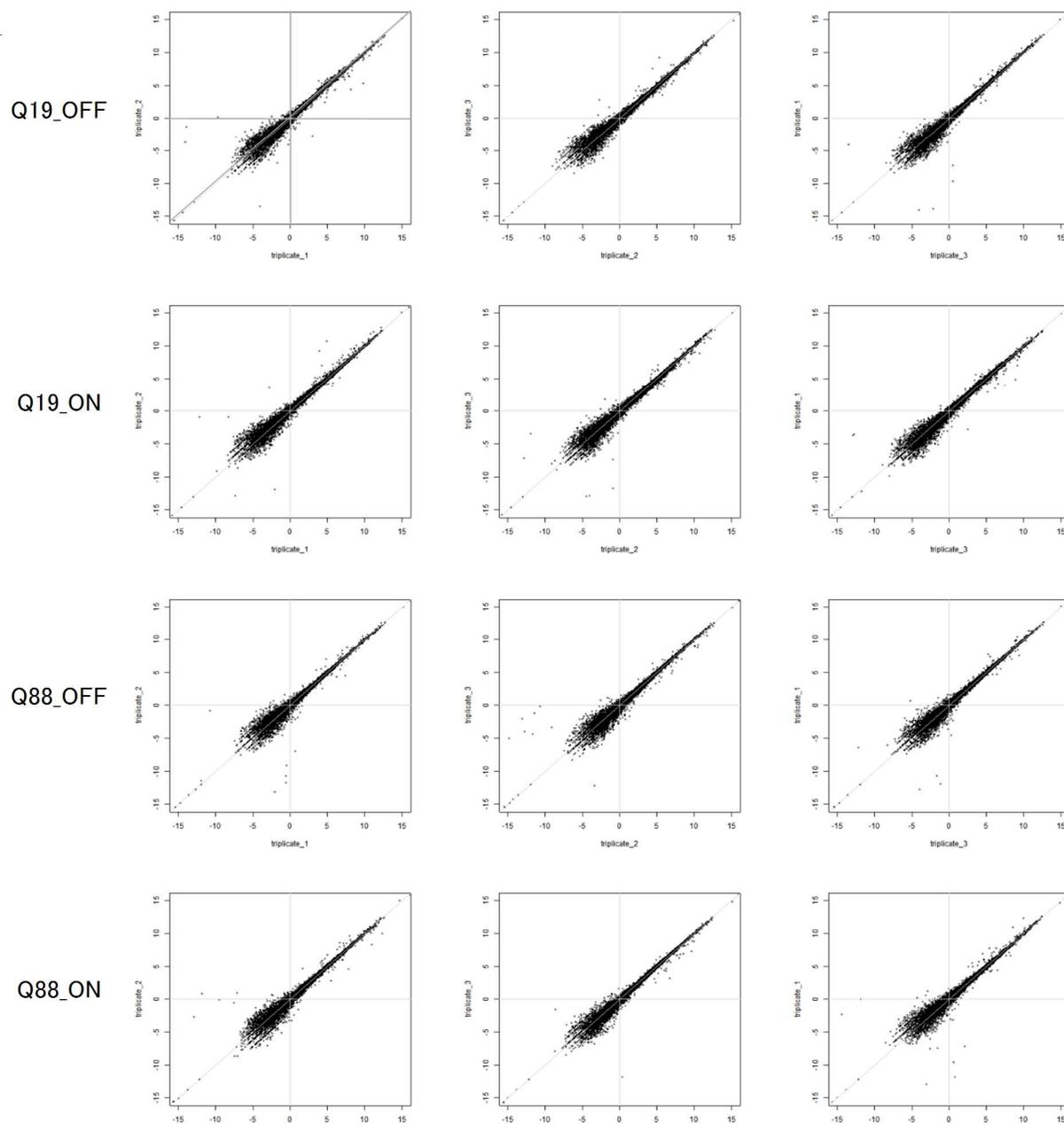
Q88\_OFF



Q88\_ON



# set 4



縦軸・横軸共に  $\log_2$  (FPKM) であり、散布図上の直線は  $y = 0$ 、 $x = 0$ 、 $y = x$  である。各遺伝子を示す各プロットは、いずれの組においても概ね  $y = x$  の直線上にプロットされた。

表 4. 各 set の各条件における 3 サンプル同士の Spearman's correlation coefficient  $\rho$

set	実験条件		Triplicate内の組		
			1 vs 2	2 vs 3	3 vs 1
1	Q19	Off	0.997	0.996	0.996
		On	0.996	0.996	0.997
	Q88	Off	0.992	0.997	0.992
		On	0.996	0.996	0.998
2	Q19	Off	0.996	0.998	0.996
		On	0.997	0.997	0.998
3	Q19	Off	0.995	0.997	0.992
		On	0.996	0.995	0.997
	Q88	Off	0.997	0.998	0.998
		On	0.997	0.997	0.997
4	Q19	Off	0.997	0.997	0.997
		On	0.997	0.996	0.996
	Q88	Off	0.997	0.996	0.997
		On	0.996	0.997	0.996

## (2) *DRPLA* 遺伝子の発現誘導と DEGs 数の確認

*DRPLA* 遺伝子が Tet-On により誘導されていることを示すために、Cufflinks 2、edgeR を用いた *DRPLA* 遺伝子の FPKM 値、CPM 値から、Q19 Dox (+) vs Dox (-)・Q88 Dox (+) vs Dox (-) の比較における fold change を算出した (表 5)。*DRPLA* 遺伝子は 2 つのソフトウェアによる解析で、全ての set において有意 ( $q < 0.05$ ) に発現誘導されていることが確認された。

また、内在性の *DRPLA* 遺伝子と、*GFP-DRPLA* 融合遺伝子の発現量について検討を行った結果を図 17 に示す。この検討から、Dox (-)の条件下でも *GFP-DRPLA* 融合遺伝子は少量発現しており、Dox 添加により発現量が増加することが示された。

Set 毎に、2 つのソフトウェアにより Q19 Dox (+) vs Dox (-)・Q88 Dox (+) vs Dox (-)

の比較における発現変動遺伝子 (DEGs) を抽出した。DEGs 数は用いたソフトウェア間で異なり、edgeR で多く認められる傾向があった (表 6)。

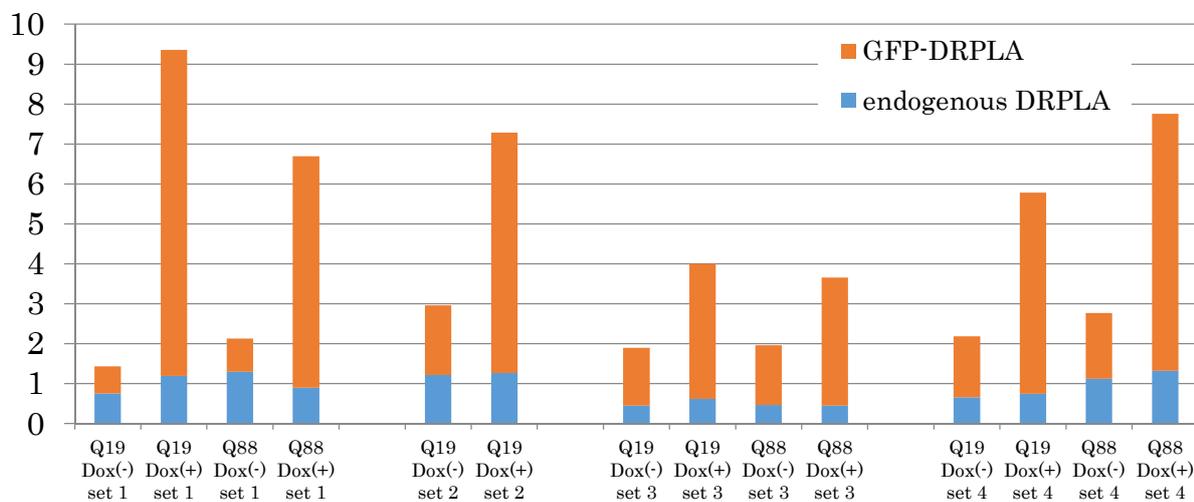
表 5. 2つのソフトウェアによる *DRPLA* 遺伝子の発現誘導

	set	Cufflinks 2										edgeR	
		fold change	q値	FPKM Dox (-) 3サンプル			SD	FPKM Dox (+) 3サンプル			SD	fold change	q値
<b>Q19</b> Dox (+) vs Dox (-)	1	<b>5.2</b>	0.0003	4	4	5	0.6	18	33	19	8.5	<b>9.4</b>	0.00325
	2	<b>2.1</b>	0.0068	7	9	9	0.9	20	18	14	3.3	<b>3.4</b>	3.07E-38
	3	<b>4.1</b>	0.0146	3	5	6	1.2	20	28	10	9.1	<b>10.8</b>	8.75E-39
	4	<b>2.8</b>	0.0116	9	7	8	1.4	10	37	20	13.4	<b>5.8</b>	3.13E-47
<b>Q88</b> Dox (+) vs Dox (-)	1	<b>2.4</b>	0.0004	11	4	3	4.5	17	11	16	2.9	<b>2.8</b>	2.09E-57
	3	<b>1.9</b>	0.0034	6	7	6	0.7	12	14	9	2.8	<b>4.8</b>	4.70E-23
	4	<b>3.2</b>	0.0099	13	15	10	2.3	45	43	32	6.7	<b>4.9</b>	1.26E-08

表 6. 2つのソフトウェアによる各 set における DEGs 数

	set	Cufflinks 2	edgeR
<b>Q19</b> Dox (+) vs Dox (-)	1	6498	7995
	2	172	899
	3	62	34
	4	88	218
<b>Q88</b> Dox (+) vs Dox (-)	1	2686	2810
	3	536	1964
	4	101	531

図 17. 内在性 *DRPLA* 遺伝子と *GFP-DRPLA* 融合遺伝子の発現量の比較



内在性 *DRPLA* 遺伝子と *GFP-DRPLA* 融合遺伝子の発現量について検討を行った。縦軸は、どちらかのシーケンスに alignment されたリードの数を million mapped reads 数で各々正規化した後、set 内の 3 サンプルで平均した値を示す。Dox (+) により、*GFP-DRPLA* 遺伝子の発現が増加していることが確認された。

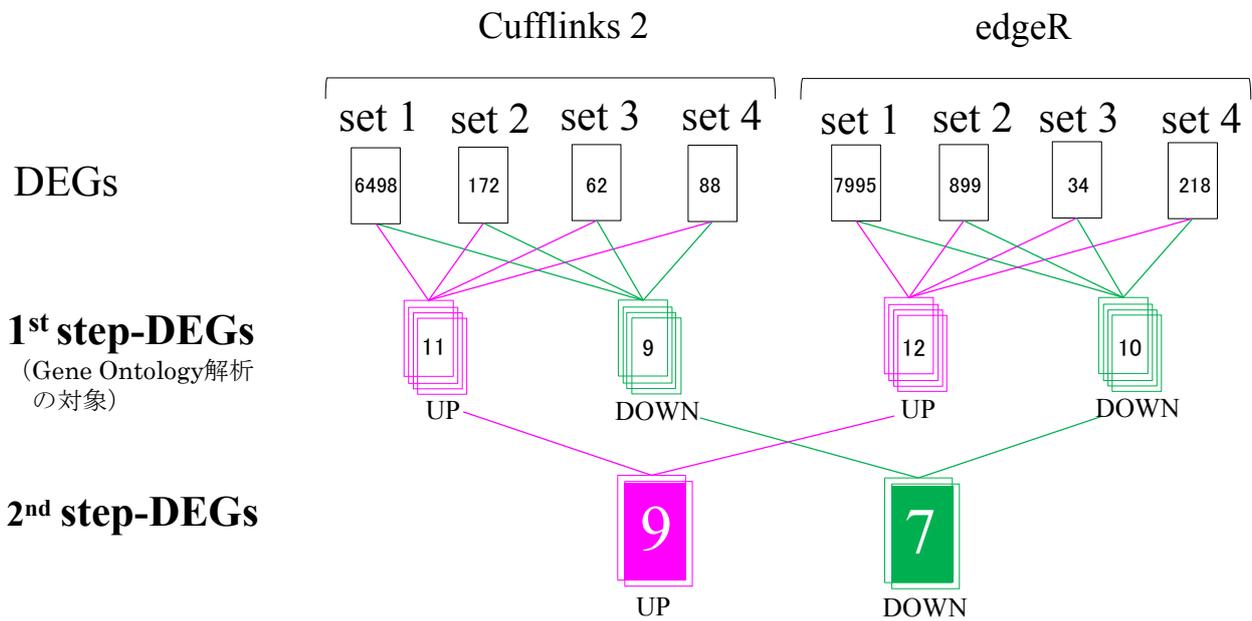
#### 4. Q19 Dox (+) vs Dox (-) の比較

ここではまず、wild type DRPLAp の転写の標的遺伝子を同定することを目的に、Q19 の Dox (+) vs Dox (-) の比較について検討した。

##### (1) 1<sup>st</sup> step-DEGs、2<sup>nd</sup> step-DEGs 数と遺伝子名

DEGs の数には、set 間で大きく異なることが判明したため、再現性のある結果を得るために、set 間で共通する発現変動遺伝子 (1<sup>st</sup> step-DEGs) を抽出した。次に、Cufflinks 2 と edgeR とで共通に有意に発現変動する遺伝子 (2<sup>nd</sup> step-DEGs) を抽出したところ、up-regulated DEGs が 9 遺伝子、down-regulated DEGs が 7 遺伝子であった。1<sup>st</sup> step-DEGs、2<sup>nd</sup> step-DEGs の数と遺伝子名を図 18 と表 7 に示した。

図 18 . Q19 Dox (+) vs Dox (-) における 1<sup>st</sup> step-DEGs と 2<sup>nd</sup> step-DEGs の数



1<sup>st</sup> step-DEGs は、Cufflinks 2 による up-regulated DEGs が 11 遺伝子、down-regulated DEGs が 9 遺伝子、edgeR による up-regulated DEGs が 12 遺伝子、down-regulated DEGs が 10 遺伝子であった。2<sup>nd</sup> step-DEGs は、up-regulated DEGs が 9 遺伝子、down-regulated DEGs が 7 遺伝子であった。

表 7. Q19 Dox (+) vs Dox (-) の 1<sup>st</sup> step-DEGs と 2<sup>nd</sup> step-DEGs

1 <sup>st</sup> step-DEGs				2 <sup>nd</sup> step-DEGs	
Cufflinks 2		edgeR		up-regulated 9遺伝子	down-regulated 7遺伝子
up-regulated 11遺伝子	down-regulated 9遺伝子	up-regulated 12遺伝子	down-regulated 10遺伝子		
<i>ARID5B</i>	<i>AMOT</i>	<i>ARID5B</i>	<i>AMOT</i>	<i>ARID5B</i>	<i>AMOT</i>
<i>ATN1</i>	<i>ANXA1</i>	<i>ATN1</i>	<i>ANXA1</i>	<i>ATN1</i>	<i>ANXA1</i>
<i>DDIT4</i>	<i>CCT4</i>	<i>DDIT4</i>	<i>CCT4</i>	<i>DDIT4</i>	<i>CCT4</i>
<i>FOXB1</i>	<i>KAL1</i>	<i>EMP1</i>	<i>GLO1</i>	<i>FOXB1</i>	<i>KAL1</i>
<i>HOXD11</i>	<i>PPP2R2A</i>	<i>FOXB1</i>	<i>KAL1</i>	<i>HOXD11</i>	<i>PYGL</i>
<i>HOXD9</i>	<i>PYGL</i>	<i>HOXC6</i>	<i>MID1</i>	<i>ITGA8</i>	<i>RHOA</i>
<i>ITGA8</i>	<i>RHOA</i>	<i>HOXD11</i>	<i>PYGL</i>	<i>PRSS12</i>	<i>TSC22D3</i>
<i>MTRNR2L8</i>	<i>TMEM45A</i>	<i>ITGA8</i>	<i>RHOA</i>	<i>RASL11A</i>	
<i>PRSS12</i>	<i>TSC22D3</i>	<i>PRSS12</i>	<i>THAP9-AS1</i>	<i>SEMA3C</i>	
<i>RASL11A</i>		<i>RASL11A</i>	<i>TSC22D3</i>		
<i>SEMA3C</i>		<i>RNLS</i>			
		<i>SEMA3C</i>			

## (2) 1<sup>st</sup> step-DEGs を用いた Gene Ontology 解析

Q19 Dox (+) vs Dox (-) の比較における、1<sup>st</sup> step-DEGs (ソフトウェアごとに、4つの set で共通して有意に変動する遺伝子) につき、各遺伝子に付与された Gene Ontology term の特徴抽出を、DAVID Bioinformatics Resources 6.7 を用いて行った。Cufflinks 2、edgeR のそれぞれにおいて up-regulate された 1<sup>st</sup> step-DEGs について (表 8) は、Cufflinks 2、edgeR 両者とも、骨格の発生、転写調節、RNA 代謝に関する GO term が cluster を形成した。Cufflinks 2 と edgeR のそれぞれにおいて down-regulate された 1<sup>st</sup> step-DEGs は、Cufflinks 2、edgeR 両者とも、cluster を形成しなかった。

表 8. Q19 Dox (+) vs Dox (-) の比較において up-regulate された 1<sup>st</sup> step-DEGs (Cufflinks 2 による 11 遺伝子、edgeR による 12 遺伝子)

Cufflinks 2による 11遺伝子によるAnnotation Cluster

Category	Term	PValue	Benjamini	Genes
GOTERM_BP_FAT	<b>skeletal system morphogenesis</b>	0.002	0.129	<i>HOXD9, ARID5B, HOXD11</i>
GOTERM_MF_FAT	<b>transcription regulator activity</b>	0.009	0.270	<i>HOXD9, ATN1, ARID5B, FOXB1, HOXD11</i>
GOTERM_BP_FAT	<b>skeletal system development</b>	0.018	0.470	<i>HOXD9, ARID5B, HOXD11</i>
GOTERM_BP_FAT	<b>regulation of transcription, DNA-dependent</b>	0.021	0.471	<i>HOXD9, ATN1, ARID5B, FOXB1, HOXD11</i>
GOTERM_BP_FAT	<b>regulation of RNA metabolic process</b>	0.023	0.445	<i>HOXD9, ATN1, ARID5B, FOXB1, HOXD11</i>
SP_PIR_KEYWORDS	<b>dna-binding</b>	0.049	0.606	<i>HOXD9, ARID5B, FOXB1, HOXD11</i>

edgeRによる 12遺伝子によるAnnotation Cluster

Category	Term	PValue	Benjamini	Genes
GOTERM_MF_FAT	<b>transcription regulator activity</b>	0.009	0.257	<i>HOXC6, ATN1, ARID5B, FOXB1, HOXD11</i>
GOTERM_MF_FAT	<b>transcription repressor activity</b>	0.015	0.227	<i>HOXC6, ATN1, ARID5B</i>
GOTERM_BP_FAT	<b>skeletal system development</b>	0.026	0.647	<i>HOXC6, ARID5B, HOXD11</i>
GOTERM_BP_FAT	<b>regulation of transcription, DNA-dependent</b>	0.045	0.763	<i>HOXC6, ATN1, ARID5B, FOXB1, HOXD11</i>
GOTERM_BP_FAT	<b>regulation of RNA metabolic process</b>	0.049	0.725	<i>HOXC6, ATN1, ARID5B, FOXB1, HOXD11</i>

### (3) qRT-PCR による validation

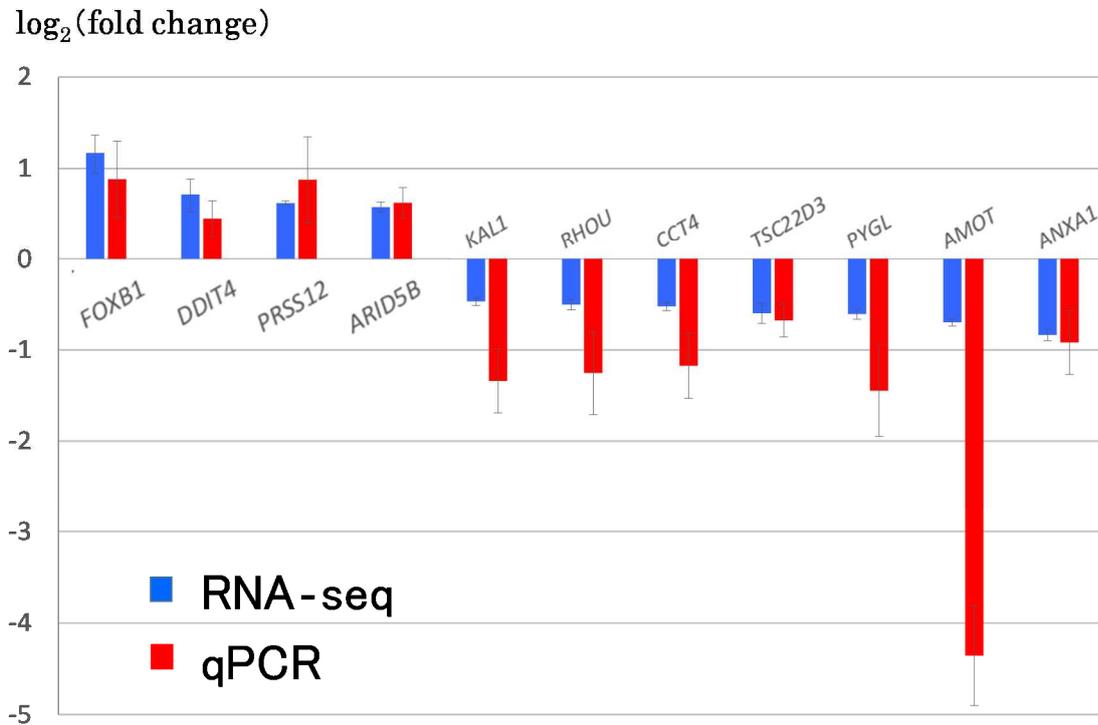
Q19 Dox (+) vs Dox (-) の 2<sup>nd</sup> step-DEGs のうち 11 遺伝子に対して qRT-PCR による validation を行った。11 遺伝子の gene symbol と probe の Assay ID を示した (表 9)。

ここで解析した 11 遺伝子については、RNA-seq と qRT-PCR における regulation が同方向であることを確認した (図 19)。qRT-PCR により validate された発現変動遺伝子 11 遺伝子とそのディスクリプションを示した (表 10)。

表 9. qRT-PCR による validation を行った 2<sup>nd</sup> step-DEGs 11 遺伝子と Assay ID

gene symbol	Assay ID
<i>ARID5B</i>	Hs01382781_m1
<i>DDIT4</i>	Hs01111686_g1
<i>FOXB1</i>	Hs00247213_s1
<i>PRSS12</i>	Hs00186221_m1
<i>AMOT</i>	Hs00611096_m1
<i>ANXA1</i>	Hs00167549_m1
<i>CCT4</i>	Hs00272345_m1
<i>KAL1</i>	Hs01085107_m1
<i>PYGL</i>	Hs00958087_m1
<i>RHOU</i>	Hs00221873_m1
<i>TSC22D3</i>	Hs00272345_m1

図 19. 2<sup>nd</sup> step-DEGs のうち 11 遺伝子に対する qRT-PCR による validation



縦軸は Q19 Dox (+) vs Dox (-) の相対的発現量比の底を 2 とする対数軸である。RNA-seq における相対的発現量比は、set 毎の FPKM 値の、(Dox (+) の発現量の平均) / (Dox (-) の発現量の平均) の 4 set における平均値であり、エラーバーは (Dox (+) の発現量の平均) / (Dox (-) の発現量の平均) の 4 set における標準誤差より算出した。qRT-PCR における相対的発現量比は 3 回の実験の  $-\Delta\Delta Ct$  の平均から算出した ( $-\Delta\Delta Ct = \log_2(2^{-\Delta\Delta Ct \text{ ave.}})$ )。qRT-PCR のエラーバーは  $-\Delta\Delta Ct$  の標準誤差である。ここに挙げた 11 遺伝子は、RNA-seq と qRT-PCR における regulation が同方向であることを確認した。

表 10. qRT-PCR により validate された 11 遺伝子とそのディスクリプション

Gene symbol	Description
<b>up-regulated 4遺伝子</b>	
<i>ARID5B</i>	AT-rich interaction domain 5B
<i>DDIT4</i>	DNA damage inducible transcript 4
<i>FOXB1</i>	forkhead box B1
<i>PRSS12</i>	protease, serine 12
<b>down-regulated 7遺伝子</b>	
<i>AMOT</i>	angiominin
<i>ANXA1</i>	annexin A1
<i>CCT4</i>	chaperonin containing TCP1, subunit 4 (delta)
<i>KAL1</i>	Kallmann syndrome 1 sequence
<i>PYGL</i>	phosphorylase, glycogen, liver
<i>RHOU</i>	ras homolog family member U
<i>TSC22D3</i>	TSC22 domain family member 3

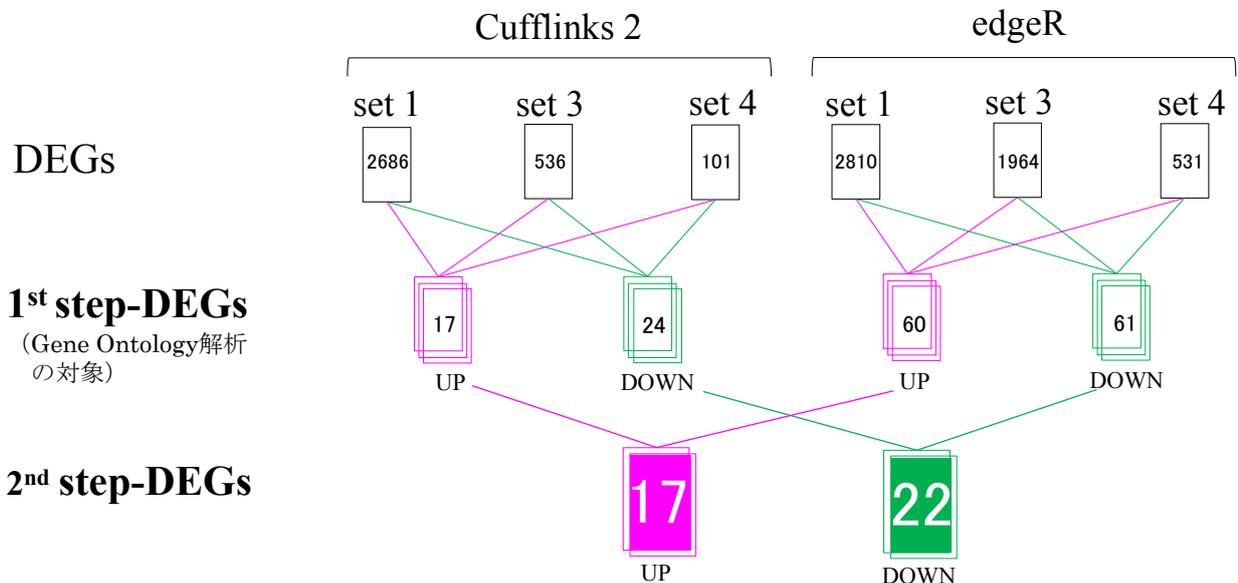
## 5. Q88 Dox (+) vs Dox (-) の比較

Mutant DRPLAp の転写調節標的遺伝子の同定のため、同じ set 内の Q88 の Dox (+) vs Dox (-) の 3 サンプルについて比較を行った結果を以下に示す。

### (1) 1<sup>st</sup> step-DEGs、2<sup>nd</sup> step-DEGs 数と遺伝子名

Q88 Dox (+) vs Dox (-) の解析により、up-regulated DEGs が 17 遺伝子、down-regulated DEGs が 22 遺伝子が抽出された。1<sup>st</sup> step-DEGs、2<sup>nd</sup> step-DEGs の数と遺伝子名を図 20 と表 11 に示した。

図 20. Q88 Dox (+) vs Dox (-) における 1<sup>st</sup> step-DEGs と 2<sup>nd</sup> step-DEGs の数



1<sup>st</sup> step-DEGs は、Cufflinks 2 による up-regulated DEGs が 17 遺伝子、down-regulated DEGs が 24 遺伝子、edgeR による up-regulated DEGs が 60 遺伝子、down-regulated DEGs が 61 遺伝子であった。2<sup>nd</sup> step-DEGs は、up-regulated DEGs が 17 遺伝子、down-regulated DEGs が 22 遺伝子であった。

表 11. Q88 Dox (+) vs Dox (-) の 1<sup>st</sup> step-DEGs と 2<sup>nd</sup> step-DEGs

		1 <sup>st</sup> step-DEGs						2 <sup>nd</sup> step-DEGs	
Cufflinks 2		edgeR						up-regulated 17遺伝子	down-regulated 22遺伝子
up-regulated 17遺伝子	down-regulated 24遺伝子	up-regulated 60遺伝子			down-regulated 61遺伝子				
ARID5B	ACAD11	AGPAT9	HOXD11	SH3BGR3	AKIRIN2	MEF2C	SLC4A7	ARID5B	AKIRIN2
ATN1	AKIRIN2	ALDH1A2	HOXD13	SLC12A9	AMOT	MID1	SLC9A2	ATN1	AMOT
CDKN2C	AMOT	ALYREF	HOXD9	SRP14	ANXA1	MID2	SMS	CDKN2C	ANXA1
DDIT4	ANXA1	ANXA2	HSD17B10	SSR2	CCT4	NHS	SNTB1	DDIT4	CSNK1E
EMP1	BDNF	ARID5B	ID3	SSR4	CCT7	NUP62CL	SPEN	EMP1	EDEM1
FOXB1	CSNK1E	ATN1	IMPDH2	STRA13	CSNK1E	OTUD1	STC1	FOXB1	EIF4G3
H1FO	EDEM1	ATP6AP1	INPP5K	SURF1	CXorf57	PCDH19	SYTL5	GBE1	GBE1
HOXD10	EIF4G3	C11orf48	ITGA8	TES	DCAF17	PDK1	THAP9-AS1	HOXD10	KAL1
HOXD11	GBE1	CDKN2C	LAGE3	TSPAN12	E2F3	PDK3	TMEM45A	HOXD11	KCNJ8
HOXD13	KAL1	CNPY2	MCM5	VAT1L	EDEM1	PIGA	WAC-AS1	HOXD13	LOC100506548
ITGA8	KCNJ8	CRELD2	MYL6	WNT16	EIF4G3	PLOD2	XK	ITGA8	LOX
NTS	LOC100506548	DDIT4	NDUFB10	ZNF608	ENPP2	PPP2R2A	ZBTB21	NTS	MID1
PRSS12	LOX	EMP1	NEU1		ERAP2	PYGL	ZNF711	PRSS12	NHS
RASL11A	MID1	FOXB1	NR2F1		FAT4	RAB9A		RASL11A	PCDH19
SEMA3C	NHS	FRZB	NTS		GBE1	RHO		SEMA3C	PPP2R2A
VAT1L	PCDH19	FSTL1	OST4		GK	RPL22L1		VAT1L	PYGL
WNT16	PPP2R2A	GALNS	POLR2F		GLO1	RPL24		WNT16	RHO
	PYGL	GAMT	PRSS12		KAL1	RPL34			SLC9A2
	RHO	GAS1	PSAP		KCNJ8	RPL6			SMS
	SLC9A2	GRN	RASL11A		KIAA2022	RPL7L1			STC1
	SMS	H1FO	RNLS		LIMA1	RPS6KA3			THAP9-AS1
	STC1	HOXB-AS3	SCAMP3		LOC10050654	SAMD9			TMEM45A
	THAP9-AS1	HOXC6	SELENBP1		LOX	SCML1			
	TMEM45A	HOXD10	SEMA3C		LPL	SERBP1			

## (2) 1<sup>st</sup> step-DEGs を用いた Gene Ontology 解析

Q88 Dox (+) vs Dox (-) の比較における、1<sup>st</sup> step-DEGs につき、各遺伝子に付与された Gene Ontology (GO) term の特徴抽出を、DAVID Bioinformatics Resources 6.7 を用いて行い、 $p < 0.05$  を満たす GO term が形成する cluster を観察した。

Cufflinks 2 と edgeR のそれぞれの解析において up-regulate された 1<sup>st</sup> step-DEGs については (表 12、表 13)、Q19 Dox (+) vs Dox (-) の比較において up-regulate された 1<sup>st</sup> step-DEGs と類似して、主に四肢・骨格の発生に関する GO term が cluster を形成した。

Cufflinks 2 の解析において down-regulate された 1<sup>st</sup> step-DEGs からは  $p < 0.05$  を満たす GO term は得られなかった。edgeR の解析において down-regulate された 1<sup>st</sup> step-DEGs から得られた  $p < 0.05$  を満たす GO term は、細胞質タンパク質、糖代謝に関する cluster を形成した (表 14)。

表 12. Q88 Dox (+) vs Dox (-) で up-regulate された 1<sup>st</sup> step-DEGs (Cufflinks 2 による 17 遺伝子)

Annotation Cluster				
Category	Term	PValue	Benjamini	Genes
GOTERM_BP_FAT	urogenital system development	0.0003	0.0673	<i>ITGA8, ARID5B, HOXD13, HOXD11</i>
SP_PIR_KEYWORDS	DNA binding	0.0026	0.1281	<i>H1FO, HOXD13, HOXD10, HOXD11</i>
SP_PIR_KEYWORDS	developmental protein	0.0033	0.0836	<i>WNT16, ITGA8, HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	embryonic limb morphogenesis	0.0046	0.4490	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	embryonic appendage morphogenesis	0.0046	0.4490	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	embryonic morphogenesis	0.0052	0.3605	<i>ITGA8, HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	skeletal system development	0.0058	0.2581	<i>ARID5B, HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	appendage morphogenesis	0.0060	0.2256	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	limb morphogenesis	0.0060	0.2256	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	appendage development	0.0064	0.2108	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	limb development	0.0064	0.2108	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	skeletal system morphogenesis	0.0076	0.2163	<i>ARID5B, HOXD10, HOXD11</i>
UP_SEQ_FEATURE	DNA-binding region:Homeobox	0.0108	0.5362	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	anterior/posterior pattern formation	0.0116	0.2834	<i>HOXD13, HOXD10, HOXD11</i>
SP_PIR_KEYWORDS	dna-binding	0.0151	0.2352	<i>H1FO, ARID5B, HOXD13, FOXB1, HOXD10, HOXD11</i>
SP_PIR_KEYWORDS	Homeobox	0.0168	0.2015	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_MF_FAT	transcription regulator activity	0.0173	0.5815	<i>ATN1, ARID5B, HOXD13, FOXB1, HOXD10, HOXD11</i>
INTERPRO	IPR017970:Homeobox, conserved site	0.0204	0.6201	<i>HOXD13, HOXD10, HOXD11</i>
INTERPRO	IPR001356:Homeobox	0.0209	0.3909	<i>HOXD13, HOXD10, HOXD11</i>
INTERPRO	IPR012287:Homeodomain-related	0.0214	0.2872	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	regionalization	0.0221	0.4073	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_MF_FAT	sequence-specific DNA binding	0.0252	0.4713	<i>HOXD13, FOXB1, HOXD10, HOXD11</i>
SMART	SM00389:HOX	0.0262	0.2917	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	pattern specification process	0.0388	0.5717	<i>HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	regulation of transcription, DNA-dependent	0.0484	0.6253	<i>ATN1, ARID5B, HOXD13, FOXB1, HOXD10, HOXD11</i>

表 13. Q88 Dox (+) vs Dox (-) で up-regulate された 1<sup>st</sup> step-DEGs (edgeR による 60 遺伝子)

Annotation Cluster 1				
Category	Term	PValue	Bonferroni	Genes
GOTERM_BP_FAT	skeletal system development	0.00001	0.0059	<i>HOXD9, HOXC6, ARID5B, HOXD13, GAS1, FRZB, HOXD10, HOXD11, ANXA2</i>
GOTERM_BP_FAT	embryonic limb morphogenesis	0.00001	0.0061	<i>HOXD9, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	embryonic appendage morphogenesis	0.00001	0.0061	<i>HOXD9, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	limb morphogenesis	0.00002	0.0115	<i>HOXD9, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	appendage morphogenesis	0.00002	0.0115	<i>HOXD9, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	appendage development	0.00002	0.0139	<i>HOXD9, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	limb development	0.00002	0.0139	<i>HOXD9, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	regionalization	0.00005	0.0283	<i>HOXD9, HOXC6, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	proximal/distal pattern formation	0.00006	0.0344	<i>HOXD9, ALDH1A2, HOXD10, HOXD11</i>
GOTERM_BP_FAT	anterior/posterior pattern formation	0.00011	0.0588	<i>HOXD9, HOXC6, ALDH1A2, HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	pattern specification process	0.00027	0.1410	<i>HOXD9, HOXC6, ALDH1A2, HOXD13, GAS1, HOXD10, HOXD11</i>
SP_PIR_KEYWORDS	DNA binding	0.00050	0.0652	<i>HOXD9, HOXC6, H1F0, HOXD13, HOXD10, HOXD11, NR2F1</i>
GOTERM_BP_FAT	urogenital system development	0.00052	0.2535	<i>ALDH1A2, ITGA8, ARID5B, HOXD13, HOXD11</i>
GOTERM_BP_FAT	skeletal system morphogenesis	0.00056	0.2687	<i>HOXD9, ARID5B, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	embryonic morphogenesis	0.00057	0.2735	<i>HOXD9, ALDH1A2, ITGA8, HOXD13, GAS1, HOXD10, HOXD11</i>
GOTERM_BP_FAT	embryonic skeletal system development	0.00226	0.7198	<i>HOXD9, HOXC6, GAS1, HOXD10</i>
UP_SEQ_FEATURE	DNA-binding region:Homeobox	0.00248	0.4016	<i>HOXD9, HOXC6, HOXD13, HOXD10, HOXD11</i>
GOTERM_BP_FAT	forelimb morphogenesis	0.00250	0.7560	<i>HOXD9, ALDH1A2, HOXD10</i>
GOTERM_BP_FAT	embryonic organ development	0.00272	0.7838	<i>HOXD9, ALDH1A2, ITGA8, GAS1, HOXD10</i>
GOTERM_BP_FAT	embryonic digit morphogenesis	0.00323	0.8381	<i>HOXD13, GAS1, HOXD11</i>
GOTERM_BP_FAT	kidney development	0.00422	0.9073	<i>ALDH1A2, ITGA8, ARID5B, HOXD11</i>
GOTERM_BP_FAT	chordate embryonic development	0.00513	0.9446	<i>HOXD9, HOXC6, ALDH1A2, GRN, GAS1, HOXD10</i>
GOTERM_BP_FAT	embryonic development ending in birth or egg hatching	0.00532	0.9505	<i>HOXD9, HOXC6, ALDH1A2, GRN, GAS1, HOXD10</i>
SMART	SM00389:HOX	0.00550	0.1754	<i>HOXD9, HOXC6, HOXD13, HOXD10, HOXD11</i>
SP_PIR_KEYWORDS	Homeobox	0.00573	0.5423	<i>HOXD9, HOXC6, HOXD13, HOXD10, HOXD11</i>
INTERPRO	IPR017970:Homeobox, conserved site	0.00717	0.6454	<i>HOXD9, HOXC6, HOXD13, HOXD10, HOXD11</i>
INTERPRO	IPR001356:Homeobox	0.00750	0.6618	<i>HOXD9, HOXC6, HOXD13, HOXD10, HOXD11</i>
INTERPRO	IPR012287:Homeodomain-related	0.00784	0.6779	<i>HOXD9, HOXC6, HOXD13, HOXD10, HOXD11</i>
SP_PIR_KEYWORDS	developmental protein	0.00791	0.6603	<i>HOXD9, HOXC6, WNT16, ITGA8, HOXD13, FRZB, HOXD10, HOXD11</i>
GOTERM_BP_FAT	embryonic organ morphogenesis	0.01036	0.9972	<i>HOXD9, ITGA8, GAS1, HOXD10</i>
UP_SEQ_FEATURE	compositionally biased region:Poly-Gly	0.01042	0.8856	<i>HOXD9, HOXC6, GAS1, HOXD11, NR2F1</i>
GOTERM_BP_FAT	embryonic skeletal system morphogenesis	0.01603	0.9999	<i>HOXD9, GAS1, HOXD10</i>
GOTERM_MF_FAT	sequence-specific DNA binding	0.02145	0.9732	<i>HOXD9, HOXC6, HOXD13, FOXB1, HOXD10, HOXD11, NR2F1</i>
GOTERM_BP_FAT	regulation of transcription from RNA polymerase II promoter	0.03543	1.0000	<i>HOXD9, HOXC6, ATN1, HOXD13, ID3, HOXD10, NR2F1</i>
SP_PIR_KEYWORDS	dna-binding	0.04618	0.9984	<i>HOXD9, HOXC6, H1F0, STRA13, ARID5B, HOXD13, FOXB1, HOXD10, MCM5, HOXD11, NR2F1</i>

Annotation Cluster 2				
Category	Term	PValue	Bonferroni	Genes
SP_PIR_KEYWORDS	nad	0.01853	0.9214	<i>HSD17B10, ALDH1A2, NDUFB10, IMPDH2</i>
GOTERM_BP_FAT	oxidation reduction	0.02036	1.0000	<i>HSD17B10, RNLS, VAT1L, ALDH1A2, NDUFB10, SURF1, IMPDH2</i>
UP_SEQ_FEATURE	nucleotide phosphate-binding region:NAD	0.02163	0.9892	<i>HSD17B10, ALDH1A2, IMPDH2</i>
SP_PIR_KEYWORDS	oxidoreductase	0.02519	0.9689	<i>HSD17B10, RNLS, VAT1L, ALDH1A2, NDUFB10, IMPDH2</i>

Annotation Cluster 3				
Category	Term	PValue	Bonferroni	Genes
GOTERM_BP_FAT	skeletal muscle tissue development	0.02114	1.0000	<i>MYL6, HOXD9, HOXD10</i>
GOTERM_BP_FAT	skeletal muscle organ development	0.02114	1.0000	<i>MYL6, HOXD9, HOXD10</i>
GOTERM_BP_FAT	muscle organ development	0.03476	1.0000	<i>MYL6, HOXD9, ARID5B, HOXD10</i>

Annotation Cluster 4				
Category	Term	PValue	Bonferroni	Genes
GOTERM_BP_FAT	positive regulation of epithelial cell proliferation	0.00853	0.9919	<i>ATP6AP1, GRN, GAS1</i>
GOTERM_BP_FAT	regulation of epithelial cell proliferation	0.02423	1.0000	<i>ATP6AP1, GRN, GAS1</i>

Annotation Cluster 5				
Category	Term	PValue	Bonferroni	Genes
KEGG_PATHWAY	hsa04142:Lysosome	0.01017	0.3356	<i>PSAP, ATP6AP1, GALNS, NEU1</i>
GOTERM_CC_FAT	vacuole	0.04111	0.9952	<i>PSAP, ATP6AP1, GALNS, NEU1</i>
SP_PIR_KEYWORDS	lysosome	0.07217	1.0000	<i>PSAP, GALNS, NEU1</i>

Annotation Cluster 6				
Category	Term	PValue	Bonferroni	Genes
SP_PIR_KEYWORDS	signal	0.01100	0.7779	<i>RNLS, WNT16, CRELD2, PSAP, ATP6AP1, CNPY2, FSTL1, GAS1, FRZB, NTS, ITGA8, GRN, GALNS, SEMA3C, NEU1, SSR4, SSR2, PRSS12</i>
UP_SEQ_FEATURE	signal peptide	0.01170	0.9126	<i>RNLS, WNT16, CRELD2, PSAP, ATP6AP1, CNPY2, FSTL1, GAS1, FRZB, NTS, ITGA8, GRN, GALNS, SEMA3C, NEU1, SSR4, SSR2, PRSS12</i>

Annotation Cluster 7				
Category	Term	PValue	Bonferroni	Genes
GOTERM_BP_FAT	oxidation reduction	0.02036	1.0000	<i>HSD17B10, RNLS, VAT1L, ALDH1A2, NDUFB10, SURF1, IMPDH2</i>

表 14. Q88 Dox (+) vs Dox (-) で down-regulate された 1<sup>st</sup> step-DEGs (edgeR による 61 遺伝子)

Annotation Cluster 1

Category	Term	PValue	Benjamini	Genes
GOTERM_CC_FAT	large ribosomal subunit	0.002	0.252	<i>RPL6, RPL34, RPL7L1, RPL24</i>
SP_PIR_KEYWORDS	ribosomal protein	0.002	0.162	<i>RPL6, RPL34, RPL7L1, RPL24, RPL22L1</i>
GOTERM_CC_FAT	cytosolic part	0.003	0.167	<i>CCT7, CCT4, RPL6, RPL34, RPL24</i>
GOTERM_MF_FAT	structural constituent of ribosome	0.004	0.497	<i>RPL6, RPL34, RPL7L1, RPL24, RPL22L1</i>
KEGG_PATHWAY	hsa03010:Ribosome	0.004	0.156	<i>RPL6, RPL34, RPL24, RPL22L1</i>
GOTERM_BP_FAT	translation	0.007	0.958	<i>EIF4G3, RPL6, RPL34, RPL7L1, RPL24, RPL22L1</i>
SP_PIR_KEYWORDS	ribonucleoprotein	0.009	0.301	<i>RPL6, RPL34, RPL7L1, RPL24, RPL22L1</i>
GOTERM_CC_FAT	ribosome	0.009	0.340	<i>RPL6, RPL34, RPL7L1, RPL24, RPL22L1</i>
GOTERM_MF_FAT	structural molecule activity	0.009	0.582	<i>RPL6, RPL34, KAL1, RPL7L1, ANXA1, NUP62CL, RPL24, RPL22L1</i>
GOTERM_CC_FAT	cytosolic large ribosomal subunit	0.009	0.271	<i>RPL6, RPL34, RPL24</i>
GOTERM_CC_FAT	ribosomal subunit	0.013	0.298	<i>RPL6, RPL34, RPL7L1, RPL24</i>
GOTERM_CC_FAT	cytosolic ribosome	0.038	0.589	<i>RPL6, RPL34, RPL24</i>

Annotation Cluster 2

Category	Term	PValue	Benjamini	Genes
GOTERM_BP_FAT	glucose metabolic process	0.018	0.986	<i>PDK1, GBE1, PYGL, PDK3</i>
SP_PIR_KEYWORDS	carbohydrate metabolism	0.021	0.383	<i>PDK1, PYGL, PDK3</i>
GOTERM_BP_FAT	hexose metabolic process	0.032	0.994	<i>PDK1, GBE1, PYGL, PDK3</i>
GOTERM_BP_FAT	monosaccharide metabolic process	0.046	0.996	<i>PDK1, GBE1, PYGL, PDK3</i>
SP_PIR_KEYWORDS	transferase	0.048	0.585	<i>PDK1, RPS6KA3, GBE1, PYGL, CSNK1E, PDK3, GK, SMS, PIGA</i>

Annotation Cluster 3

Category	Term	PValue	Benjamini	Genes
SP_PIR_KEYWORDS	phosphotransferase	0.021	0.346	<i>PDK1, RPS6KA3, PDK3, GK</i>
SP_PIR_KEYWORDS	transferase	0.048	0.585	<i>PDK1, RPS6KA3, GBE1, PYGL, CSNK1E, PDK3, GK, SMS, PIGA</i>
SP_PIR_KEYWORDS	nucleotide-binding	0.053	0.589	<i>PDK1, CCT7, RPS6KA3, CCT4, RAB9A, PYGL, CSNK1E, PDK3, GK, RHOU</i>

### (3) qRT-PCR による validation

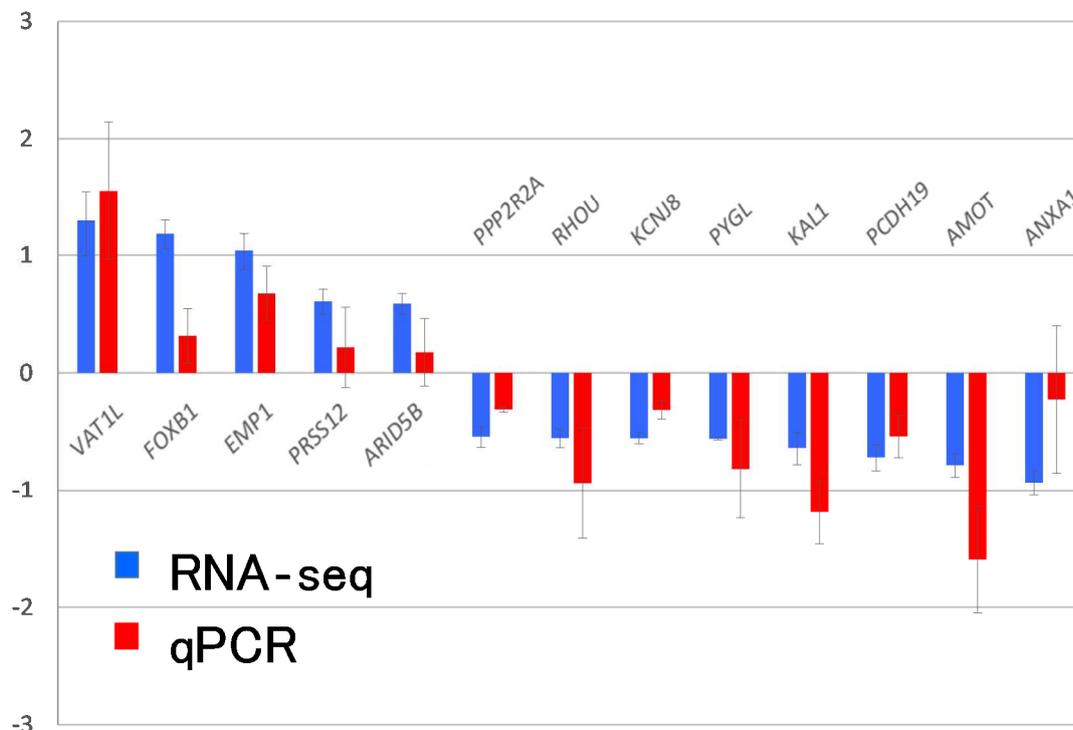
Q88 Dox (+) vs Dox (-) の 2<sup>nd</sup> step-DEGs のうち 13 遺伝子に対して qRT-PCR による発現変動の確認を行った。13 遺伝子の gene symbol と probe の Assay ID を示した (表 15)。解析した 13 遺伝子のうち、*PRSS12*、*ARID5B*、*ANXA1* を除く 10 遺伝子は、RNA-seq と qRT-PCR は発現変動の方向が同じであり validate されたと考えた。*PRSS12*、*ARID5B*、*ANXA1* の 3 遺伝子は、RNA-seq の発現変動と、qRT-PCR による発現変動の平均値は同一の方向性を示したが、qRT-PCR においては 3 回の実験で毎回同一方向という結果が得られなかったこともあり、今回の検討では validate されたとは考えなかった (図 21)。qRT-PCR により validate された発現変動遺伝子 10 遺伝子とそのディスクリプションを示した (表 16)。

表 15. qRT-PCR による発現変動の確認を行った 2<sup>nd</sup> step-DEGs 13 遺伝子と probe Assay ID

Gene symbol	Assay ID
<i>ARID5B</i>	Hs01382781_m1
<i>EMP1</i>	Hs00608055_m1
<i>FOXB1</i>	Hs00247213_s1
<i>PRSS12</i>	Hs00186221_m1
<i>VAT1L</i>	Hs00326103_m1
<i>AMOT</i>	Hs00611096_m1
<i>ANXA1</i>	Hs00167549_m1
<i>KAL1</i>	Hs01085107_m1
<i>KCNJ8</i>	Hs00958961_m1
<i>PCDH19</i>	Hs00403382_m1
<i>PPP2R2A</i>	Hs00953658_m1
<i>PYGL</i>	Hs00958087_m1
<i>RHOA</i>	Hs00221873_m1

図 21. 2<sup>nd</sup> step-DEGs のうち 13 遺伝子に対する qRT-PCR による発現変動の確認

$\log_2$  (fold change)



縦軸は Q88 Dox (+) vs Dox (-) の相対的発現量比の底を 2 とする対数軸である。RNA-seq における相対的発現量比は、set 毎の FPKM 値の、(Dox (+)の発現量の平均) / (Dox (-)の発現量の平均) の 3 set における平均値であり、エラーバーは (Dox (+) の発現量の平均) / (Dox (-) の発現量の平均) の 3 set における標準誤差より算出した。qRT-PCR における相対的発現量比は 3 回の実験の  $-\Delta\Delta Ct$  の平均から算出した ( $\log_2 (2^{-\Delta\Delta Ct \text{ ave.}})$ )。qRT-PCR のエラーバーは  $-\Delta\Delta Ct$  の標準誤差である。ここに挙げた 13 遺伝子のうち、*PRSS12*、*ARID5B*、*ANXA1* を除く 10 遺伝子は、RNA-seq と qRT-PCR は発現変動の方向が同じであり validate されたと考えた。*PRSS12*、*ARID5B*、*ANXA1* の 3 遺伝子は、3 回中 2 回 RNA-seq と同じ方向に発現変動が観察された。

表 16. qRT-PCR により validate された 10 遺伝子とそのディスクリプション

Gene symbol	Description
<b>up-regulated 3遺伝子</b>	
<i>EMP1</i>	epithelial membrane protein 1
<i>FOXB1</i>	forkhead box B1
<i>VAT1L</i>	vesicle amine transport 1-like
<b>down-regulated 7遺伝子</b>	
<i>AMOT</i>	angiominin
<i>KAL1</i>	Kallmann syndrome 1 sequence
<i>KCNJ8</i>	potassium channel, inwardly rectifying subfamily J, member 8
<i>PCDH19</i>	protocadherin 19
<i>PPP2R2A</i>	protein phosphatase 2 regulatory subunit B, alpha
<i>PYGL</i>	phosphorylase, glycogen, liver
<i>RHOU</i>	ras homolog family member U

## 第 2 章 考察

本研究では、DRPLA の転写の co-regulator としての標的遺伝子を同定することを目的に、*GFP-DRPLA* 遺伝子定常発現 HEK293 細胞を用い、RNA-seq による網羅的な解析を行った。実験間における再現性を最優先し、qRT-PCR において validate された遺伝子として、wild type DRPLAp (Q19) の強制発現により 11 個の遺伝子を見いだすことができた。同様の方法で、mutant DRPLAp (Q88) の強制発現により 10 個の遺伝子を見いだした。

本研究においては、HEK293 細胞由来の Flp-In™ T-REx™ 293 細胞を用いて、発現変動遺伝子を検討した。剖検脳やモデルマウスの脳を用いた解析では検体間のばらつきによる再現性を得ることの難しさ、混在した細胞種の homogenate を用いた研究にならざるを得ないことなどの問題点があり、また RNA-seq 自体の手法の頑健性の問題については検討する必要があると考えられたことから、今回は培養細胞を用いて検討する方針とした。ライン間のばらつきを取り除くため、Flp-In™ T-REx™ 293 細胞の FRT 部位に Q19 もしくは Q88 の全長 *GFP-DRPLA* 遺伝子を導入した定常発現細胞を構築したことにより、一定の条件下で Q19 もしくは Q88 を定常発現させて RNA-seq を検討することが可能となった。

RNA-seq の再現性という点で、同一条件で培養した、同一セットに属する 3 サンプルのシーケンスによる遺伝子の発現量を検討したところ、Spearman's correlation coefficient が 0.98 以上となる、高い再現性を得られることが判明した (図 16)。

発現変動遺伝子の抽出にあたり、今回 Cufflinks 2 と edgeR という 2 つのソフトウェアを用いて解析を行い、最終的に両方で有意に変動する遺伝子として得られたものを選択した。Cufflinks 2 と edgeR は各々遺伝子の発現量として FPKM 値、CPM 値を用いており、おのおの算出方法が異なるが、Zhang らは同一 RNA サンプルを用いた RNA-seq 解析で両者のパイプラインを比較しており、発現変動遺伝子は edgeR でより多く検出されることを報告している<sup>[64]</sup>。今回の解析でも同様に、edgeRの方が Cufflinks 2 より多く、有意に変動する発現変動遺伝子を検出する結果となった。

1<sup>st</sup> step-DEGs に対しては、Gene Ontology 解析を行った。Q19、Q88 において up-regulate された 1<sup>st</sup> step-DEGs については、共通して、骨格の発生に関する GO term が観察された (表 8、表 12、表 13)。これについては、1<sup>st</sup> step-DEGs において共通してホメオボックス遺伝子 (*HOXC6*, *HOXD9*, *HOXD11*) が含まれたためと考えられた。

今回、Cufflinks2 と edgeR の両方のパイプラインで、Q19 では 4 set、Q88 では 3 set で常に有意に変動する発現変動遺伝子を抽出した (2<sup>nd</sup> step-DEGs)。このように再現性に重点を置いて抽出された遺伝子については、qRT-PCR の解析を行った。qRT-PCR 解析で算出された fold change の平均値については、解析した 24 遺伝子全てで RNA-seq と同一の方向であることが示され、中でも 21 遺伝子については 3 組の RNA サンプルで一貫して同一方向へ変動していることが確認されたため、本研究ではこの 21 遺伝子を validate されたと考えた。また、この 21 遺伝子のうち、20 遺伝子については、RNA-seq で算出された fold change と qRT-PCR で算出された fold change の差が 2 倍以下に収まってお

り、RNA-seq の定量性を支持するものと考えた。Q19 の解析における *AMOT* 遺伝子の  
み、RNA-seq と qRT-PCR のデータが大きくずれたが、発現量が少ない (FPKM 値  
<10) ことが影響していたのではないかと考えられた。

次に、DRPLAp (Q19) の強制発現により発現変動した遺伝子について述べる。今回の  
検討から、11 個の遺伝子を見だし、そのうち 4 個は up-regulate、7 個は down-  
regulate されていることが判明した。なお、序章で述べた通り、赤岩らは、wild type マ  
ウスに比して DRPLA ノックアウトマウスの全脳で発現変動する遺伝子を見だしている  
が、今回の DRPLAp (Q19) の強制発現により発現変動する遺伝子として見いだされた 11  
個の遺伝子とはオーバーラップは認めなかったが、発現変動遺伝子の抽出に関する統計学  
的な条件が異なっていること、マウスの脳の homogenate とヒト由来培養細胞と種や細胞  
種が異なること、*DRPLA* 遺伝子の減少による変化と強制発現による変化という、比較対  
象の違いによるものと考えられた。将来的に、統計学的な解析条件を一致させた解析をす  
ることやヒト由来培養細胞の系においても *DRPLA* 遺伝子のノックアウト もしくはノッ  
クダウンによる発現変動遺伝子を検討することは有意義と考えられる。

今回 DRPLAp (Q19) の強制発現により発現変動した 11 遺伝子が、病態にどのように  
関わっているかについては、今後の検討が必要であるが、DRPLAp の生理機能について  
は不明な点が多く、これらの遺伝子は非常に興味深いと言える。今回の結果を基に、モデ  
ルマウスやヒト剖検脳を用いた分析を行っていく必要があると考えられる。

DRPLAp (Q88) の強制発現により、本研究では 10 個の発現変動遺伝子を見いだした。

この 10 個の遺伝子から、脳での発現量が GTE<sub>x</sub> (Genotype-Tissue Expression) データベース<sup>[65]</sup> 上で RPKM (Reads Per Kilobase of transcript per Million mapped reads) 値が 1 以上となり、また FPKM 値より算出した遺伝子発現量比 (fold change) が、 $|\log_2(\text{fold change})| \geq 0.5$  となる遺伝子を抽出したところ、5 個の遺伝子が得られた。内訳は、up-regulation された遺伝子が *VAT1L*, *EMP1* の 2 個、down-regulation された遺伝子が *PPP2R2A*, *KCNJ8*, *PCDH19* の 3 個である。

このうち、*PCDH19* はカルシウム依存性細胞間接着分子であるプロトカドヘリン 19 をコードし、EFMR (epilepsy and mental retardation restricted to females) の原因遺伝子とされている<sup>[66]</sup>。臨床像は heterogenous であるが、生後 6 ヶ月から 36 ヶ月の女兒の発熱を契機としたてんかん発作の群発と精神発達遅滞を主徴とし、自閉症を伴う場合がある。変異の約半数は、ナンセンス変異やフレームシフト変異などの機能喪失性変異であり、残りの半数のミスセンス変異も、プロトカドヘリン 19 のオルソログやパラログとも相同性の高い細胞外ドメインに存在することがほとんどとされている<sup>[67]</sup>。これらの知見からは、EFMR は *PCDH19* の loss-of-function による疾患であることが強く示唆される。*KCNJ8* は内向き整流カリウムチャネル ( $K_{ATP}$ ) サブユニット Kir6.1 をコードする。*KCNJ8* 遺伝子変異が Brugada syndrome、early repolarization syndrome といった突然死の原因となる疾患において、疾患感受性遺伝子であることが示唆されている<sup>[68]</sup>。<sup>69]</sup> これらの疾患では中枢神経症状は指摘されていないが、*KCNJ8* は中枢神経における

発現は確認されている。DRPLA は、特に幼少時～若年成人発症においてはてんかんを合併することが多いが、てんかんを起こす機序については不明な点が多い。本研究から、Flp-In™ T-REx™ 293 細胞において mutant DRPLA<sub>p</sub> の強制発現により、てんかんを起こしうる病因遺伝子の一つである *PCDH19* と、カリウムチャネルのサブユニットである *KCNJ8* の発現変動が認められるという結果が得られたことは、てんかんの病態を考える上で興味深い知見と言えるかもしれない。

なお、Q19 と Q88 の各 set における Dox (+) vs Dox (-) の比較における発現変動遺伝子数については、set 毎のばらつきが観察された。Tet-On system を RNA-seq により定量評価した報告はないが、この理由として、細胞培養時の confluency など、完全にはコントロールしきれない培養条件による違いなどが影響している可能性がある。また、発現量の差の検定の際に、replicate 数が 3 と多くはないことから、母分散がたまたま小さく見積もられた時、検定において帰無仮説が棄却される遺伝子数が増えるといったことが影響している可能性も考えられた。そのため、本研究においては、Q19 は 4 set、Q88 は 3 set において共通する再現性の高い発現変動遺伝子を抽出する方針としたが、qRT-PCR によって大部分が validate されたことから、今回の方針は真の発現変動遺伝子を抽出するための一つの手法であると考えられた。

最後に、polyQ 鎖長の伸長に伴う発現変動遺伝子の探索のため、Q88 Dox (+) と Q19 Dox (+) の比較を、前述と同様に分析した。比較可能な、set 1, 3, 4 において共通して発現変動する遺伝子 (1<sup>st</sup> step-DEGs) を抽出した後、Cufflinks 2 と edgeR で共通して発現変

動するとされる遺伝子を抽出したところ、再現性を持って有意に変動する遺伝子は認めないという結果となった。実際両者に、今回の実験回数や、再現性に重視した解析方法では有意となる程の大きな差がない可能性も考えられるが、そもそも今回の RNA-seq の解析では、Q88 Dox (+) と Q19 Dox (+) の *DRPLA* 遺伝子の発現量に差があること (Set 1 では Cufflinks 2 と edgeR の両者で、set 3 と set 4 では Cufflinks 2 で  $q < 0.05$ ) も観察されており、このように単純に比較することが、polyQ 鎖長の伸長に伴う発現変動遺伝子を探索するには適切な設定ではない可能性も考えられ、結果については慎重な解釈が必要と思われた。さらに、ポリグルタミン病の病態機序においては、伸長ポリグルタミン鎖を有する変異タンパクの核内集積が重要であると考えられている。本実験は、24 時間という短時間の遺伝子発現の変化に焦点を絞って解析していることから、ポリグルタミン病の病態機序を解析するためには、伸長ポリグルタミン鎖を有する変異タンパクが十分に核内集積するような実験系を用いる必要があると考えられる。DRPLAp の生理機能という点では、本実験条件の範囲においては、Q19, Q88 のポリグルタミン鎖の鎖長の違いは大きな変化を示さないと考えられる。

本研究における 1 つの limitation として、今回の検討では、wild type DRPLAp、mutant DRPLAp を強制発現させた時に発現変動する遺伝子を同定したが、これらの遺伝子転写制御に直接 DRPLAp が影響を及ぼしていることまでは示せていないことが挙げられる。TAF II 130 以外に DRPLAp と共役する転写因子を同定し、共役する転写因子に対

する抗体を用いたクロマチン免疫沈降 - 大量並列シーケンシング (chromatin-immunoprecipitation followed by massively parallel sequencing (ChIP-seq))<sup>[70]</sup>により、共役因子がクロマチンに結合する部位と、今回見いだされた遺伝子との関係を検討し、DRPLAp が直接制御しているのか否かを解明することが今後の課題と考えられる。もう一つは、ヒト脳における DRPLAp の生理的機能、伸長 polyQ による病態生理について今後さらに検討を行う必要があることである。RNA-seq の解析手法については、今回の検討からその定量性が示されたこともあり、今後ヒト脳組織を用いた RNA-seq を行うことで発現変動遺伝子を探索することが可能であろう。

### 第3章 結語と今後の課題

DRPLAの病態を考える上で、*DRPLA* 遺伝子の機能について検討することが必要であるが未解明の点が多かった。これまで、DRPLApは転写 co-regulator として働くことが示唆されているものの、DRPLApの生理的な標的遺伝子については明らかではなかった。そこで、DRPLApの転写 co-regulator としての標的遺伝子を同定することを目的とし、再現性に重点をおいてRNA-seqを行ったところ、wild type DRPLAp (Q19)の強制発現により、11個の発現変動遺伝子を明らかにした。また、mutant DRPLAp (Q88)の強制発現により、10個の発現変動遺伝子を明らかにした。この中にはヒトでてんかんを起こす病因遺伝子が含まれており、DRPLAの病態を考える上で興味深い知見と思われた。

今後の課題は、DRPLAの病態への関与についての検討である。今回同定した21個の発現変動遺伝子が、実際にDRPLApの転写 co-regulator としての標的遺伝子であることを確認すること、polyQ鎖長の異常伸長による発現変動遺伝子の変化を、解析方法の検討により探索すること、本研究で見いだされた21個の発現変動遺伝子の*in vivo*における検証を行うことが必要である。今回見出した21遺伝子が病態に関わっていることが判明すれば、治療介入のシーズになる可能性があり、注目すべきである。

## 謝辞

本稿を終えるにあたり、東京大学大学院医学系研究科脳神経医学専攻神経内科学教室  
辻 省次 教授、石浦浩之 助教、伊達英俊 学術支援専門職員、三井純 助教、田中真生 特  
任臨床医、国際医療福祉大学三田病院神経内科 後藤 順 教授、東京大学大学院新領域学  
術創成科学研究科情報生命科学専攻 森下真一 教授、土井晃一郎 特任講師、吉村淳 特  
任助教をはじめとする教室員各位の皆様に深謝申し上げます。

実験の補助をして頂きました竹山未央様、平山圭子様には深謝申し上げます。

## 参考文献

1. Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, Kondo R, Ishikawa A, Hayashi T, Saito M, Tomoda A, Miike T, Naito H, Ikuta F, Tsuji S, Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nature Genetics*, 6, 9-13, 1994
2. Nagahuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, Bundo M, Takeda T, Tadokoro K, Kondo I, Murayama N, Tanaka Y, Kikushima H, Umino K, Kurosawa H, Furukawa T, Nihei K, Inoue T, Sano A, Komure O, Takahashi M, Yoshizawa T, Kanazawa I, Yamada M, Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nature Genetics*, 6, 14-18, 1994
3. La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH, Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*, 352, 77-79, 1991
4. The Huntington's Disease Collaborative Research Group, A Novel Gene Containing a Trinucleotide Repeat That is Expanded and Unstable on Huntington's Disease Chromosomes. *Cell*, 72, 971-963, 1993
5. Orr HT, Chung MY, Banfi S, Kwiatkowski TJ Jr, Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LP, Zoghbi HY, Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nature Genetics*, 4, 221-226, 1993
6. Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M, Akiguchi I, Kimura J, Narumiya S, Kakizuka A, CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nature Genetics*, 8, 221-228, 1994

7. Imbert G, Saudou F, Yvert G, Devys D, Trottier Y, Garnier JM, Weber C, Mandel JL, Cancel G, Abbas N, Dürr A, Didierjean O, Stevanin G, Agid Y, Brice A, Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nature Genetics*, 14, 285-291, 1996
8. Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunkes A, DeJong P, Rouleau GA, Auburger G, Korenberg JR, Figueroa C, Sahba S, Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nature Genetics*, 14, 269-276, 1996
9. Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, Wakisaka A, Tashiro K, Ishida Y, Ikeuchi T, Koide R, Saito M, Sato A, Tanaka T, Hanyu S, Takiyama Y, Nishizawa M, Shimizu N, Nomura Y, Segawa M, Iwabuchi K, Eguchi I, Tanaka H, Takahashi H, Tsuji S, Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nature Genetics*, 14, 277-284, 1996
10. Zhuchenko O, Bailey J, Bonnen P, Ashizawa T, Stockton DW, Amos C, Dobyns WB, Subramony SH, Zoghbi HY, Lee CC, Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nature Genetics*, 15, 62-69, 1997
11. David G, Abbas N, Stevanin G, Dürr A, Yvert G, Cancel G, Weber C, Imbert G, Saudou F, Antoniou E, Drabkin H, Gemmill R, Giunti P, Benomar A, Wood N, Ruberg M, Agid Y, Mandel JL, Brice A, Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nature Genetics*, 17, 65-70, 1997
12. Koide R, Kobayashi S, Shimohata T, Ikeuchi T, Maruyama M, Saito M, Yamada M, Takahashi H, Tsuji S, A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? *Hum Mol Genet*, 8, 2047-2053, 1999

13. Nakamura K, Jeong SY, Uchihara T, Anno M, Nagashima K, Nagashima T, Ikeda S, Tsuji S, Kanazawa I, SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet*, 10, 1441-1448, 2001
14. Naito H, Oyanagi S, Familial myoclonus epilepsy and choreoathetosis: hereditary dentatorubral-pallidolusian atrophy. *Neurology*, 32, 798-807, 1982
15. Tsuji S, Onodera O, Kanazawa I, Mizusawa H, Hattori T, Sobue G, Yamada M, Kuroiwa Y, Kakizuka A, Takeda A, Hasegawa K, Yuasa T, Kanda T, Yoshida K, Kachi T, Nakajima T, Osame M, Kuno S, Nakajima K, Kawakami H, Nagai Y, Nishizawa M, Sporadic ataxias in Japan--a population-based epidemiological study. *Cerebellum*, 7, 189-197, 2008
16. Farmer TW, Wingfield MS, Lynch SA, Vogel FS, Hulette C, Katchinoff B, Jacobson PL, Ataxia, chorea, seizures, and dementia: pathologic features of a newly defined familial disorder. *Arch Neurol*, 46, 774-779, 1989
17. Burke JR, Wingfield MS, Lewis KE, Roses AD, Lee JE, Hulette C, Pericak-Vance MA, Vance JM, The Haw River syndrome: dentatorubropallidolusian atrophy (DRPLA) in an African-American family. *Nature Genetics*, 7, 521-524, 1994
18. Warner T T, Williams L, Harding AE, DRPLA in Europe. (Letter) *Nature Genet*, 6, 225, 1994
19. Connarty M, Dennis N R, Patch C, Macpherson JN, Harvey JF, Molecular re-investigation of patients with Huntington's disease in Wessex reveals a family with dentatorubral and pallidolusian atrophy. *Hum Genet*, 97, 76-78, 1996
20. Takano H, Onodera O, Takahashi H, Igarashi S, Yamada M, Oyake M, Ikeuchi T, Koide R, Tanaka H, Iwabuchi K, Tsuji S, Somatic Mosaicism of Expanded

CAG Repeats in Brains of Patients with Dentatorubral-Pallidoluysian Atrophy: Cellular Population-Dependent Dynamics of Mitotic Instability. *Am J Hum Genet*, 58, 1212-1222, 1996

21. Kim HJ, Jeon BS, Lee WY, Chung SJ, Yong SW, Kang JH, Lee SH, Park KW, Park MY, Kim BC, Kim JW, Kim HT, Ha CK, Koh SB, Kim JM, Choi KD, Sung YH, Ahn TB, Lee GH, Lee JH, Lee HW, Kim SJ, Park JH, Kwon DY, Kim MJ, Kim YJ, Kim JS, Cho J, Kwon JH, Kim EJ, Kim JH, Sung KB, Song IU, Oh HG, Lee SB, Lee SH, Lee JY, Lee TK, Cho AH, Yoon WT, Kim SR, Kim HJ, SCA in Korea and its regional distribution: a multicenter analysis. *Parkinsonism Relat Disord*, 17, 72-75, 2011
22. Jin DK, Oh MR, Song SM, Koh SW, Lee M, Kim GM, Lee WY, Chung CS, Lee KH, Im JH, Lee MJ, Kim JW, Lee MS, Frequency of spinocerebellar ataxia types 1, 2, 3, 6, 7 and dentatorubral pallidoluysian atrophy mutations in Korean patients with spinocerebellar ataxia. *J Neurol*, 246, 207-210, 1999
23. Silveira I, Miranda C, Guimarães L, Moreira MC, Alonso I, Mendonça P, Ferro A, Pinto-Basto J, Coelho J, Ferreirinha F, Poirier J, Parreira E, Vale J, Januário C, Barbot C, Tuna A, Barros J, Koide R, Tsuji S, Holmes SE, Margolis RL, Jardim L, Pandolfo M, Coutinho P, Sequeiros J, Trinucleotide repeats in 202 families with ataxia: a small expanded (CAG)<sub>n</sub> allele at the SCA17 locus. *Arch Neurol*, 59, 623-629, 2002
24. Martins S, Matama T, Guimara L, Vale J, Guimarães J, Ramos L, Coutinho P, Sequeiros J, Silveira I, Portuguese families with dentatorubropallidoluysian atrophy (DRPLA) share a common haplotype of Asian origin. *European Journal of Human Genetics*, 11, 808-811, 2003
25. Nørremølle A, Nielsen JE, Sørensen SA, Hasholt L, Elongated CAG repeats of the B37 gene in a Danish family with dentato-rubro-pallido-luysian atrophy. *Hum Genet*, 95, 313-318, 1995

26. Brusco A, Gellera C, Cagnoli C, Saluto A, Castucci A, Michielotto C, Fetoni V, Mariotti C, Migone N, Di Donato S, Taroni F, Molecular Genetics of Hereditary Spinocerebellar Ataxia: Mutation Analysis of Spinocerebellar Ataxia Genes and CAG/CTG Repeat Expansion Detection in 225 Italian Families. *Arch Neurol*, 61, 727-733, 2004
27. Le Ber I, Camuzat A, Castelnovo G, Azulay JP, Genton P, Gastaut JL, Broglin D, Labauge P, Brice A, Durr A, Prevalence of dentatorubral-pallidoluysian atrophy in a large series of white patients with cerebellar ataxia. *Arch Neurol* 60, 1097-1099, 2003
28. Filla A, Mariotti C, Caruso G, Coppola G, Coccozza S, Castaldo I, Calabrese O, Salvatore E, De Michele G, Riggio MC, Pareyson D, Gellera C, Di Donato S, Relative frequencies of CAG expansions in spinocerebellar ataxia and dentatorubropallidoluysian atrophy in 116 Italian families. *Eur Neurol*, 44, 31-36, 2000
29. Vinton A, Fahey MC, O'Brien TJ, Shaw J, Storey E, Gardner RJ, Mitchell PJ, Du Sart D, King JO, Dentatorubral-pallidoluysian atrophy in three generations, with clinical courses from nearly asymptomatic elderly to severe juvenile, in an Australian family of Macedonian descent. *Am J Med Genet A*, 136, 201-204, 2005
30. Paradisi I, Ikonomu V, Arias S, Spinocerebellar ataxias in Venezuela: genetic epidemiology and their most likely ethnic descent. *Journal of Human Genetics*, 1-8, 2015
31. Zádori D, Tánzos T, Jakab K, Vécsei L, Klivényi P, The first identified Central-Eastern European patient with genetically confirmed dentatorubral-pallidoluysian atrophy. *Ideggyogy Sz.* 68, 68-71, 2015
32. Tang B, Liu C, Shen L, Dai H, Pan Q, Jing L, Ouyang S, Xia J, Frequency of SCA1, SCA2, SCA3/MJD, SCA6, SCA7, and DRPLA CAG trinucleotide repeat

expansion in patients with hereditary spinocerebellar ataxia from Chinese kindreds. *Arch Neurol*, 57, 540-544, 2000

33. Sura T, Eu-ahsunthornwattana J, Youngcharoen S, Busabaratana M, Dejsuphong D, Trachoo O, Theerasasawat S, Tunteeratum A, Noparutchanodom C, Tunlayadechanont S, Frequencies of spinocerebellar ataxia subtypes in Thailand: window to the population history? *J Hum Genet*, 54, 284-288, 2009
34. Cintra VP, Lourenço CM, Marques SE, de Oliveira LM, Tumas V, Marques W Jr, Mutational screening of 320 Brazilian patients with autosomal dominant spinocerebellar ataxia. *J Neurol Sci*, 15, 375-379, 2014
35. Dubourg O, Dürr A, Chneiweiss H, Does the ataxo-choreic form of DRPLA exist in Europe? Search of mutation in 120 families. *Rev Neurol (Paris)*, 11, 657-660, 1995
36. Yazawa I, Nukina N, Hashida H, Goto J, Yamada M, Kanazawa I, Abnormal gene product identified in hereditary dentatorubral-pallidolusian atrophy (DRPLA) brain. *Nature Genetics*, 10, 99-103, 1995
37. Yazawa I, Nukina N, Goto J, Kurisaki H, Hebisawa A, Kanazawa I, Expression of dentatorubral-pallidolusian atrophy (DRPLA) proteins in patients. *Neurosci Lett*, 225, 53-56, 1997
38. Knight SP, Richardson MM, Osmand AP, Stakkestad A, Potter NT, Expression and distribution of the dentatorubral-pallidolusian atrophy gene product (atrophin-1/drplap) in neuronal and non-neuronal tissues. *J Neurol Sci*, 146, 19-26, 1997
39. Sato A, Shimohata T, Koide R, Takano H, Sato T, Oyake M, Igarashi S, Tanaka K, Inuzuka T, Nawa H, Tsuji S, Adenovirus-mediated expression of mutant DRPLA proteins with expanded polyglutamine stretches in neuronally

differentiated PC12 cells. Preferential intranuclear aggregate formation and apoptosis. *Hum Mol Genet*, 8, 997-1006, 1999

40. Nucifora FC Jr, Ellerby LM, Wellington CL, Wood JD, Herring WJ, Sawa A, Hayden MR, Dawson VL, Dawson TM, Ross CA, Nuclear Localization of a Non-caspase Truncation Product of Atrophin-1, with an Expanded Polyglutamine Repeat, Increases Cellular Toxicity. *J Biol Chem*, 278, 13047-13055, 2003
41. Onodera O, Oyake M, Takano H, Ikeuchi T, Igarashi S, Tsuji S, Molecular cloning of a full-length cDNA for dentatorubral-pallidoluysian atrophy and regional expressions of the expanded alleles in the CNS. *Am J Hum Genet*, 57, 1050-1060, 1995
42. Yamada M, Wood JD, Shimohata T, Hayashi S, Tsuji S, Ross CA, Takahashi H, Widespread occurrence of intranuclear atrophin-1 accumulation in the central nervous system neurons of patients with dentatorubral-pallidoluysian atrophy. *Ann Neurol*, 49, 14-23, 2001
43. Sato T, Miura M, Yamada M, Yoshida T, Wood JD, Yazawa I, Masuda M, Suzuki T, Shin RM, Yau HJ, Liu FC, Shimohata T, Onodera O, Ross CA, Katsuki M, Takahashi H, Kano M, Aosaki T, Tsuji S, Severe neurological phenotypes of Q129 DRPLA transgenic mice serendipitously created by en masse expansion of CAG repeats in Q76 DRPLA mice. *Hum Mol Genet*, 18, 723-7736, 2009
44. 赤岩靖久, DRPLA ノックアウトマウスにおける遺伝子発現プロファイリング, 新潟医学会雑誌, 119, 672-682, 2005
45. Zhang S, Xu L, Lee J, Xu T, *Drosophila* Atrophin Homolog Functions as a Transcriptional Corepressor in Multiple Developmental Processes. *Cell*, 108, 45-56, 2002

46. Shen Y, Lee G, Choe Y, Zoltewicz JS, Peterson AS, Functional architecture of atrophins. *J Biol Chem*, 282, 5037-5044, 2007
47. Wood JD, Nucifora FC Jr, Duan K, Zhang C, Wang J, Kim Y, Schilling G, Sacchi N, Liu JM, Ross CA, Atrophin-1, the dentato-rubral and pallidoluysian atrophy gene product, interacts with ETO/MTG8 in the nuclear matrix and represses transcription. *J Cell Biol*, 150, 939-948, 2000
48. Shimohata T, Nakajima T, Yamada M, Uchida C, Onodera O, Naruse S, Kimura T, Koide R, Nozaki K, Sano Y, Ishiguro H, Sakoe K, Ooshima T, Sato A, Ikeuchi T, Oyake M, Sato T, Aoyagi Y, Hozumi I, Nagatsu T, Takiyama Y, Nishizawa M, Goto J, Kanazawa I, Davidson I, Tanese N, Takahashi H, Tsuji S, Expanded polyglutamine stretches interact with TAFII130, interfering with CREB-dependent transcription. *Nature Genetics*, 26, 29-36, 2000
49. Felinski EA, Quinn PG. The CREB constitutive activation domain interacts with TATA-binding protein-associated factor 110 (TAF110) through specific hydrophobic residues in one of the three subdomains required for both activation and TAF110 binding. *J Biol Chem*, 274, 11672-11678, 1999
50. Nakajima T, Uchida C, Anderson SF, Parvin JD, Montminy M, Analysis of a cAMP-responsive activator reveals a two-component mechanism for transcriptional induction via signal-dependent factors. *Genes Dev*, 11, 738-47, 1997
51. Shimohata M, Shimohata T, Igarashi S, Naruse S, Tsuji S, Interference of CREB-dependent transcriptional activation by expanded polyglutamine stretches – augmentation of transcriptional activation as a potential therapeutic strategy for polyglutamine diseases. *Journal of Neurochemistry*, 93, 654-663, 2005
52. Suzuki K, Zhou J, Sato T, Takao K, Miyagawa T, Oyake M, Yamada M, Takahashi H, Takahashi Y, Goto J, Tsuji S, DRPLA transgenic mouse

substrains carrying single copy of full-length mutant human DRPLA gene with variable sizes of expanded CAG repeats exhibit CAG repeat length- and age-dependent changes in behavioral abnormalities and gene expression profiles. *Neurobiol Dis*, 46, 336-350, 2012

53. Pinsky L, Kaufman M, Levitsky LL, Partial androgen resistance due to a distinctive qualitative defect of the androgen receptor. *Am J Med Genet*, 27, 459-466, 1987
54. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, 10, 57-63, 2009
55. Gossen M, Bujard H, Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci USA*, 89, 5547-5551, 1992
56. Trapnell C, Pachter L, Salzberg SL, TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, 25, 1105-1111, 2009
57. Langmead B, Salzberg SL, Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357-359, 2012
58. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protoc*, 7, 562-578, 2012
59. Robinson MD, McCarthy DJ, Smyth GK, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 139-140, 2010
60. Anders A, Pyl PT, Huber W, HTSeq — A Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31, 166-169, 2015

61. Li H, Durbin R, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 15, 1754-1760, 2009
62. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G, Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25, 25-29, 2000
63. Huang DW, Sherman BT, Lempicki RA, Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc*, 4, 44-57, 2009
64. Zhang ZH, Jhaveri DJ, Marshall VM, Bauer DC, Edson J, Narayanan RK, Robinson GJ, Lundberg AE, Bartlett PF, Wray NR, Zhao QY, A comparative study of techniques for differential expression analysis on RNA-Seq data. *PLoS One*, 9, e103207, 2014
65. The GTEx Consortium, The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, 45, 580-585, 2013
66. Dibbens LM, Tarpey PS, Hynes K, X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. *Nature Genetics*, 40, 776-781, 2008
67. Depienne C, Trouillard O, Bouteiller D, Gourfinkel-An I, Poirier K, Rivier F, Berquin P, Nabbout R, Chaigne D, Steschenko D, Gautier A, Hoffman-Zacharska D, Lannuzel A, Lackmy-Port-Lis M, Maurey H, Dusser A, Bru M, Gilbert-Dussardier B, Roubertie A, Kaminska A, Whalen S, Mignot C, Baulac S, Lesca G, Arzimanoglou A, LeGuern E, Mutations and deletions in PCDH19 account for various familial or isolated epilepsies in females. *Hum Mutat*, 32, E1959-1975, 2011

68. Barajas-Martínez H, Hu D, Ferrer T, Onetti CG, Wu Y, Burashnikov E, Boyle M, Surman T, Urrutia J, Veltmann C, Schimpf R, Borggreffe M, Wolpert C, Ibrahim BB, Sánchez-Chapula JA, Winters S, Haïssaguerre M, Antzelevitch C, Molecular genetic and functional association of Brugada and early repolarization syndromes with S422L missense mutation in KCNJ8. *Heart Rhythm*, 9, 548-555, 2012
69. Veeramah KR, Karafet TM, Wolf D, Samson RA, Hammer MF, The KCNJ8-S422L variant previously associated with J-wave syndromes is found at an increased frequency in Ashkenazi Jews. *Eur J Hum Genet*, 22, 94-98, 2014
70. Johnson DS, Mortazavi A, Myers RM, Wold B, Genome-wide mapping of in vivo protein-DNA interactions. *Science*, 316, 1497-1502, 2007