

## 第4章 ジヒドロフラボノール 4-レダクターゼ遺伝子配列とその5'上流領域配列の単離と解析

### 1. 緒言

第2章・第2節において、*CHI*、*F3H*、*DFR*、*ANS*、*3GT*、*5GT*および*3AT*のアントシアニン生合成遺伝子のRT-PCR分析により、ダッチアイリスの白色花品種「White Wedgewood」の内・外両花被および雌蕊の白色化は*DFR*遺伝子発現の欠損または顕著な減少に、また複色花品種「Surprise」の外花被および雌蕊の白色化は*DFR*遺伝子発現の顕著な減少によることが明らかにされた。このような遺伝子発現の欠損または減少機構を解明するためには、構造遺伝子ばかりでなく、転写調節（プロモーター）領域がある5'上流領域のDNA配列や転写因子の単離・解析する必要がある。また、構造遺伝子およびその5'上流領域のゲノム配列を単離・解析することは、色素形成の時期や器官に特異的に機能するプロモーターを得ることも可能である。さらに、そのプロモーターを利用することにより、アントシアニン生合成に関わる調節機構の基礎研究および育種ツールとしても利用することが期待できる。

これまでのところ、アントシアニン生合成に関与する構造遺伝子やその転写調節領域のゲノム構造の解析は、変異体解析の進んでいるトウモロコシ、キンギョソウ、ペチュニア、シロイヌナズナおよびアサガオなどで主に行われてきており（Chopra *et al.* 2006）、転写調節因子であるMYB因子、bHLH (basic helix-loop-helix) 因子およびWD40因子の相互作用によって、一連のアントシアニンの生合成と蓄積に関わる遺伝子の協調した発現が制御されていることが明らかにされてきた（Quattrocchio *et al.* 2006）。しかしながら、他の植物種でのプロモーターおよび転写因子の解析に関する情報は少なく、特に*Iris*属植物ではその報告例が全く見当たらないのが実情である。

そこで、本章ではダッチアイリスの「ブルーダイヤモンド」、「White Wedgewood」および「Surprise」の3品種を用いて、*DFR*構造遺伝子とその5'上流領域配列を単離・解析し、

「White Wedgewood」および「Surprise」における *DFR* 遺伝子発現の欠損または減少の原因を解明した。

## 2. 材料および方法

### 1) 植物材料

ダッチアイリスの3品種、「ブルーダイヤモンド」、「White Wedgewood」および「Surprise」は、第2章・第2節と同じものを供試した。

### 2) DNA 抽出

ダッチアイリスのゲノム DNA は、CTAB 法を用いて各品種の葉から抽出した。抽出した DNA は TE に溶解した後、4°C で冷蔵保存した。

### 3) Nested-PCR を用いた *DFR* 遺伝子のゲノム DNA 配列の獲得

1回目のPCR反応は、20  $\mu$ l の1× ExTaq buffer 中で、鋳型として500 ng のゲノム DNA、200  $\mu$ M dNTPs、0.25  $\mu$ M プライマー (IhDFR.N1、5'-CCC GTT GTC GTG ACC GGA G-3' および IhDFR.P2、5'-TCA CTT GAC CTT TTT TTT TGG-3')、および5 units TaKaRa ExTaq® (タカラバイオ株式会社) を用いて行った。PCR プログラムは、初期変性5分間の後、変性94°C、30秒間、アニーリング48°C、15秒間、伸長72°C、2.5分間を40サイクルとした。1回目のPCR増幅産物1  $\mu$ l は、2回目のPCR反応の鋳型として用いた。2回目のPCRは、1回目の反応溶液のプライマーをIhDFR.N2 (5'-GCT ACG TCG GTT CAT GGC TT-3') およびIhDFR.P2-2 (5'-CAA TTC GGG CAT GTT CCT CTG-3') に変更したものをを用いた。PCR プログラムは、初期変性5分間の後、変性94°C、30秒間、アニーリング50°C、15秒間、伸長72°C、2.5分間を40サイクル、3'-A オーバーハング付加反応72°C、10分間で行った。

#### 4) Inverse PCR を用いた *DFR* 遺伝子のプロモーター配列の獲得

4 µg ゲノム DNA を、1 × A buffer 20 µl 中で 40 unit の *SacI* (Roche Diagnostics K.K.) を用いて 37°C、16 時間、制限酵素処理した。*SacI* の失活は、65 °C、15 分間の加熱により行った。次いで、*SacI* 処理溶液 5 µl を等量の Ligation high (東洋紡績株式会社) と混合し、16 °C、16 時間インキュベートすることで、DNA 断片をセルフライゲーションさせた。

セルフライゲーションさせた DNA (環状化 DNA) 1 µl を鋳型として使用し、20 µl の 1 × ExTaq buffer 中で、200 µM dNTPs、0.25 µM プライマー (IhDFR.N>C-1, 5'-GGC AGA GAG TTA GAC CTA C-3' および IhDFR.C>N-1, 5'-GTC GCG AAG GAG CTT CAT AAC A-3')、および 5 units TaKaRa Ex Taq® (タカラバイオ株式会社) を用いて 1 回目の PCR 反応を行った。PCR プログラムは初期変性 5 分間の後、変性 94°C、30 秒間、アニーリング 50 °C、15 秒間、伸長 72 °C、2.5 分間を 40 サイクルとした。1 回目の PCR 増幅産物 1 µl は、2 回目の PCR 反応の鋳型として用いた。2 回目の PCR 反応は、プライマーとして IhDFR.N>C-2 (5'-CGG TGG ACG TGA AAG AAC ATC-3') および IhDFR.C>N-2 (5'-CAA GCC ATG AAC CGA CGT AG-3') を使用し、PCR プログラムの最後に 3'-A オーバーハング付加反応 72°C、10 分間を追加した以外、反応溶液の組成およびプログラムは 1 回目と同様に行った。

#### 5) ベクターへのサブクローニング

上記の 2) および 3) にて増幅した DNA 断片は、pGEM®-T Easy Vector System (Promega Co. Ltd.) を用いてサブクローニングした。サブクローニングはキットの使用法に従った。

#### 6) 塩基配列の解析

塩基配列は第 2 章・第 1 節と同様の方法により決定し、転写因子結合部位などの核酸

のモチーフ検索は、植物のシスエレメントデータベースである PLACE の Signal Scan Search program (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>) および Genome Net 提供の TFSEARCH (<http://www.cbrc.jp/research/db/TFSEARCHJ.html>) を用いて行った。

### 3. 結果および考察

#### 1) *IhDFR* のゲノム構造

ダッチアイリスの品種、「White Wedgewood」および「Surprise」における *DFR* 遺伝子発現の欠損または減少の原因を解明するために、まず、第2章・第1節で単離した「ブルーダイヤモンド」の *IhDFR* cDNA の5'末端および3'末端配列を利用した nested-PCR により、「ブルーダイヤモンド」、「White Wedgewood」および「Surprise」のゲノム *DFR* (*gDFR*) 配列を獲得した。各品種の *gDFR* 配列の全長（開始メチオニンから終止コドンまで）は、「ブルーダイヤモンド」から *gDFR-BD1* (1,968 bp) および *gDFR-BD2* (1,962 bp)、「White Wedgewood」 *gDFR-WW7* (1,957 bp)、*gDFR-WW13* (1,969 bp) および *gDFR-WW15* (1,966 bp)、「Surprise」 *gDFR-Su1* (1,958 bp) および *gDFR-Su2* (1,960 bp) であり、5つのエクソンからなる 1,086 bp の ORF と、4つのイントロンから構成されていた (Fig. 4-1A)。また、イントロンのスプライス部位はすべて GT...AG という真核生物のイントロンに共通の配列が見出された。一方、第1から第5エクソンは 109、170、195、160 および 452 bp の塩基対からなり、第2から第4エクソンに含まれる塩基対数は、多くの植物種の *gDFR* 構造と類似していた (Table 4-1、Fig. 4-1A)。また、*IhDFR* のゲノム DNA および cDNA (第2章・第1節) 配列を比較したところ、*IhDFR* と他種植物の *DFR* の推定アミノ酸配列のマルチプルアライメントにより明らかになった *DFR* の基質認識部位 (Fig. 2-3A) が、ゲノム DNA では第3エクソンにコードされていることが示された (Fig. 4-1A)。さらに、*DFR* の補因子である NADP<sup>+</sup>との結合に関与するアミノ酸残基は、第1、第2および第4エクソン上にコードされていた (Fig. 4-1A)。

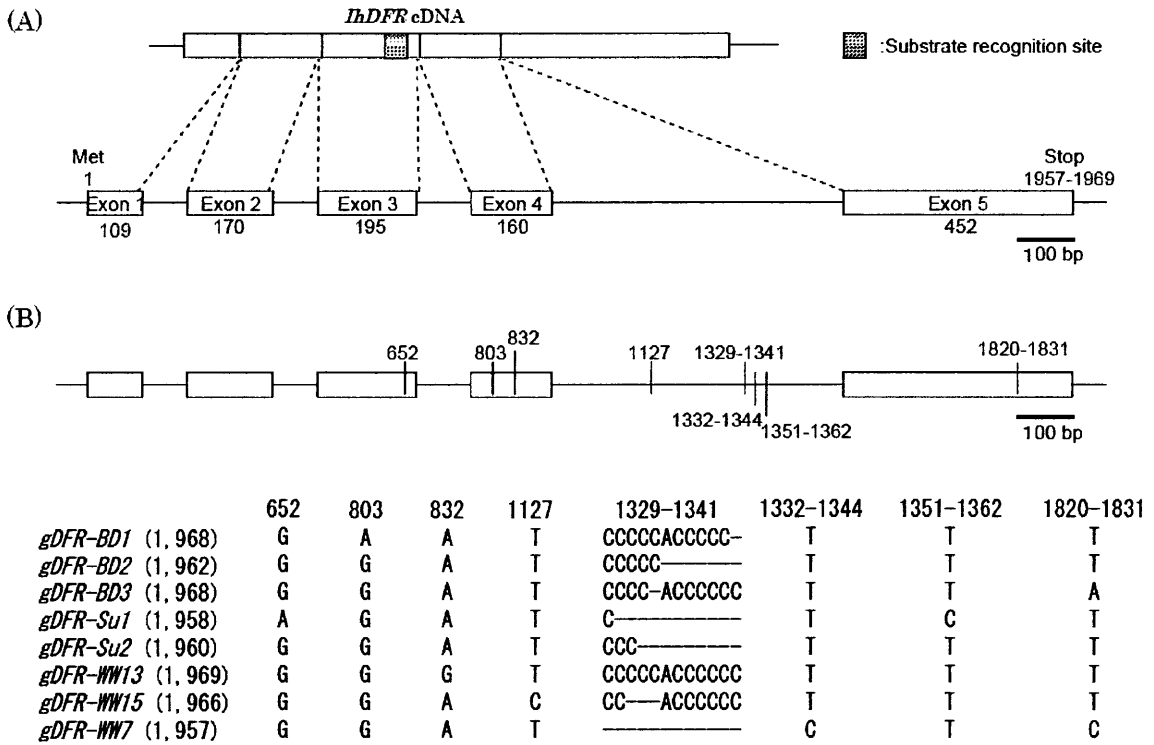


Fig. 4-1(A). Genomic and cDNA structures of *IhDFRs*. Exons are indicated by boxed regions. Introns and 5'- and 3'- untranslated regions are represented by lines. (B) Positions of different sequences among *IhDFR* genome sequence of *Iris hollandica* cv. Blue Diamond (BD), cv. White Wedgewood (WW) and cv. Surprise (Su).

Table 4-1. Species comparison of exon and intron length of *IhDFR* gDNA with those of other plant *DFR* gDNAs

	Accession no.	Exon 1	Intron 1	Exon 2	Intron 2	Exon 3	Intron 3	Exon 4	Intron 4	Exon 5	Intron 5	Exon 6
<i>Agapanthus praecox</i>	AB099529	142	142	170	464	355	1842	193	80	277		
<i>Allium cepa</i>	AY221250	210	91	170	104	195	63	160	81	193	79	332
<i>Anthurium andraeanum</i>	DQ364060	115	219	170	77	198	261	160	95	193	151	208
<i>Arabidopsis thaliana</i>	AB033294	118	91	170	79	195	86	160	79	193	90	319
<i>Brassica juncea</i>	EF128034	118	78	170	76	195	87	160	73	193	84	322
<i>Brassica rapa</i>	AY953249	118	101	170	77	195	88	160	71	193	74	322
<i>Citrus sinensis</i> cv. Tarocco	DQ084723	118	98	170	294	195	111	160	136	193	78	181
<i>Citrus sinensis</i> cv. Navel	DQ084722	118	98	170	293	195	111	160	136	193	78	181
<i>Fragaria x ananassa</i>	AY695813	127	94	170	209	195	110	160	153	193	102	184
	AY575057											
<i>Glycine max</i>	EF187612	124	143	170	1025	195	630	160	344	193	103	223
<i>Ipomoea batatas</i> DFR-B	EF187612	124	233	170	127	195	419	160	93	193	554	343
<i>Ipomoea nil</i> (DFR-A)	AB112545	139	300	170	138	195	483	160	114	193	527	208
<i>Ipomoea nil</i> (DFR-B)	AB006793	133	165	170	266	195	827	160	93	193	898	361
<i>Ipomoea nil</i> (DFR-C)		142	146	176	290	195	109	160	81	193	76	205
<i>Ipomoea purpurea</i> (DFR-A)		139	570	170	111	195	517	160	118	193	267	190
<i>Ipomoea purpurea</i> (DFR-B)	AB011667	133	201	170	806	195	460	160	87	193	629	310
<i>Ipomoea purpurea</i> (DFR-C)		142	105	176	283	195	102	160	81	193	76	205
<i>Iris hollandica</i>	this study	109	92	170	90	195	110	160	580-591	453		
<i>Lophopyrum ponticum</i> x <i>Triticum aestivum</i>	AY208999	118	76	170	75	195	85	582				
<i>Lophopyrum ponticum</i>	AY208997	118	99	170	86	195	82	582				
<i>Lotus corniculatus</i>	AY633707	118	153	170	383	195	147	160	179	193	127	175
<i>Oryza sativa</i> (japonica cultivar group)	AB003495	121	112	365	387	633						
<i>Petunia hybrida</i> (V30)	X79723	148	149	170	92	195	89	160	138	193	86	277
<i>Solanum tuberosum</i> (clone 07506-01)	AY289924	154	150	170	102	195	105	160	100	193	88	277
<i>Solanum tuberosum</i> (clone W5281.2)	AY289923	154	158	170	102	195	94	160	109	193	88	277
<i>Triticum aestivum</i> (TaDFR-D)	AB162140	118	109	170	103	195	83	582				
<i>Triticum aestivum</i> (TaDFR-B)	AB162139	118	81	170	93	195	99	582				
<i>Triticum aestivum</i> (TaDFR-A)	AB162138	118	76	170	75	195	85	582				
<i>Triticum aestivum</i>	AY208998	118	81	170	92	195	99	582				
<i>Triticum aestivum</i>	AY210885	118	73	230	32	195	85	582				

ダッチアイリスの品種、「ブルーダイヤモンド」、「White Wedgewood」および「Surprise」の間で *gDFR* の塩基配列を比較したところ、第3、第4および第5エクソンと第4イントロンに配列の差異が認められるものの、各配列間の相同性は99%以上であり、白色花品種「White Wedgewood」においても停止コドン(UAA、UAGおよびUGA)の出現により早期翻訳中止を導くナンセンス変異は確認されなかった(Fig. 4-1B)。また、3品種の構造遺伝子内にトランスポゾンやレトロトランスポゾンなどの転移因子の挿入も確認されなかったため、次に *gIhDFR* の5'上流領域配列を単離・解析し、プロモーター領域の獲得を試みた。

## 2) *IhDFR* のプロモーター領域の構造

第2章・第1節において獲得した *IhDFR* cDNA、ならびに本章で獲得した *gDFR* 配列を基とした Inverse PCR により、*gIhDFR* の5'上流領域を獲得した。獲得した配列は、「ブルーダイヤモンド」において *gDFR-BD5f1* (791 bp)、*gDFR-BD5f2* (3,034 bp)、*gDFR-BD5f3* (2,946 bp) および *gDFR-BD5f4* (2,944 bp) 「White Wedgewood」において *gDFR-WW5f1* (2,945 bp)、*gDFR-WW5f2* (2,945 bp)、*gDFR-WW5f3* (2,943 bp) および *gDFR-WW5f4* (2,945 bp)、「Surprise」*gDFR-Su5f1* (2,714 bp)、*gDFR-Su5f2* (2,714 bp) および *gDFR-Su5f3* (2,713 bp) であった。これらの配列を PLACE の Signal Scan Search program および TFSEARCH にて解析し、単子葉植物で検出されている配列を Table 4-2 および Fig. 4-2 に示した。遺伝子発現制御に関わるシス因子の候補として、植物に特異的な転写因子である Dof 結合部位 (Yanagisawa and Schmidt 1999)、植物遺伝子の光応答に関与している GATA および GT-1 の結合部位 (Gilmartin *et al.* 1990、Villain *et al.* 1996)、フェニルプロパノイド生合成に関与する MYB と (Sablowski *et al.* 1994、Tamagnone *et al.* 1998) そのホモログであるトウモロコシ P タンパク質結合部位 (Grotewold *et al.* 1994) が検出された (Table 4-2、Fig. 4-2)。さらに、開始メチオニンからから 13 および 98 bp 上流に基本転写因子の結合モチーフである TATA box [TATA(A/T)A(A/T), Bucher

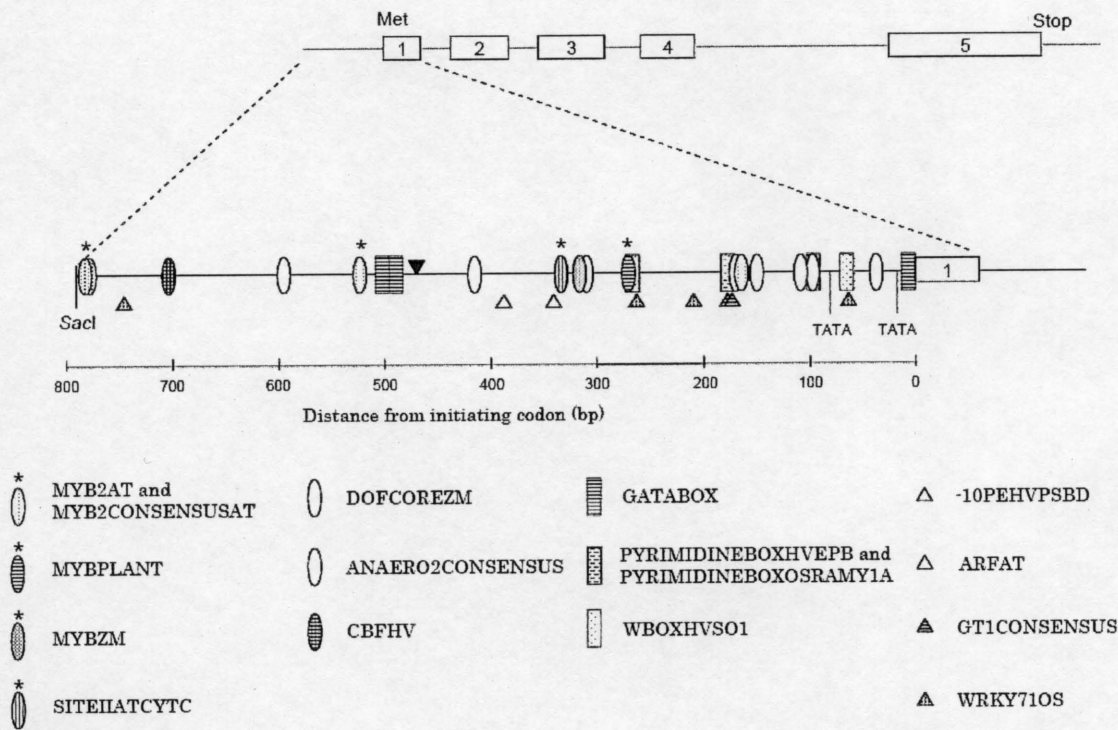


Fig. 4-2. Diagrammatic representation of various putative motifs observed in 5'-upstream region of *IhDFR* gene from *Iris hollandica* cv. Blue Diamond. The black arrow shows retrotransposon insertion site and other symbols indicate the putative motifs. PLACE accession numbers of putative motifs are shown in Table 4-2.



Table 4-2. Cis-element or trans-acting factor binding sites in *IhDFR* promoter

Factor or site name	Loc. <sup>a</sup> (Str. <sup>b</sup> )			Signal sequence	ID
	gDFR-BD1/2	gDFR-WW1	gDFR-Su1		
-10PEHVPSBD	386 (+)	386 (+)	384 (+)	TATTCT	S000392
ANAERO2CONSENSUS	589 (+)/-	-	-	AGCAGC	S000478
	592 (+)/-	-	-		
ARFAT	338 (-)	-	-	TGTCTC	S000270
ARR1AT	438 (+)	438 (+)	413 (+)	NGATT	S000454
BIHD1OS	745 (+)/-	-	-	TGTCA	S000498
CACGTGMOTIF	-	-	222 (+)	CACGTG	S000042
	-	-	222 (-)		
CAREOSREP1	325 (+)	325 (+)	323 (+)	CAACTC	S000421
	407 (-)	407 (-)	-		
CBFHV	699 (+)/-	-	-	RYCGAC	S000497
DOFCOREZM	311 (+)	311 (+)	309 (+)		
	170 (+)	170 (+)	170 (+)		
	96 (+)	96 (+)	96 (+)		
	413 (-)	413 (-)	-	AAAG	S000265
	150 (-)	150 (-)	150 (-)		
	108 (-)	108 (-)	-		
	34 (-)	34 (-)	34 (-)		
GATABOX	490 (+)	490 (+)	465 (+)		
	500 (-)	500 (-)	475 (-)	GATA	S000265
	-	-	460 (-)		
	4 (-)	4 (-)	4 (-)		
GT1CONSENSUS	173 (+)	173 (+)	-	GRWAAW	S000198
	172 (+)	172 (+)	171 (+)		
MYB2AT	777 (+)/-	-	-	TAACTG	S000177
MYB2CONSENSUSAT	777 (+)/-	-	-		
	-	-	58 (+)	YAACKG	S000409
	520 (-)/-	-	-		
	162 (-)	162 (-)	162 (-)		
MYBPLANT	215 (+)	215 (+)	213 (+)	MACCWAMC	S000167
MYBPZM	315 (-)	315 (-)	313 (-)	CCWACC	S000179
PYRIMIDINEBOXHVEPB1	171 (-)	171 (-)	-	TTTTTTCC	S000298
PYRIMIDINEBOXOSRAMY	95 (-)	95 (-)	95 (-)	CCTTTT	S000259
SITEIATCYTC	329 (-)	329 (-)	-	TGGGCT	S000474
TATAPVTRNALEU	13 (-)	13 (-)	13 (-)	TTTATATA	S000340
V\$TATA-C	98 (+)	98 (+)	98 (+)	TATAAAA	M00216
WBOXHVIS01	264 (-)	264 (-)	262 (-)	TGACT	S000442
	61 (-)	61 (-)	-		
WRKY71OS	210 (+)	210 (+)	208 (+)		
	745 (-)/-	-	-	TGAC	S000447
	264 (-)	264 (-)	262 (-)		
	61 (-)	61 (-)	-		

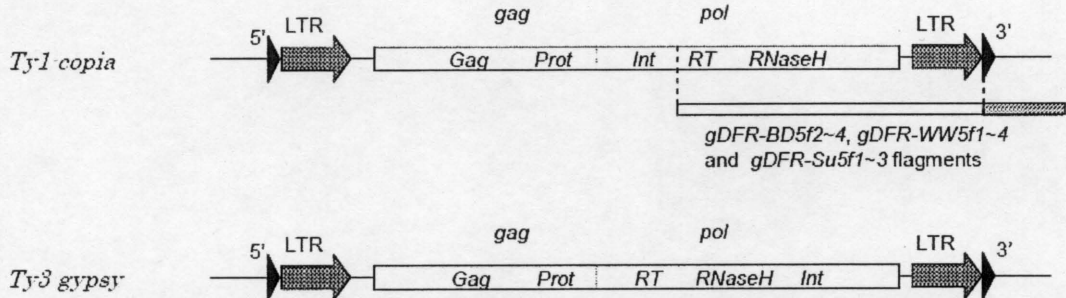
<sup>a</sup> Loc., Distance from Met.

<sup>b</sup> Str., Strand

1990] と推測される領域も見出されたことから、本章で獲得した *gDFR-BD5f1*~4、*gDFR-WW5f1*~4 および *gDFR-Su5f1*~3 はダッチアイリス *DFR* 遺伝子のプロモーター領域をコードしていると思なされる。

獲得した各品種の 5' 上流領域の比較により、*gDFR-BD5f2*~4 および *gDFR-WW5f1*~4 の各クローンでは *gDFR* 配列の開始メチオニンから 513 bp 上流に、*gDFR-Su5f1*~3 では 488 bp 上流に、挿入配列が存在していることが明らかになった (Fig. 4-2)。この挿入配列を BLAST において検索した結果、クラス I 可動性遺伝因子に属する LTR レトロトランスポゾン内に存在する *pol*-polyprotein をコードする遺伝子 (*pol* 遺伝子) と 40 - 50 % の同一性が得られた (以下、上記の挿入配列を *rTih1* と称する)。LTR レトロトランスポゾンは *pol* 遺伝子にコードされる 3 つの酵素の順序およびアミノ酸配列の類似性により、*Ty1-copia* と *Ty3-gypsy* の 2 つのグループに分類されている (Xiong *et al.* 1990) (Fig. 4-3)。本章において単離した挿入配列は、*pol* 遺伝子の構造およびアミノ酸配列の相同性から *Ty1-copia* グループに属するものと考えられる (Fig. 4-3)。そのため、*Ty1-copia* の *pol* 遺伝子内に含まれる逆転写酵素 (RT) について、*rTih1* と高い相同性を示したトマト (*Lycopersicon esculentum*, AAW22873)、タバコ (*Nicotiana tabacum*, P10978) およびイネの RT とのマルチプルアライメントに加え、植物および動物を含めた既知の *Ty1-copia* と *Ty3-gypsy* の RT との分子系統樹解析を行った (Figs. 4-4, 4-5)。マルチプルアライメントにおいて、*rTih1* の RT は他の *Ty1-copia* グループの RT との間で、共通のアミノ酸残基を保存していた。しかしながら、本章で単離された *rTih1* の RT はすべて ORF 内にストップコドンやフレームシフトを含んでいたことから、これらの RT は活性を持たないことが示された (Fig. 4-4)。分子系統樹において *rTih1* の RT は *Ty1-copia* グループに属しており、その中でダッチアイリス単独のグループを形成した (Fig. 4-5)。さらに、「ブルーダイヤモンド」と「White Wedgewood」*rTih1* の RT 配列は近接しており、「Surprise」とは分岐していることが明らかになった。これは、「ブルーダイヤモンド」と「White Wedgewood」の両品種がともに「Wedgewood」に由来することによるものと考えられる (Fig. 2-8)。

### LTR retrotransposons



### Non-LTR retrotransposons

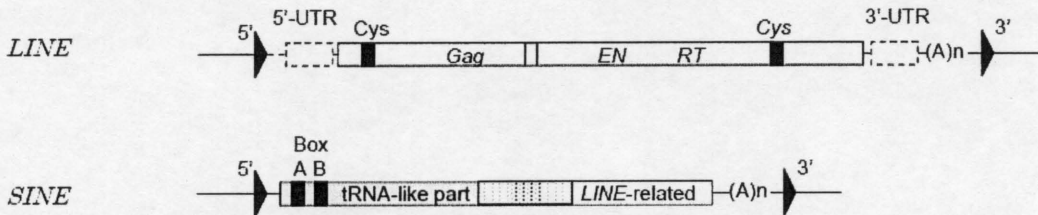


Fig. 4-3. Structure and organization of LTR and non-LTR retrotransposons. Black arrows indicate target site duplication. The LTR retrotransposons (*Ty1-copia* and *Ty3-gypsy* retrotransposons) are both flanked long terminal repeats (LTRs). The LTR retrotransposons encode a nucleocapsid protein (*gag*) with a nucleic acid-binding domain, a protease (Prot), and a polyprotein with reverse transcriptase and RNaseH activity (RT and RNaseH), and an integrase (Int), respectively. The structure of fragments of *gDFR-BD5f2*~4, *gDFR-WW5f1*~4 and *gDFR-Su5f1*~3 of *Iris hollandica* cv. Blue Diamond (BD), cv. White Wedgewood (WW) and cv. Surprise (Su) are presumed from the structure of *Ty1-copia*. The flanking genomic sequence of *gDFR-BD5f2*, *gDFR-WW5f1* and *gDFR-Su5f1* are shown by stripe diagonal box. Non-LTR retrotransposons (LINE and SINE) are terminated by 3'-poly(A) stretch. The LINEs have two ORFs encoding a *gag*, an endonuclease (EN) and RT. Cystein-rich nucleic acid-binding domains (Cys) are shown by black boxes. The light gray box between the *gag* and *EN-RT* domain reflects the overlapping of ORF1 and ORF2. The ORFs are flanked by untranslated regions (UTR), the broken line indicates the 5'-truncations found in many LINEs. SINEs are composed of a tRNA-derivative region (dark gray), an unrelated related DNA sequence (dotted box) and a LINE related region (open box). The two black boxes A and B are similar to RNA polymerase III promoters.

LeRtr	1	AMQEDMSLKKNETYDLVKFPKGGKVLKNRWLFKPKKDGK—KLVKARKARLVVKGCHQKKGDFDEIFAPVVKMTSIRMI
NtRtr	1	AMQEEESLCKNGTYKLVLPKGGKPLKCKWVFKKKDGDG—KLVRYKARLVVKGFEQKKGDFDEIFSPVVKMTSIRTI
OsRtr	1	AMTEEMQSLKKNSTWVFLPKKKKTISCKWIFRKEGLSSSEPPYKARLVARGYSCLPGVDYNDVFSVVKHSSIRTF
rTih-BD2	1	TMEEEEMGSLEKNQWVSLVKLPVEKVKLNK*VIRI KEHDGS—KRYKERLVVKGFCQKRRVDYNKIFAPVVKMTTIRLV
rTih-WW1	1	TMEEEEMGSLEKNQWVSLVKLPVEKKA*LNK*VIRI KEHDGS—KRYKERLVVKGFCQKRRVDYNKIFAPVVKMTTIRLV
rTih-Su1	1	AMEEEEMGSLKKNQWVSLVRLPVGK*ALYNKVVFR I KEHDGN—KRYKTRLV—GFPTKRS—LQDLRSDSEDDYYOIG
LeRtr	79	LGATCLNLELEQLDVKTAFLHGDLEEIYMEQPEGFEVKGKENVCKLKKSLYGLKQAPRWYHKFDSFMSNNEYKRIT
NtRtr	80	LSLAASLDLEVEQLDVKTAFLHGDLEEIYMEQPEGFEVAGKKHVCKLTKKSLYGLKQAPRWYHKFDSFMSKSTYLYKTY
OsRtr	81	LSIVASLDLELEQLDVKTAFLHGLEEDIYMQPEGFIYVPGKEIVCKLKRSLYGLKQSPRWYHKRFSDFMLSHSFKRSK
rTih-BD2	78	LSIVVAEDLHLEQLDVKTFIYGDLEEDIYMAELGFRVAEKENLCKLKKS*YGLKQAPRWYKIFKFSDFMQRNRYRRC
rTih-WW1	78	LSIVVAEDLHLEQLDVKTFIYGDLEEDIYMAELGFRVAEKENLCKLKKS*YGLKQAPRWYKIFKFSDFMQRNRYRRC
rTih-Su1	75	PEYCCRRLASGTIRQDDFLVWRLEEDIYMAQPEGFOVAEKENLVYK*YGLKQAPRWYKIFKFSDFMQRNRYRRC
LeRtr	159	ADPCVYFRKRFSEGNFIIQLLYVDDMLVYGQVEMICRLKEDL*SKSFD*MKDLGPAKQILGMEIARDKAGK*LVLSQENYIE
NtRtr	160	SDPCVYFRKRFSENNFIIQLLYVDDMLVYGQVEMIAKLGKDL*SKSFD*MKDLGPAKQILGMIYRERTSRK*LVLSQEKYIE
OsRtr	161	YDSCVYIKHVN—GSPYLLLYVDDMLIAAKSKIEITKLLKLLSSEFD*MKDLGSAKQILGMEISRDKAGGL*LVLSQENYIK
rTih-BD2	158	MDI*CYLKRFG—SSYIILLLYVEDMLVAGSDMKEIN*KLKRYLSQDFV*MKDLGAAK*IFGMSIIRDKAGGT*LVLSQENYIR
rTih-WW1	158	MDI*CYLKRFG—SSYIILLLYVEDMLVAGSDMKEIN*KLKRYLSQDFV*MKDLGAAK*IFGMSIIRDKAGGT*LVLSQENYIR
rTih-Su1	155	IDRC*CYLKRFG—SSYIILLLYVEDMLVAGSDMKEIN*KLKRYLSQDFV*MKDLGAAKQILGMSIIRDKA*GTLKLVLSQENYIM
LeRtr	239	RVLERFNMKNKPVNTPLAAH—
NtRtr	240	RVLERFNMKNKPVSTPLAGH—
OsRtr	240	KVLCRFNMKNKAVSTPIAPHF—
rTih-BD2	236	KVLEKFSMKDAKSRNTPLS—
rTih-WW1	236	KVLEKFSMKDAKSRNTPLS—
rTih-Su1	234	NVLEKFSMKDAKSRSTPLS—

Fig. 4-4. Multiple alignment of deduced amino acid sequences of rTih (reverse transcriptase region) of *Iris hollandica* cv. Blue Diamond (BD), cv. White Wedgewood (WW) and cv. Surprise (Su) and other plant group II reverse transcriptases. Black shading and gray boxes show identical and similar amino acids, respectively. Genbank accession numbers are as follows: LeRtr (*Lycopersicon esculentum*, AAW22873), NtRtr (*Nicotiana tabacum*, P10978), OsRtr (*Oryza sativa*, ABA98656).

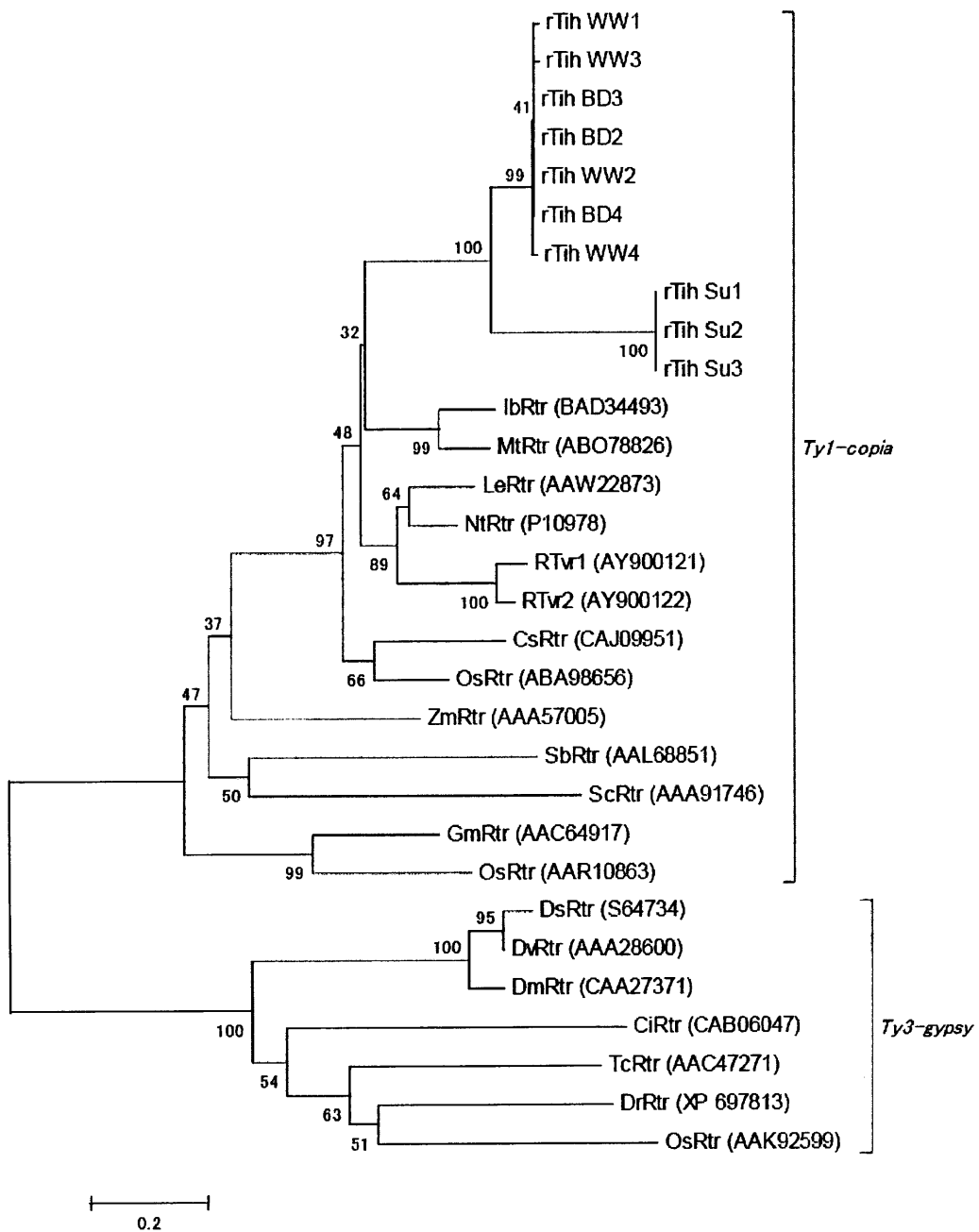


Fig. 4-5. A molecular phylogenetic tree of the deduced amino sequences of rTih (reverse transcriptase region) of *Iris hollandica* cv. Blue Diamond (BD), cv. White Wedgewood (WW) and cv. Surprise (Su) and other plant group II reverse transcriptases. The tree was constructed by the neighbor-joining method. The numbers at the branches are the bootstrap values that indicate confidence level each branch based on 1000 repetitions of the analysis. The lengths of the lines indicate the relative distances between nodes.

```

gDFR-BD5f1 -513: AGCTAGTAGCTATCAAATTTGATAGTATGAAGTTAAACATTTTGAATAGTTAAGTGCATCAAGTTACAATTAGATTAAC -434
gDFR-BD5f2 -513: AGCTAGTAGCTATCAAATTTGATAGTATGAAGTTAAACATTTTGAATAGTTAAGTGCATCAAGTTACAATTAGATTAAC -434
gDFR-WW5f1 -513: AGCTAGTAGCTATCAAATTTGATAGTATGAAGTTAAACATTTTGAATAGTTAAGTGCATCAAGTTACAATTAGATTAAC -434
gDFR-Su5f1 -488: AGCTAGTAGCTATCAAATTTGATAGTATCAAGTTAAACATTTTGAATAGCTAAGGACATCAAGTTACAATTAGATTAAC -434
*****

gDFR-BD5f1 -433: ATATATTGAAGTACTCAGTTGAGTTGTGTTCAATATTGGACTATTCTCACTAGCAAGCTTCAACACAAGTAGTTCATAC -354
gDFR-BD5f2 -433: ATATATTGAAGTACTCAGTTGAGTTGTGTTCAATATTGGACTATTCTCACTAGCAAGCTTCAACACAAGTAGTTCATAC -354
gDFR-WW5f1 -433: ATATATTGAAGTACTCAGTTGAGTTGTGTTCAATATTGGACTATTCTCACTAGCAAGCTTCAACACAAGTAGTTCATAC -354
gDFR-Su5f1 -408: ATATATTGAAGTACTCA-----ACTATTCTCCTTGGCAAAGTTCAGCACAAGTAGTTCATAC -352
*****

gDFR-BD5f1 -353: GGTTAGAGTGGAGACATAAGGCCCAACTCATATGGTTGGAAGCCAGCAACGACATCACCTATACAACCTAATTGGAAC -274
gDFR-BD5f2 -353: GGTTAGAGTGGAGACATAAGGCCCAACTCATATGGTTGGAAGCCAGCAACGACATCACCTATACAACCTAATTGGAAC -274
gDFR-WW5f1 -353: GGTTAGAGTGGAGACATAAGGCCCAACTCATATGGTTGGAAGCCAGCAACGACATCACCTATACAACCTAATTGGAAC -274
gDFR-Su5f1 -351: GGTTAGAGTGGAGACGTAAGACCAACTCAATGGTTGGAAGCCAGCAACGACATCACCTATACAACCTAATTGGAAC -272
*****

gDFR-BD5f1 -273: TTGCCAGTCAATGCGCCTACTAATTTGGTGTCTTCTGCGAGGATGCACGCGCCACCAAACATGACGAATGGCGAGTCTGGA -194
gDFR-BD5f2 -273: TTGCCAGTCAATGCGCCTACTAATTTGGTGTCTTCTGCGAGGATGCACGCGCCACCAAACATGACGAATGGCGAGTCTGGA -194
gDFR-WW5f1 -273: TTGCCAGTCAATGCGCCTACTAATTTGGTGTCTTCTGCGAGGATGCACGCGCCACCAAACATGACGAATGGCGAGTCTGGA -194
gDFR-Su5f1 -271: TTGCCAGTCAATGCGCCTACTAATTTGGTGTCTTCTGCGAGGATGCACGCGCCACCAAACATGACGAATGGCGAGTCTGGA -192
*****

gDFR-BD5f1 -193: ATGGGAATGGTAGAGGGAAAAAGAGCAGTTGTGTCTGCTCTTTTGAAGAGGGGAACTCGTCTCTTAGCTGTTGGTAC -114
gDFR-BD5f2 -193: ATGGGAATGGTAGAGGGAAAAAGAGCAGTTGTGTCTGCTCTTTTGAAGAGGGGAACTCGTCTCTTAGCTGTTGGTAC -114
gDFR-WW5f1 -193: ATGGGAATGGTAGAGGGAAAAAGAGCAGTTGTGTCTGCTCTTTTGAAGAGGGGAACTCGTCTCTTAGCTGTTGGTAC -114
gDFR-Su5f1 -191: ATGGGAATGGTAGAGGGA—AAAGAGCAGTTGTGTCTGCTCTTTTGAAGAGGGGAACTCGTCTCTTAGCTGTTGGTAC -114
*****

gDFR-BD5f1 -113: CTCTTTCCGTATAAAAAGGACCTCCATCATAAACCATTTGCTTACACTAGTCACTGCCACACCTATACTACAATTTCTTT -34
gDFR-BD5f2 -113: CTCTTTCCGTATAAAAAGGACCTCCATCATAAACCATTTGCTTACACTAGTCACTGCCACACCTATACTACAATTTCTTT -34
gDFR-WW5f1 -113: CTCTTTCCGTATAAAAAGGACCTCCATCATAAACCATTTGCTTACACTAGTCACTGCCACACCTATACTACAATTTCTTT -34
gDFR-Su5f1 -113: CTCTTTCCGTATAAAAAGGACCTCCATCATAAACCATTTGCTTACACTAGTCACTGCCACACCTATACTACAATTTCTTT -34
**

gDFR-BD5f1 -33: CTTCTTGAGATCATATATAACAATATATCGAGATGATGAGCCCGTTGTCGTGACCCGAGCGAGCGGCTACGTCGGTTC 47
gDFR-BD5f2 -33: CTTCTTGAGATCATATATAACAATATATCGAGATGATGAGCCCGTTGTCGTGACCCGAGCGAGCGGCTACGTCGGTTC 47
gDFR-WW5f1 -33: CTTCTTGAGATCATATATAACAATATATCGAGATGATGAGCCCGTTGTCGTGACCCGAGCGAGCGGCTACGTCGGTTC 47
gDFR-Su5f1 -33: CTTCTTGAGATCATATATAACAATATATCGAGATGATGAGCCCGTTGTCGTGACCCGAGCGAGCGGCTACGTCGGTTC 47
*****

```

Fig. 4-6. The nucleotide sequence of gDFR promoter region and its downstream sequences of *Iris hollandica* cv. Blue Diamond (BD), cv. White Wedgewood (WW) and cv. Surprise (Su). Identical nucleotides are indicated by asterisks and gaps in the alignment by dots. The ATG initiation codon of *DFR* is underlined. TATA boxes are indicated by double underlined.

ダッチアイリスの *DFR* 遺伝子のプロモーター領域への *rTih1* の挿入の影響は、次のように考えられる。この *rTih1* の挿入位置よりも上流には、*bronze-1* (トウモロコシのアントシアニン生合成に関与) の発現を誘導する MYB2AT タンパク質の結合部位が位置しているため (Fig. 4-2、Urao *et al.* 1993) その MYB2AT 結合部位と *DFR* 構造遺伝子の配列が *rTih1* の挿入によって分断された結果、*DFR* の発現の欠損または減少がもたらされたと考えられる。すなわち、「White Wedgewood」や「Surprise」はプロモーター領域に *rTih1* の挿入した *DFR* 遺伝子のみを持つため、「White Wedgewood」ではアントシアニン生合成の欠損あるいは著しい低下により内、外両花被および雌蕊の白色化、また「Surprise」ではアントシアニン生合成の著しい低下により外花被および雌蕊の白色化がもたらされ、一方「ブルーダイヤモンド」は *rTih1* が挿入した *DFR* と *rTih1* が挿入していない *DFR* の両方を有するため、アントシアニン生合成が正常に機能して青紫色花を発現したものと見なされる。さらに、3品種の *rTih1* 挿入配列より下流のプロモーター領域の比較により、「Surprise」において「ブルーダイヤモンド」や「White Wedgewood」とは異なる配列が見出され (Fig. 4-6)、転移因子の結合部位も異なることが示唆される (Table 4-2)。今後、これらの知見をもとにプロモーターおよび転写調節因子の解析を行うことにより、器官ごとに異なるアントシアニン生合成の制御機構が明らかされ、第2章・第2節で示された「White Wedgewood」および「Surprise」の花器官間で *DFR* 遺伝子の発現程度が異なる原因も究明されるであろう。

#### 4. 摘要

「White Wedgewood」および「Surprise」の両品種における *DFR* 遺伝子発現の欠損もしくは減少を解明する目的で、両品種に「ブルーダイヤモンド」を加えた3品種を供試した。3品種のゲノムDNAを鋳型として nested PCR および Inverse PCR を行い、*DFR* 構造遺伝子とその5'上流領域配列を単離して、それらの配列の比較解析を行った。その結果、*DFR* 構造遺伝子が5つのエクソンからなる 1,086 bp の ORF と、4つのイントロンから構

成されていること、その内部に発現の欠損につながるナンセンス変異やフレームシフト変異を持たないことが明らかにされた。さらに、5'上流領域の解析を行った結果、TATA box やトウモロコシのアントシアニン生合成に関与する P 因子の結合部位と類似した領域が見出されたことから、得られた配列はプロモーター領域であることが推測された。次に、供試した 3 品種の *DFR* 遺伝子のプロモーター領域の配列を詳細に調査したところ、*Ty1-copia* 型 LTR レトロトランスポゾンの挿入が見出された。また、「ブルーダイヤモンド」はプロモーター領域にレトロトランスポゾンが挿入した *DFR* と挿入していない *DFR* の両方を有していたが、他の 2 品種はレトロトランスポゾンの挿入を持つもののみが検出された。このことから、「White Wedgewood」および「Surprise」における *DFR* 遺伝子発現の欠損または減少はレトロトランスポゾンに起因することが指摘された。