琉球大学学術リポジトリ

鳥類のミトコンドリアDNAにおけるエクソンのスペクトル解析と塩基配列部分置換による強度特性

メタデータ	言語:
	出版者: 琉球大学理学部
	公開日: 2012-12-21
	キーワード (Ja):
	キーワード (En):
	作成者: 中村, 圭太, 清野, 光弘, Nakamura, Keita, Seino, Mitsuhiro
	メールアドレス:
	所属:
URL	http://hdl.handle.net/20.500.12000/25535

鳥類のミトコンドリアDNAにおけるエクソンのスペクトル解析と塩基配列部分置換による強度特性

中村圭太　清野光弘 *
琉球大学理学部物質地球科学科物理系

Spectral Analysis for Partial Substitution with Base Sequences of Exons in Mitchondrial DNA of Birds and its Intensities of Spectra.

Keita Nakamura , Mitsuhiro Seino*
Department of Physics and Earth Science, Faculty of Science, University of the Ryukyus, Nishihara, Okinawa 903-0213,Japan

Abstract

Spectrum analysis for sequences of exons and introns with genome sequences for mitchondrial DNA (mtDNA) of 20 birds are discussed. Spectrum peaks of power spectra for exon are come into existence remarkably and for intron are disappeared. Consequently, power spectrum indicated coding DNA has the spectrum of maximum for approximate to 0.33 and 0.66. In this paper,Fourier Spectra with new sequences joined exon codes to intron codes divided into are analyzed.

1 はじめに

系統樹を作る場合, 分子系統樹による推定は大きな役割を果たしており, 鳥類では近年のゲノム解析により大規模な系統樹の改変が行われた [1]. しかし, 分子系統樹の推定は仮定したモデルに依存するため, モデルが現実的でない場合, 偏った系統樹推定が行われる可能性がある [2]. このモデルはDNAの塩基置換に関するもので, 分子進化の中立説などさまざまな仮説を取り入れて提案される. より信頼性が高く精度の良いモデルを提案することは, 分子系統樹の推定において重要なことである. そこで, 物理的な手法でDNA塩基置換のメカニズムを調べるため, スペクトル解析を行った.DNAにおけるスペクトル解析はH.E.StanleyやR.F.Voss,W.Li等によって研究され [3,4,5,6], 多くの報告がなされている [7,8,9,10,11].DNA塩基配列の完全長に関する統計学的解釈を一種の信号理論的概念に置き換えて, フーリエ解析によりスペクトル解析を行う. このとき, パワースペクトルから, 塩基配列のLocation間での長距離相関の強さが決定できる. 本研究では, 鳥類のミトコンドリアDNA(mtDNA)のエクソンおよびイントロンにおける様相について, パワースペクトルのある特定の周波数に分布して最大となるような強度を調べる.mtDNAは核DNAに比べ塩基置換が頻繁に起こることが知られており, 種による差異が生じやすい. その中で, どの種にも共通するメカニズムを見い出すことで, 新たな仮説の提案につながり, より信頼性の高い系統樹推定につながる可能性があると期待される.

* Email: seino@sci.u-ryukyu.ac.jp

2 DNA塩基配列におけるパワースペクトル

本研究では, エクソンおよびイントロンに対してパワースペクトルを解析する.DNA塩基配列を一種の信号とみて解析するには, 塩基を数値にマッピングしてフーリエスペクトルを求める必要がある. 塩基を数値にマッピングする場合, 対象とするゲノムの全塩基に含まれるアデニン (A), チミン (T), グアニン (G), シトシン (C) の存在比を計算し対応させる. 例えば, ある塩基配列 $\{\mathrm{A,T,G,C,A,T,A,G}\}$ を存在比に置換する場合を考える. 配列 $\{\mathrm{A,T,G,C,A,T,A,G}\}$ にLocation番号 τ を左側から 1,2,3, $\cdots$,8 の順に割り当てる. 総塩基数を N, その中に含まれるA,T,G,Cの各塩基数を N_τ^η, η =A,T,G,Cとすると,

$$N = 8, N_\tau^{\mathrm{A}} = 3, N_\tau^{\mathrm{T}} = 2, N_\tau^{\mathrm{G}} = 2, N_\tau^{\mathrm{C}} = 1$$

となる. また,A,T,G,Cの各存在比を Z_τ^η とすると,

$$Z_\tau^\eta = \frac{N_\tau^\eta}{N}, \ \eta = \mathrm{A,T,G,C}, \ \tau = 1,2,3,\cdots,8$$

で与えられる. よって, 配列 $\{\mathrm{A,T,G,C,A,T,A,G}\}$ に対して, 存在比 Z_τ^η を割り当てて, τ =1,2,3, $\cdots$,8 の順に並べて作った数列の集合 $\boldsymbol{X}_\tau$ は,

$$\boldsymbol{X}_\tau = \left\{\frac{3}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}, \frac{3}{8}, \frac{2}{8}, \frac{3}{8}, \frac{2}{8}\right\}$$

となる. 本研究で取り扱う鳥類の塩基配列 B_τ に対して, 塩基の配列の順に存在比 Z_τ^η, τ =1,2,3, $\cdots$,8 を割り当てて, Z_τ^η のフーリエスペクトル $S(f)$ は,f を

周波数として,

$$S(f) = \left| \frac{1}{N} \sum_{\tau=1}^{N} Z_{\tau}^{\eta} e^{-i2\pi f \tau} \right|^2$$

で与えられる.

Ciconia boyciana(コウノトリ) の完全長 mtDNA 配列におけるエクソンを図 1 に示す. うぐいす色のボックスは,*Ciconia boyciana* の完全長 mtDNA 配列である. 完全長 mtDNA 配列の中でのエクソン部分の配置を抽出し, 青色のボックスで示す. エクソンで示した部分以外はイントロンであり,10 のイントロンに分けられる. エクソンの塩基配列およびたんぱく質名を表 1 に示す. イントロンの塩基配列および 5' から 3' 方向に向けて, イントロンに順番に intron01~intron10 と名前を付け表 2 に示す.

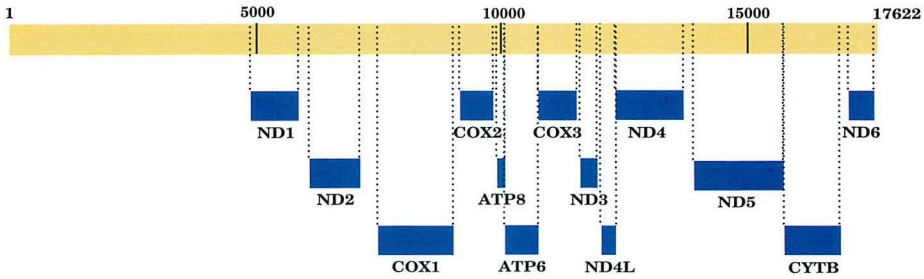

図 1: *Ciconia boyciana* の完全長 mtDNA 配列におけるエクソンの位置.

表 1: *Ciconia boyciana* におけるコーディング領域.

コーディング領域	エクソン名
4866-5843	NADH dehydrogenase subunit 1(=ND1)
6061-7101	NADH dehydrogenase subunit 2(=ND2)
7462-9012	Cytochrome c oxidase subunit I(=COX1)
9150-9833	Cytochrome c oxidase subunit II(=COX2)
9908-10075	ATP synthase F0 subunit 8(=ATP8)
10066-10749	ATP synthase F0 subunit 6(=ATP6)
10749-11532	Cytochrome c oxidase subunit III(=COX3)
11602-11952	NADH dehydrogenase subunit 3(=ND3)
12027-12323	NADH dehydrogenase subunit 4L(=ND4L)
12317-13694	NADH dehydrogenase subunit 4(=ND4)
13901-15712	NADH dehydrogenase subunit 5(=ND5)
15723-16865	Cytochrome b(=CYTB)
16867-16934	NADH dehydrogenase subunit 6(=ND6)

表 2: *Ciconia boyciana* における非コーディング領域.

非コーディング領域	イントロン名
1-4865	intron01
5844-6060	intron02
7102-7461	intron03
9013-9149	intron04
9832-9907	intron05
11533-11601	intron06
11953-12026	intron07
13695-13900	intron08
15713-15722	intron09
16866-17025	intron10

3 スペクトル解析

3.1 *Ciconia boyciana*(コウノトリ) におけるエクソンのスペクトル解析

Ciconia boyciana の完全長 mtDNA 配列データ [12] からエクソン部分を抽出し, スペクトル解析を行ったものを図 2 に示す.(i) は ND1 のパワースペクトルであり, 顕著なスペクトルのピークが周波数 0.33 および 0.66 で観測され, ピークの値は 0.071752 となった.(ii)ND2 および (iii)COX1, (iv)COX2, (vi)ATP6, (vii)COX3, (viii)ND3, (ix)ND4L, (x)ND4, (xi)ND5, (xii)CYTB は,(i) の ND1 同様に,0.33 および 0.66 で顕著なピークが観察された.(v)ATP8 において,0.33 および 0.66 では顕著なスペクトルのピークはみられず, 周波数 0.18 および 0.82 においてピークがみられ, このときピークの値は,0.038293 となった.(xiii) は ND6 であり,0.33 および 0.66 で顕著なスペクトルのピークはみられず, 周波数 0.01 および 0.99 においてピークがみられ, このときピークの値は 0.058126 となった. また,(xiv)Complete Genome Sequences においてもスペクトル解析を行い, 周波数 0.33 および 0.66 において顕著なピークがみられた. ピークの値は 0.185483 であった.

エクソンにおいて塩基配列が 168 と一番短い ATP8 の最大強度は 0.038293 であり, 塩基配列が 1812 と一番長い ND5 の最大強度は,0.47242 であった. 各エクソンによってピーク値が異なることから, ピークと塩基配列の長さの関係を図 3 に示す. 図 3 に,*Ciconia boyciana* の各エクソンのピーク値 (I_{MAX}) と塩基配列 N の長さとの関係が示されている.

$$I_{MAX} \propto \alpha N, \qquad \alpha = 2.2 \times 10^{-4}$$

これにより, 最大強度の値が緩やかに増大する傾向にあることがわかった.

3.2 *Ciconia boyciana* におけるイントロンのスペクトル解析

Ciconia boyciana の完全長 mtDNA 配列からイントロンを抽出し, スペクトル解析を行ったものを図 4 に示す. 図 4(i) の intron01 において, 顕著なスペクトルのピークは観察されなかった.intron01 の塩基配列の長さは 4865 であり, 完全長 mtDNA 配列を除いて最も長い配列となっていた.(ii) の intron02 から (x) の intron10 までは塩基配列が 10~360 までの長さの配列であり,(i) の intron01 と比較して, 短い塩基配列となっていた. これらの塩基配列では顕著なスペクトルのピークは観察されなかった.

3.3 20 種の鳥類におけるスペクトル強度と位置

Ciconia boyciana の ATP8 および ND6 を除いたエクソンや完全長 mtDNA 配列において, 顕著なスペクトルのピークが観測される. この特徴を鳥類全般にわたって調べるために, *Ciconia boyciana* を含め,20 種の鳥類の完全長 mtDNA 配列データを NCBI[12] より取得しスペクトル解析を行った. 例えば, *Ciconia boyciana* の ND1 において, コーディング領域 4866-5843 を 1-978 に置き換えて解析する. ND1 に関するパワースペクトルにおいて, 周波数 0.33 および 0.66 で最大強度が出現し, この場合の Position はそれぞれ 327 および 653 となっている. 表 3 および 4,5 に,20 種の鳥類における最大強度 (I_{MAX}) や塩基上で最大強度の出現する位置 (Position), 各エクソンおよび完全長 mtDNA 配列についての塩基の長さ (Length) をそれぞれ示す. 20 種について, ND3 のように 2 つの Location 番号で開始される遺伝子をもつ種は,*Ardea novaehollandiae* などの 12 種存在する. また,*Podiceps cristatus* には ND1 が存在せず,*Ardea novaehollandiae* と同様に 2 つの Location 番号で始まる ND3 を持っている.*Tachybaptus novaehollandiae* は,2 つの Location 番号で始まる ND1 および ND3 を持っている.*Thalassarche melanophrys* は,2 つの Location 番号で始まる ND6 を持っている.*Phaethon rubricauda* は mtDNA 上でのエクソンの配置が他の種のエクソンの配置と異なっており, 塩基配列の始まりに近い方から順に CYTB, ND6, ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4, ND5 の配置となっている. 他の種は図 1 の *Ciconia boyciana* と同様の配置となっている.

20 種に存在する ND1 および ND2, COX1, COX2, ATP6, COX3, ND4L, ND4, ND5, CYTB において, 周波数 0.33 および 0.66 で顕著なスペクトルのピークが観測された. これらのエクソン以外の ATP8,ND3,ND6 においては, ピークが観測される position は, 種によって異なっていた. さらに, 図 3 のようにピークの値が緩やかに増加する傾向はどの種のエクソンでも観測され, ピークの値は種による差異があることがわかった.

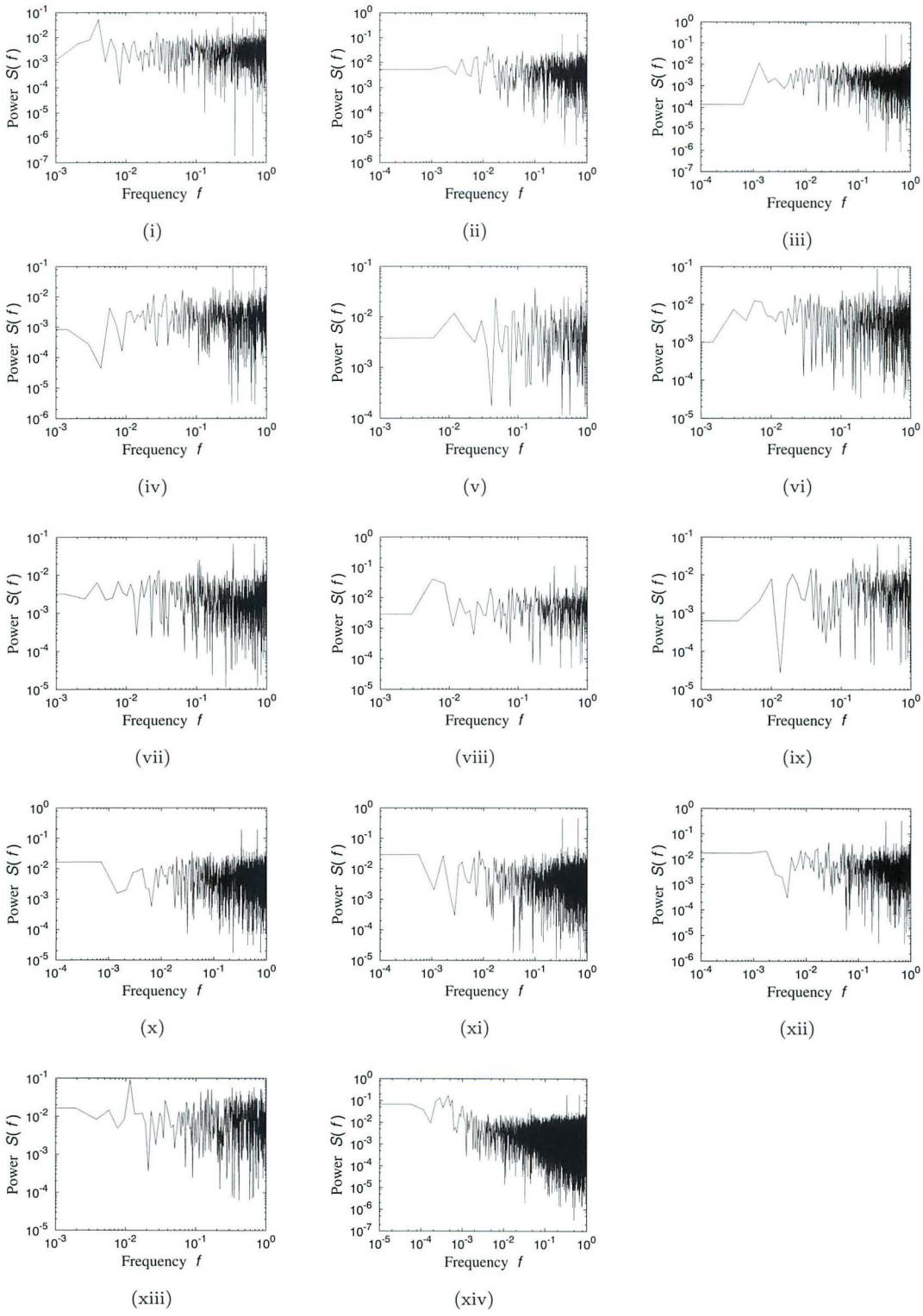

図 2: *Ciconia boyciana* のエクソンにおけるパワースペクトル. (i)ND1. (ii)ND2. (iii)COX1. (iv)COX2. (v)ATP8. (vi)ATP6. (vii)COX3. (viii)ND3. (ix)ND4L. (x)ND4. (xi)ND5. (xii)CYTB. (xiii)ND6. (xiv)Complete Genome Sequences.

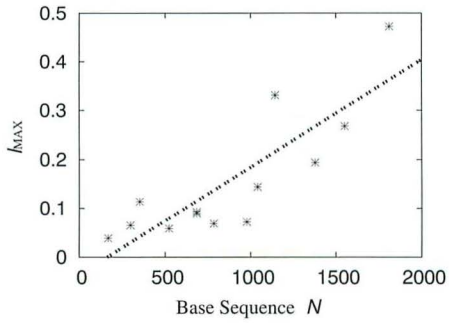

図 3: *Ciconia boyciana* のエクソンにおけるピーク値. $I_{\mathrm{MAX}} \propto \alpha N, \alpha = 2.2 \times 10^{-4}$.

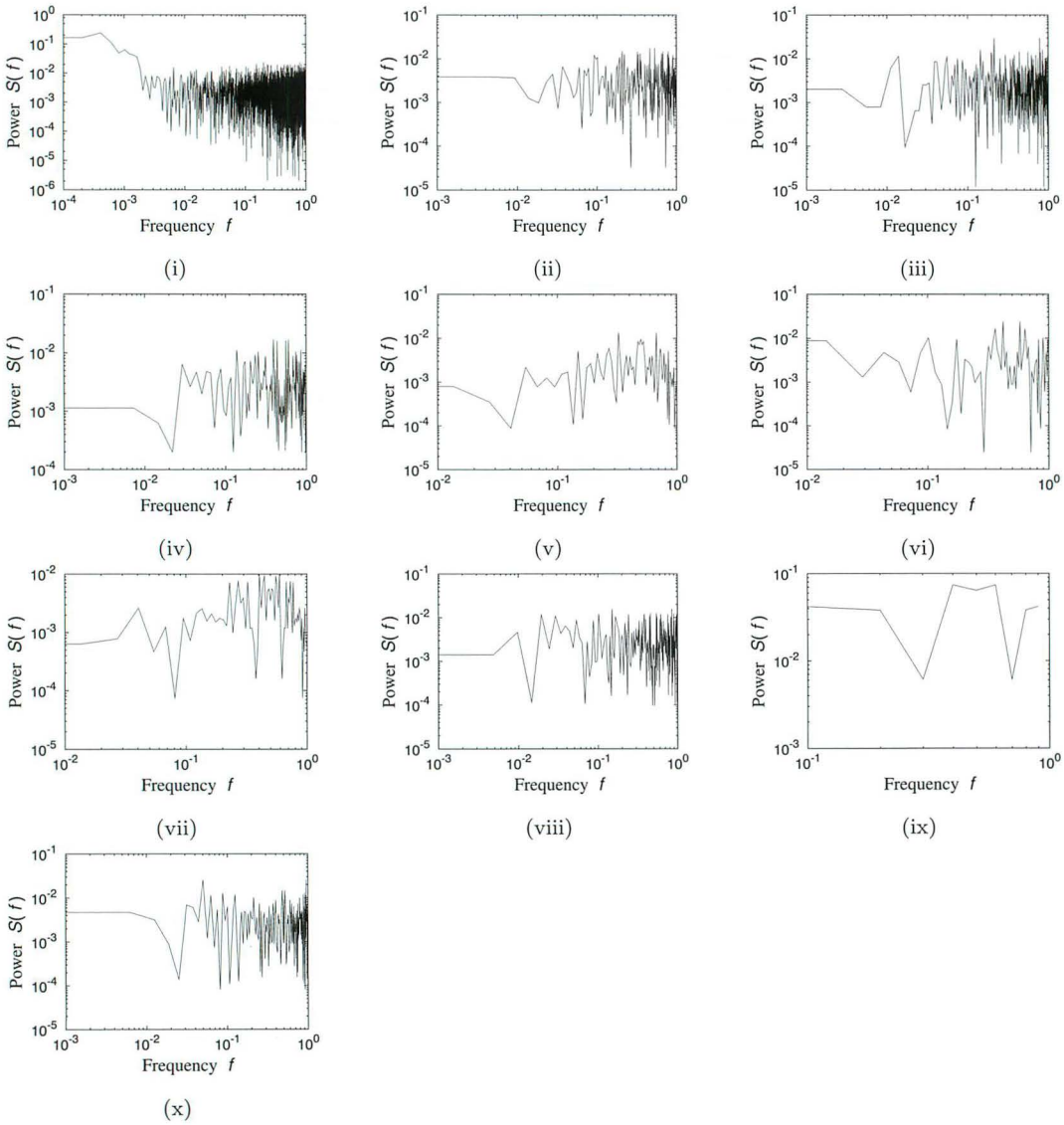

図 4: *Ciconia boyciana* のイントロンにおけるパワースペクトル. (i)1-4865(intron01). (ii)5844-6060(intron02). (iii)7102-7461(intron03). (iv)9013-9149(intron04). (v)9832-9907(intron05). (vi)11533-11601(intron06). (vii)11953-12026(intron07). (viii)13695-13900(intron08). (ix)15713-15722(intron09). (x)16866-17025(intron10).

表 3: エクソンのスペクトル強度 (I_{MAX}) とその配置. ND1, ND2, COX1, COX2, ATP8.

Protein / Scientific name	ND1				ND2				COX1				COX2				ATP8			
	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length
Ciconia boyciana	327	653	0.071752	978	348	695	0.144063	1041	518	1035	0.267558	1551	229	457	0.092612	684	31	139	0.038293	168
Ciconia ciconia	327	653	0.066757	978	348	695	0.144546	1041	518	1035	0.243671	1551	229	457	0.102139	684	33	137	0.026801	168
Ardea novaehollandiae	327	653	0.123739	978	348	695	0.114310	1041	518	1035	0.167090	1551	229	457	0.141781	684	33	137	0.045585	168
Egretta eulophotes	327	653	0.084575	978	347	694	0.091782	1039	518	1035	0.168529	1551	229	457	0.163939	684	33	137	0.029840	168
Ixobrychus cinnamomeus	327	653	0.089214	978	347	694	0.095240	1039	518	1035	0.191973	1551	229	457	0.117840	684	31	139	0.027998	168
Threskiornis aethiopicus	327	653	0.052142	978	348	694	0.050657	1040	518	1035	0.176403	1551	229	457	0.120057	684	32	138	0.032985	168
Platalea minor	327	654	0.059172	978	348	694	0.039383	1040	518	1035	0.149503	1551	229	457	0.119038	684	22	148	0.030997	168
Platalea leucorodia	327	653	0.058418	978	348	694	0.051992	1040	518	1035	0.162163	1551	229	457	0.122464	684	21	149	0.025617	168
Podiceps cristatus					348	695	0.163099	1041	518	1035	0.267940	1551	229	457	0.155897	684	57	113	0.039976	168
Tachybaptus novaehollandiae	56	111	0.024332	165	347	694	0.149812	1039	518	1035	0.351785	1551	229	457	0.193999	684	23	147	0.033931	168
	272	543	0.114214	813																
Phaethon rubricauda	327	653	0.119869	978	348	695	0.197654	1041	518	1035	0.205998	1551	229	457	0.074238	684	31	139	0.022681	168
Eudyptula minor	327	653	0.095477	978	348	695	0.125613	1041	518	1035	0.224004	1551	229	457	0.170553	684	4	163	0.038051	165
Pterodroma brevirostris	327	653	0.054251	978	348	695	0.097481	1041	518	1035	0.113163	1551	229	457	0.089377	684	3	139	0.025783	168
Thalassarche melanophrys	327	653	0.054903	978	347	694	0.056778	1039	518	1035	0.126379	1551	229	457	0.092017	684	31	139	0.026638	168
Cathartes aura	327	653	0.044196	978	348	695	0.060853	1041	518	1035	0.126499	1551	229	457	0.073161	684	2	171	0.053131	171
Buteo buteo	327	653	0.090324	978	348	695	0.158839	1041	518	1035	0.224678	1551	229	457	0.137635	684	55	115	0.040017	168
Carduelis sinica	327	653	0.053445	978	348	694	0.047101	1040	518	1035	0.206639	1551	229	457	0.106823	684	66	104	0.017799	168
Cyanopica cyanus	327	653	0.049615	978	348	695	0.066274	1041	518	1035	0.141680	1551	229	457	0.061750	684	31	139	0.025713	168
Ficedula zanthopygia	327	653	0.057755	978	348	694	0.031059	1040	518	1035	0.126174	1551	229	457	0.114804	684	27	143	0.027081	168
Garrulus glandarius	327	653	0.069455	978	348	653	0.077015	1041	518	1035	0.098495	1551	229	457	0.093485	684	79	91	0.022227	168

表 4: エクソンのスペクトル強度 (I_{MAX}) とその配置. ATP6, ND3, ND4L, ND4.

Protein / Scientific name	ATP6				COX3				ND3				ND4L				ND4			
	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length
Ciconia boyciana	229	457	0.089029	684	262	524	0.068206	784	118	235	0.113518	351	100	199	0.064621	297	460	920	0.194079	1378
Ciconia ciconia	229	457	0.105731	684	262	524	0.063085	784	118	235	0.080812	351	100	199	0.049789	297	461	922	0.215452	1381
Ardea novaehollandiae	229	457	0.089003	684	262	524	0.104441	784	59	117	0.009927	174	100	199	0.058914	297	460	920	0.205652	1378
									60	119	0.043871	177								
Egretta eulophotes	229	456	0.145697	683	262	524	0.079954	784	8	168	0.024837	174	99	197	0.058201	294	460	920	0.208918	1378
									60	119	0.053818	177								
Ixobrychus cinnamomeus	229	456	0.099506	684	262	524	0.097553	784	35	141	0.020182	174	100	199	0.055087	297	460	920	0.144180	1378
									60	119	0.024511	177								
Threskiornis aethiopicus	229	457	0.090171	684	262	524	0.068986	784	58	115	0.032802	171	65	234	0.013935	297	460	920	0.153179	1378
									61	121	0.029855	180								
Platalea minor	229	457	0.080752	684	262	524	0.071513	784	9	166	0.028667	173	100	199	0.047225	297	460	920	0.117938	1378
									16	164	0.024179	178								
Platalea leucorodia	229	457	0.070402	684	262	524	0.077084	784	9	165	0.030211	172	100	199	0.052600	297	460	920	0.139845	1378
									2	179	0.036282	179								
Podiceps cristatus	229	457	0.165931	684	262	524	0.099089	784	59	117	0.034795	174	100	199	0.064641	297	457	913	0.431015	1368
									60	119	0.068717	177								
Tachybaptus novaehollandiae	229	457	0.236988	684	262	524	0.098666	784	59	117	0.079539	174	100	199	0.105558	297	457	913	0.427431	1368
									60	119	0.044461	177								
Phaethon rubricauda	210	488	0.077449	696	262	524	0.048876	784	3	350	0.029931	351	56	243	0.020106	297	457	913	0.291489	1368
Eudyptula minor	229	457	0.127218	684	262	524	0.124097	784	8	167	0.028868	173	100	199	0.070551	297	461	921	0.381179	1380
									2	178	0.041635	178								
Pterodroma brevirostris	229	457	0.070300	684	262	524	0.034897	784	9	166	0.024896	173	100	199	0.023072	297	460	920	0.098137	1378
									2	178	0.032198	178								
Thalassarche melanophrys	229	457	0.121021	684	262	524	0.063952	784	35	140	0.021179	173	100	199	0.023447	297	460	920	0.132383	1378
									16	164	0.016207	178								
Cathartes aura	229	457	0.106231	684	262	524	0.067210	784	32	143	0.023426	173	100	199	0.041349	297	460	920	0.071724	1378
									2	178	0.025515	178								
Buteo buteo	229	457	0.078490	684	262	524	0.093966	784	118	235	0.092578	351	100	199	0.029009	297	460	920	0.205512	1378
Carduelis sinica	229	457	0.060384	684	262	524	0.076831	784	118	235	0.029987	351	100	199	0.023897	297	460	920	0.161882	1378
Cyanopica cyanus	229	457	0.060153	684	262	524	0.034142	784	3	350	0.015958	351	65	234	0.029852	297	460	920	0.088553	1378
Ficedula zanthopygia	229	457	0.116647	684	262	524	0.079470	784	119	236	0.018727	353	100	199	0.057375	297	460	920	0.183889	1378
Garrulus glandarius	229	457	0.059279	684	262	524	0.055771	784	118	235	0.023189	351	65	234	0.032672	297	460	920	0.091318	1378

表 5: エクソンのスペクトル強度 (I_{MAX}) とその配置. ND5, ND6, CYTB, 完全長 DNA 配列

Protein / Scientific name	ND5				ND6				CYTB				Complete Genome Sequence			
	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length	Position		I_{MAX}	Length
Ciconia boyciana	605	1209	0.47243	1812	172	352	0.058126	522	382	763	0.331408	1143	5874	11750	0.185483	17622
Ciconia ciconia	605	1209	0.51571	1812	7	517	0.071048	522	382	763	0.253633	1143	5782	11567	0.199644	17347
Ardea novaehollandiae	606	1211	0.60860	1815	181	343	0.063914	522	382	763	0.331979	1143	5842	11671	0.193111	17511
Egretta eulophotes	606	1211	0.56839	1815	181	343	0.083727	522	382	763	0.327448	1143	5866	11715	0.292284	17579
Ixobrychus cinnamomeus	606	1211	0.39793	1815	98	426	0.056940	522	382	763	0.242714	1143	5733	11449	0.277569	17180
Threskiornis aethiopicus	606	1211	0.28963	1815	90	434	0.105872	522	382	763	0.247243	1143	5656	11306	0.180101	16960
Platalea minor	606	1211	0.28980	1815	214	310	0.082883	522	382	763	0.231716	1143	5640	11280	0.218708	16918
Platalea leucorodia	606	1211	0.29958	1815	78	446	0.106051	522	382	763	0.276141	1143	5572	11145	0.147543	16715
Podiceps cristatus	606	1211	0.52303	1815	7	517	0.093976	522	382	763	0.251439	1143	5379	10757	0.265282	16134
Tachybaptus novaehollandiae	606	1211	0.68031	1815	8	516	0.082535	522	382	763	0.304318	1143	6003	12001	0.332118	18002
Phaethon rubricauda	605	1209	0.29605	1812	230	294	0.068465	522	382	763	0.216033	1143	5930	11849	0.279406	17777
Eudyptula minor	608	1215	0.55312	1821	180	341	0.068719	519	382	763	0.320823	1143	5869	11744	0.382227	17611
Pterodroma brevirostris	606	1211	0.24397	1815	88	436	0.083255	522	382	763	0.194490	1143	5473	10943	0.138280	16414
Thalassarche melanophrys	606	1211	0.31993	1815	186	344	0.055637	528	382	763	0.171414	1143	12	18957	0.190838	18967
					121	409	0.059442	528								
Cathartes aura	606	1211	0.41085	1815	4	520	0.074148	522	382	763	0.113839	1143	5591	11190	0.166601	16779
Buteo buteo	607	1213	0.46731	1818	152	369	0.063498	519	382	763	0.308734	1143	6224	12452	0.288636	18674
Carduelis sinica	607	1213	0.28289	1818	180	341	0.089859	519	382	763	0.225964	1143	5609	11206	0.228747	16813
Cyanopica cyanus	607	1213	0.24321	1818	4	520	0.079368	522	382	763	0.152421	1143	5633	11262	0.356603	16893
Ficedula zanthopygia	607	1213	0.43502	1818	180	341	0.095379	519	382	763	0.235348	1143	5599	11197	0.196648	16794
Garrulus glandarius	607	1213	0.19039	1818	181	343	0.045231	522	382	763	0.135565	1143	5645	11287	0.151723	16930

3.4 塩基配列部分置換におけるスペクトル解析

鳥類のエクソンにおいて, 顕著なスペクトルのピークが現れることを 3.3 節で示した. その意味を考察するため, *Ciconia boyciana* の ND1, ND2, COX1, COX2, ATP6, ND4L, ND4, ND5, CYTB, ND6 について, 塩基配列の一部を切り取り ,intron01(1-4865) の塩基配列と置換した. 図 5 において,ND1 の 3' 方向から 100 配列を切り取り ,intron01(1-4865) の塩基配列の 3' 方向からの 100 配列と置換した.ND1 の場合,100 配列,200 配列,300 配列,400 配列と 4 回置換した. ピーク値は, 置換前が,0.071751 であり ,100 配列置換すると 0.070176,200 配列置換すると 0.062769,300 配列置換すると 0.046425,400 配列置換すると 0.024899 となり, 減少していくことがわかる.500 配列置換すると, 顕著なピークはみられなくなり,400 配列置換したときの, ピーク値まで有効とした. 一般に置換される配列の長さを L とすると,L を増やしていき顕著なスペクトルのピークが表れなくなるまで解析した. 図 6 に, 置換した配列の長さと, ピーク値の変化を示す. 置換する塩基の長さが増えると, 最大強度が一定の割合で弱くなる傾向が示された. 最小二乗法により $I_{\mathrm{MAX}} \propto \beta L$, を得る. ここで, 各エクソンにおける β を以下に示す.

ND1: $\beta = -2.4 \times 10^{-4}$,
ND2: $\beta = -1.5 \times 10^{-4}$,
COX1: $\beta = -2.3 \times 10^{-4}$,
COX2: $\beta = -2.2 \times 10^{-4}$,
ATP6: $\beta = -1.3 \times 10^{-4}$,
COX3: $\beta = -1.1 \times 10^{-4}$,
ND4L: $\beta = -3.8 \times 10^{-4}$,
ND4: $\beta = -1.2 \times 10^{-4}$,
ND5: $\beta = -3.2 \times 10^{-4}$,
CYTB: $\beta = -3.5 \times 10^{-4}$.

各エクソンにおける β の取り得る値は,$-1.1 \times 10^{-4} \sim -3.8 \times 10^{-4}$ の間でばらつきが見られ, 緩やかな減少を示す.COX3 における β が -1.1×10^{-4} と最も緩やかで,ND4L における β が -3.8×10^{-4} と最も大きいことがわかる.

スペクトルにおける顕著なピークが, エクソンではみられ, イントロンではみられないことから, エクソンの塩基配列の中には, フーリエ解析によりスペクトルのピークがみられる配列が存在すると考えられる. エクソンとイントロンの配列を部分置換していくと, ピーク値が一定の割合で減少することから, ピーク値を示す配列は, エクソンの塩基配列の中に一様に組み込まれていることが示唆された.

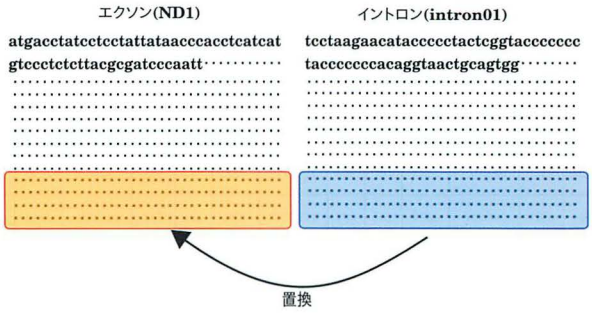

図 5: エクソンへの intron01 部分配列における配列置換.

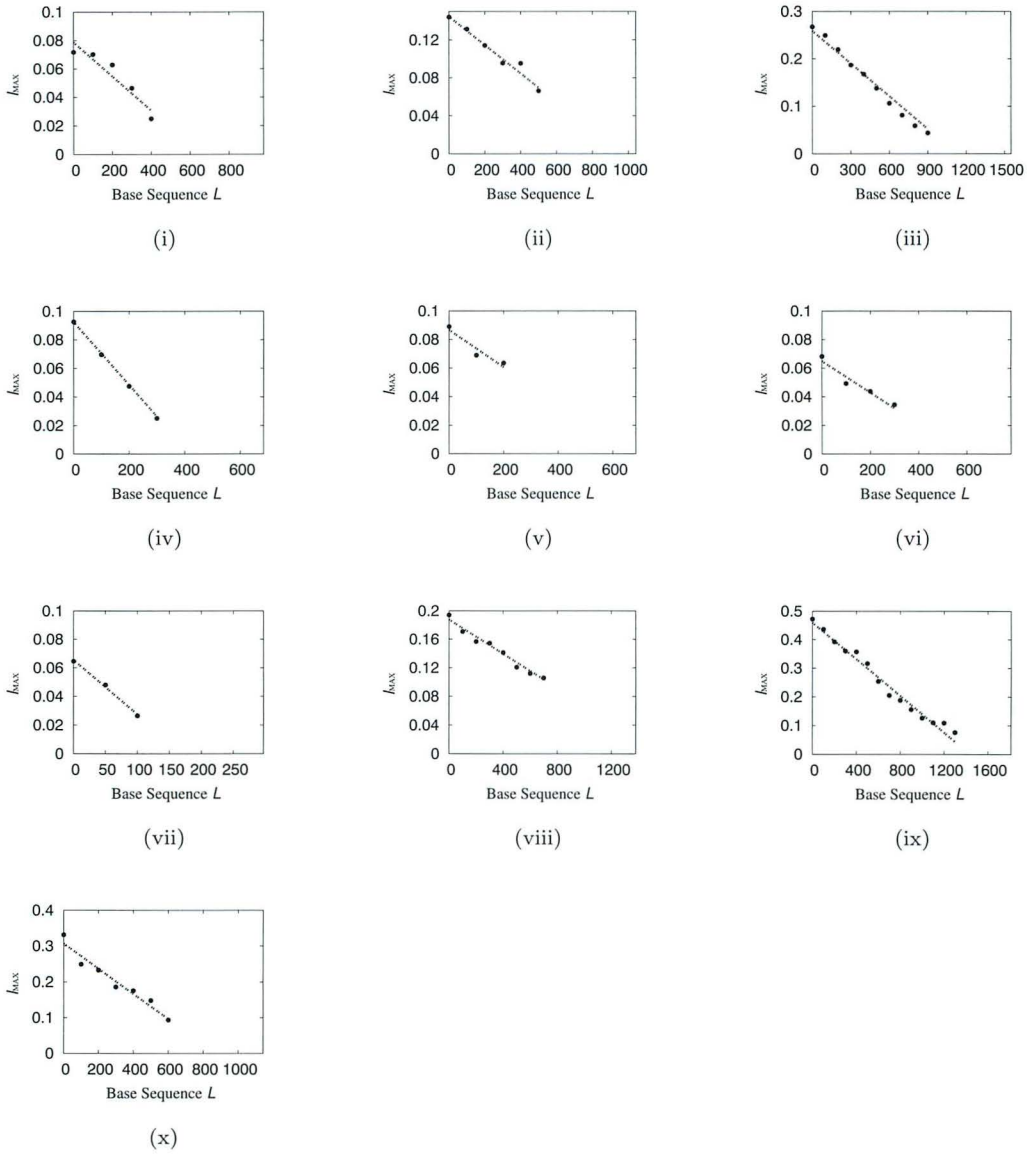

図 6: *Ciconia boyciana* のエクソンにおけるピーク値の変化.$I_{\mathrm{MAX}} \propto \beta L$. (i)ND1, $\beta = -2.4 \times 10^{-4}$. (ii)ND2, $\beta = -1.5 \times 10^{-4}$. (iii)COX1, $\beta = -2.3 \times 10^{-4}$. (iv)COX2, $\beta = -2.2 \times 10^{-4}$. (v)ATP6, $\beta = -1.3 \times 10^{-4}$.(vi)COX3, $\beta = -1.1 \times 10^{-4}$.(vii)ND4L, $\beta = -3.8 \times 10^{-4}$. (viii)ND4, $\beta = -1.2 \times 10^{-4}$. (ix)ND5, $\beta = -3.2 \times 10^{-4}$. (x)CYTB, $\beta = -3.5 \times 10^{-4}$.

4 まとめ

鳥類の完全長 mtDNA 配列に注目し,20 種の鳥類の完全長 mtDNA 配列に関するスペクトル解析を行った. エクソン, イントロン, 完全長に関してパワースペクトルを求めた.

Ciconia boyciana のエクソンおよびイントロンについて, スペクトル解析を行った. エクソンについて, 顕著なスペクトルのピークが見られた. エクソンのもつ塩基の配列長とピーク値との相対的関係を示した.

鳥類のエクソンに見られる, 顕著なスペクトルのピークの意味を探るため,*Ciconia boyciana* の ND1, ND2, COX1, COX2, ATP6, ND4L, ND4, ND5, CYTB, ND6 について, 塩基配列の一部を切り取り,intron01(1-4865) の塩基配列と置換した. 置換する塩基の配列長が増加すると, ピーク値が一定の割合で減少する傾向が示された. これにより, エクソンの塩基配列の中に, ピークを示す配列が一様に分布しており, その塩基配列の重ね合わせにより, エクソン全体で顕著なピークが見られることが示唆された.

謝辞

原稿の校閲をして頂いた理学部物質地球科学科の深水孝則教授に感謝を致します.

参考文献

[1] S. J. Hackett, R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, Kin-Lan Han, J. Harshman, C. J. Huddleston, Ben D.Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt and T.Yuri: "A Phylogenomic Study of Birds Reveals Their Evolutionary History" ,*Science* **320**,1763-1768, (2008).

[2] 曹纓, 長谷川政美:"分子系統樹推定におけるモデルのミススペシフィケーション, 統計数理",**50**,69-85, (2002).

[3] R. F. Voss: "Evolution of Long-Range Correlations and $1/f$ Noise in DNA Base Sequences",*Phys.Rev.Lett.***68**,3805-3808,(1992).

[4] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S.Havlin, F.Sciortino, M.Simons, and H. E. Stanley: "Long-range correlations in nucleotide sequences", *Nature* **356**,168-170,(1992).

[5] W. Li, and K. Kaneko: "Long-Range Correlation and Partial $1/f^{\alpha}$ Spectrum in a Non-coding DNA Sequence", *Europhys.Lett.***17**,655-660,(1992).

[6] W. Li, and K. Kaneko: "DNA correlations", *Nature* **360**,635-636, (1992).

[7] E. Takushi, H. Sakamoto and M. Seino: "Power Spectra of Bird-Influenza A Vrius(BIV) ssRNA Base Sequences: Possibility of Cooperative Genes or Interferences",*Proc.* 17th *Int. Conf. on Genome Informatics* **37**,1-37.2,(2006).

[8] M. Seino and H. Sakamoto: "Analysis and Estimation for the Substituted Sequences of DNA Sequences for Saccharomyces Cerevisiae with a Kind of Marcov Model", *Proc.* 17th *Int. Conf. on Genome Informatics* **47**,1-47.2,(2006).

[9] E. Takushi: "Powerspectra of DNA sequences in Phage ane tumor suppressor genes(TSG)", *Genome Infom.* **13**,412-413,(2002).

[10] E. Takushi, and Y. Chinen: "Powerspectra of DNA base sequence in human mitochondrial genome", *Proc. Int. Conf. Biological Physics,* **780**, Goeteborg, Sweden, (2004).

[11] W.Li and D.Holste: "Universal $1/f$ noise, crossovers of scaling exponents,and chromosome-specific patterns of guanine-cytosine content in DNA sequences of the human genome", *Phys.Rev.E* **71**,041910,(2005).

[12] NCBI (National Center for Biotechnology Information), http://www.ncbi.nlm.nih.gov/ (2011/10/26 アクセス).