

第2章 第1節

カンキツゲノム解析におけるESTの解析とその利用

静岡大学名誉教授 大村 三男

1) EST解析

発現遺伝子配列断片 (expressed sequence tag, EST) は, その解析の対象とする塩基配列がゲノムから転写されるmRNAに由来することから, その網羅的な解析プログラムはゲノム解析におけるアプローチの1つとして, 低コストで遺伝子にたどり着く方法として広く採用される.

1980年代後半からヒトゲノム解析プロジェクト (Watson 1990) で, 全ゲノム塩基配列解読が進められたが, 全塩基配列の決定は開始後5年を経過した当時においても, さらに12~15年を要する予測された (Baltimore 1988). 当初予測されていたゲノム中の5万~10万の遺伝子のコード配列領域はゲノムDNAの3%にすぎないと見積もられていたことから, ゲノム解析を効率よく進めるためには, 遺伝子の転写産物であるmRNAに基づいてcDNAの塩基配列を決定することが早道であるとの考えが提案された (Brenner 1990). このような時代背景のもとに, cDNAクローンの部分配列を解析することにより発現遺伝子を包括的に検索することが, 遺伝子にたどり着くタグとして, どのように使用できるのかを検定するパイロット研究が行われた (Adams *et al.* 1991). この研究は, ヒト脳から得たcDNAクローン600あまりについて部分配列を決定し, EST解析の最初の事例として報告された. また, 読み取ったcDNA配列が150塩基以上であればデータベースに対して相同遺伝子の検索に利用できる

こと、その結果337が新規な遺伝子であったこと、さらに得られた配列がヒト染色体地図 (chromosome map) にマッピングできることが明らかにされた。この論文では、「ESTとは、mRNAに対応するcDNAクローンの単なる部分配列である」とだけ記しているが、ゲノム中の遺伝子を網羅的に解析することで、ゲノムの機能解明にアプローチする手段として重要であることを確信させた。ほぼ同時期に、マウス (*Mus musculus* L.) や線虫 (*Caenorhabditis elegans*) においても、一連の異なる組織や異なる発育段階の試料由来のmRNAから作製したcDNAクローン対象にEST解析が開始された (Höög 1991, McCombie *et al.* 1992)。これらの取り組みから、EST解析が網羅的に遺伝子情報を獲得する基盤になることが、生物科学に関わる研究者に認知されてきた。

EST解析は植物のゲノム解析においても取り入れられた。イネ (*Oryza sativa*) では、懸濁培養細胞に由来する830cDNAクローンの塩基配列を解析された。公的データベース上に登録されたDNA配列に対して相同性を検索すると、約8%のEST配列が既知遺伝子と相同性を示し、140の遺伝子が同定された (Uchimiya *et al.* 1992)。イネEST配列は、制限酵素断片長多型 (restriction fragment length polymorphism, RFLP, Botstein *et al.* 1980) による連鎖地図 (linkage map) 作製におけるcDNAプローブとしても利用された。その後、EST解析は、大規模化 (Sasaki *et al.* 1994) が図られるとともに、シロイヌナズナ (*Arabidopsis thaliana* (L.) Heynh.), *Brassica*属など、多様な植物種での解析が行われた (Höfte *et al.* 1993; Park *et al.* 1993)。これらの初期のEST解析の論文からは、cDNAの部分配列がもたらす情報として、(i) 遺伝子を染色体地図 (chromosome map) 上あるいは物理地図 (physical map) 上に位置づける、(ii) エクソン-イントロン構造 (exon-intron structure) を知る、(iii) プロモータ (promoter) 配列など遺伝子発現制御情報のてがかりを得る、(iv) 配列から遺伝子機能を推察する、(v) 発現部位・時期など遺伝子の発現特異性の把握する、などへの期待があったと推察される。

本稿では、我が国におけるカンキツEST解析の取り組みがどのような背景で開始され、ゲノムの解析や育種の中でどのように貢献しようとしたのか、実例

を中心に紹介する。

2) カンキツにおけるEST解析

(1) 研究背景

カンキツ (*Citrus*) においても、遺伝育種研究などの底上げに関わる基盤として、1987年度からカンキツのゲノム解析に取り組むプログラムが検討され始めた。このような取り組みの背景としては、(i) イネなどでRFLPによる連鎖地図の作製と高密度化に関する研究が進み、多様な利用法が考えられるようになってきたことから、カンキツ育種における選抜マーカー (selection marker) への期待が高まり、(ii) 遺伝資源多様性解析の研究として、DNAマーカーの適用実績が蓄積され始めた、そして、(iii) バイオテクノロジーとくに遺伝子導入技術が進展し、重要でカンキツ独自の遺伝子が求められたこと、などをあげることができる。これら多岐にわたる目的を効率よく達成するため、育種目標 (breeding objective) に関わる形質 (trait) を分子生物学的・生理学的に解析するための遺伝子配列の確保や、目的形質との連鎖関係にある早期選抜マーカーの獲得などの研究をできるだけ統合された方針で取り組むことが有利と考えられた。ここでは、カンキツ果実の結実・発育・品質などの形質解析に重点をおくことで、果実から作製するcDNAライブラリーを、遺伝子単離・機能解析や発現解析に適用するとともに、マーカー作製などに共通して利用し、さらに、それらの解析結果を相互に関連づけする研究基盤としてEST解析が好適と期待された。

(2) カンキツにおけるEST解析の開始と進捗

EST解析を進めるに当たって、どの品種から、またどのような器官・時期を解析対象とするかの課題に加えて、質的に適正なcDNAライブラリーを作製すること、さらにはEST解析をルーチン化することが、当面の課題となった。‘バレンシア’ オレンジ (*Citrus sinensis* (L.) Osbeck.) の開花2か月後の初期球状胚期の胚珠 (ovule) に由来するcDNAライブラリー (Koltunow *et al.* 1996)

が最初のEST解析の対象となった。この解析では、胚形成 (embryogenesis) に関わって発現する遺伝子レパートリーが得られるものと期待された。初回の解析はわずかに192クローンの解析であったが、そのうち74 (38.5%) が植物遺伝子に相同性を示した。胚形成との直接的関連を示唆する遺伝子は得られなかったが、成分代謝や果実発育と関連する遺伝子など、想定以上に幅広い遺伝子のレパートリーが同定された (Hisada *et al.* 1996)。

次に、果実において発現する遺伝子レパートリーの確保を目的として、果実発育期の砂じょう (juice sac) 由来のライブラリーが解析された。ここでは解析規模を1,000以上とし、大量解析に向けての効率化における問題も同時に抽出・検討した (Hisada *et al.* 1997)。その後、成熟期果肉、剥皮開始時期果皮、開花時期の子房、着色期果皮、開花1月前の頂芽 (apical bud)、開花1月後の幼果などに由来するcDNAライブラリーを順次解析した (表2-1-1)。このような解析からは、ルーチンとして行う相同性検索においてインターネットの通信環境の改善と、検索パイプラインの強化の課題が浮き彫りにされるとともに、情報共有化のあり方やより幅広く利用するためのデータ公開の進め方についての検討も行われた。そして、カンキツにおけるEST情報の管理に必要なデータベースシステムが検討され (藤井ら 2003)、解析からデータの利用、管理までの一連の取り組みが開始された。基本方針として解析データを公開することとし、インハウスでのデータベース化とともに、ESTの公的データベースへの登録を行った。10年間のプログラムで、主として‘宮川早生’を対象とした19のcDNAライブラリーから合計29,228クローンを解析し、20,525配列を公的データベースに登録した (Fujii 2013)。この解析は、海外各チームの解析をも促進し、dbEST (URL2-1-1) に登録されたカンキツ属EST 総計は2013年1月時点で478,874件に達している。

3) ESTカタログの解析と利用

cDNAライブラリー作製の目的の1つが、カンキツの育種形質の生理的解析のための遺伝子検索・利用であることから、ESTカタログ化は機能性成分代謝、

表2-1-1 カンキツにおける果樹研究所解析のESTカタログ

解析年	cDNAライブラリー				解析 クローン数	相同遺伝 子の 出現率 (%)	二次 代謝 遺伝子 数
	名称	品種	組織・器官	発育時期			
1994	VSS	バレンシアオレンジ	幼種子	受粉1月後	577	50.5	4
1995	FRI	宮川早生	果肉	肥大期	1,051	32.7	4
1996	FRM	宮川早生	果肉	成熟期	385	52.1	8
1997	ALM	宮川早生	アルベド	成熟期	623	41.7	13
1998	OVA	宮川早生	子房	開花期	827	46.3	28
1999	ALP	宮川早生	アルベド	剥皮開始期	941	58.1	13
2000	WFY	宮川早生	全果	幼果期(1か月)	1,689	52.2	32
2001	BFC	宮川早生	果皮	着色期	1,650	49.3	60
2002	FBI	宮川早生	花芽	開花1月前	2,367	—	—
2003	GSA	宮川早生	種子	吸水4日	1,920	—	—
2003	RGP	宮川早生	根	発芽2週後	960	—	—
2003	SLG	宮川早生	苗条	発芽2週後	1,920	—	—
2003	YJS	宮川早生	砂じょう	開花60日後	1,926	—	—
2003	PCC	宮川早生	カルス	増殖培養	960	—	—
2003	EIC	宮川早生	カルス	胚誘導培養	1,152	—	—
2003	STG	宮川早生	柱頭	開花期	3,552	—	—
2003	ANT	宮川早生	葯	開花期	2,600	—	—
2003	LLL	リスボンレモン	葉	幼葉期	2,016	—	—
2003	EGJ	キシウミカン	胚珠	開花60-70日後	2,112	—	—

形態形成などに関わる遺伝子の単離・機能解析とも連携して進められた。EST解析で得られた遺伝子レポトリとして、豊富な二次代謝関連酵素遺伝子がリストアップされ (Omura *et al.* 2000, 表2-1-1), 着色期果皮や開花1か月の幼果に由来するcDNAライブラリーからは、フラボノイド (flavonoid), カロテノイド (carotenoid), テルペノイド (terpenoid) などの成分代謝に関わる遺伝子群が検出され、成分の代謝や蓄積などの研究に応用された。また、胚形

成・発育，花芽誘導（flower bud initiation）・器官発達などの関連する遺伝子も，開花1か月前の花芽や開花時期子房から多数検出された。

果実の発育時期の異なるcDNAライブラリー間のカタログ比較では，遺伝子発現の様相変化と関連して，検出される遺伝子レパートリーや重複検出される遺伝子に特徴が認められた。ESTの解析初期の研究においても，相同性検索により同一モチーフ（motif）をもつEST群が検出されることが指摘され（Adams *et al.* 1991），同一遺伝子に由来するESTが重複して同定されることは冗長性（redundancy）と呼ばれた。冗長性は，遺伝子を網羅的に同定・確保するという意味ではEST解析の効率を低下させるが，高度な冗長性を示すESTは，解析対象の組織あるいは特定の発現段階で発現性の高い遺伝子という可能性がある。‘宮川早生’果実の肥大期，成熟期など発育段階の比較解析では，冗長な遺伝子もそれぞれに特徴的であり（Moriguchi *et al.* 1998），果実成熟期に冗長に検出されるESTからメタロチオネイン（metallothionein）様遺伝子が同定された。また，その5'上流領域が果実成熟時期の高発現を制御するプロモータ（promoter）として単離された（Endo *et al.* 2007）。

4) ゲノムマッピングにおけるEST配列の利用

EST解析はゲノム中に散在する遺伝子から転写されたmRNAを解析の対象とする。そのため，EST配列からはゲノム配列中のコード領域を確認することができる。したがって，EST配列を対象としたDNAマーカーを作製・利用することによって，EST座位相互の連鎖関係を明らかにするばかりでなく，全ゲノム解析によってもたらされるDNA配列上でもESTが座乗する物理的位置を知ることができる。ゲノム上の特定領域を対象とするDNAマーカーは，配列タグ部位（sequence tagged site, STS）マーカーと呼ばれる。このようなマーカーの遺伝地図上での位置は，基本的に品種を越えて同一連鎖群の同一部位に存在するため，異なる分離集団や異なるモードで解析された連鎖地図間の相互比較や統合化（Omura *et al.* 2003a），物理地図との照合などに利用できる。

(1) DNAマーカー開発

1990年代初頭、急速にDNAマーカーの開発・普及が進み、カンキツにおいてもDNAマーカーによる遺伝地図の作製が開始された。DNAマーカーとして最初に用いられたのは、サザン分析の応用技術としてのRFLP法である。米国においても初期のカンキツDNAマーカー地図作製には、California大学Riverside校で作製したcDNA及びゲノムDNAクローンが、RFLPのプローブとして用いられ、カンキツ最初の遺伝地図が作製された (Jarrell *et al.* 1992 ; Durham *et al.* 1992)。

しかし、RFLPは技術的に煩雑で、スループットも低いことから、DNAマーカーの作製はPCR法を利用する技術へと急速に移っていった。生食用カンキツ品種間においても多型出現率の高いランダム多型増幅 (Random Amplified Polymorphic DNA, RAPD) 型のDNAマーカーがマッピングに利用された (Omura *et al.* 1993) が、このタイプのマーカーは遺伝的に優性マーカー (dominant marker) であることや異なる分離集団間での共有化しにくさがあることなどの理由から、STS化した共優性マーカー (codominant marker) の開発へと重点が移動した。特に、特定の塩基配列に基づいてSTSプライマーを作製し、PCR増幅後、制限酵素 (restriction endonuclease) による多型 (polymorphism) を検出する手法は、cleaved amplified polymorphic sequence (CAPS) と呼ばれ、広く利用されるDNAマーカーとなった (Konieczny and Ausubel 1993)。

カンキツのショ糖合成酵素遺伝子 (sucrose synthase, 小松ら 1994) では、カンキツゲノムに存在する3つのアイソフォーム (isoform) 座のそれぞれをPCR増幅し、座位ごとの制限酵素多型を検出できることが示された。この結果は、ウンシュウミカン (*Citrus unshiu* Marcow.), スイートオレンジなど生食用品種間で遺伝子領域をSTSとしたとき、PCRを用いて効率的にCAPSが検出でき、またESTの配列情報から座位特異性の高いSTSプライマー設計ができることを示した。カンキツのEST配列に基づいたCAPSマーカー化は、RFLPによる遺伝地図作製に利用されたラフレモン (*Citrus jambhiri* Lush.) cDNA及び

gDNAクローン (Jarrell *et al.* 1992) を用いて試行され、いくつかのマーカーについて、米国の遺伝地図との共有化に成功した。

ところで、カンキツゲノムサイズは、約360 Mb ($n=9$) と予測され、連鎖地図における遺伝距離総長は、当初1,700 cM (Jarrell *et al.* 1992) と見積もられた。マーカー間の平均遺伝距離を5cM間隔と設定した場合、連鎖地図作製に必要なマーカー座位 (marker locus) 数は340になることから、効率的で系統的な多型検出手順の確立が必要とされた。ゲノム塩基配列情報がなく、EST配列のみの情報からプライマー設計した場合、カンキツゲノムを鋳型 (template) としたPCRで安定的に増幅産物が得られる割合は概ね60%であった (表2-1-2)。また、‘清見’ と ‘宮川早生’ の間でSTSに多型の得られる割合は59%で、対象とした全STSの35%となった。この多型には制限酵素処理を要さないSTS断片の有無、あるいは配列反復回数差に基づくものがそれぞれ3.9%および1.4%混在した。しかし、多型マーカーの95%近くが制限酵素処理後に検出されることから、EST配列に基づいてCAPSを系統的に探索するのがマーカー取得の手順として適切と考えられた。

カンキツEST配列からSTSプライマー設計する場合には、ゲノム配列におけるエクソン-イントロンの構造に関する知見がないことと、PCR産物長を直接予測することができない点での難点があり、それがSTS化率を低下させている要因となっていた。この対策として、カンキツEST配列が由来するSTS領域におけるエクソン-イントロン構造の予測に基づいたプライマー設計を試みた。カンキツESTに相同なシロイヌナズナの遺伝子のゲノム構造を参照して、カンキツEST配列でのイントロン挿入部位を推定した。そして、そのイントロンを跨ぐエクソン配列にプライマーを設計することで、変異が多いと期待されるイントロン配列を含んだゲノム配列を増幅対象とするイントロンスパンニング (intron-spanning) PCR法を適用した。エクソン上に確実にプライマーを設計し、同時にPCRに適した産物長を予測することで、STS化率及びCAPS検出率を高めることを目的として、自動化プログラムMAEZATO (Fujii *et al.* 2010) を適用した。

表2-1-2 カンキツEST及び遺伝子cDNA配列に基づいたDNA多型の検出程度 (Omura *et al.* 2003)

cDNAライブラリー	STS化			清見と宮川早生間における多型のタイプ				
	試行数	PCR増殖数	同 (%)	CAPS	SSR ^y	STS ^x	合計 (多型STS数)	対全試行数 (%)
UCR ^z	14	4	28	1	0	0	1	7
VSS	18	11	61	6	0	0	6	33
MRF	84	60	71	28	2	3	33	39
ALM	214	122	57	64	1	1	66	31
OVA	119	79	66	55	0	3	58	48
ALP	95	47	49	25	0	0	25	26
各種遺伝子	23	20	87	12	0	1	13	56
計	567	343	60	191	3	8	202	35

^zCalifornia大学 Riverside校より分譲のラフレモン由来cDNA (Jarrell *et al.* 1992)

^y増幅領域内の単純反復配列回数に生じた多型

^xPCR増幅の有無による多型

(2) CAPSの遺伝地図作製上の特性

連鎖関係を利用したカンキツ育種形質のマッピングと連鎖マーカーによる早期選抜技術を確立して、時代のニーズに即応した品種育成の効率化を図ることが、カンキツゲノム解析における遺伝地図作製の目標の1つであった。生食用品種の育成に関わるマーカーへの要望が強いものの育種目標が幅広いため、その目的に合致したマッピング集団の選定が重要であり、それによってマーカー作製と対象形質マッピング成否が関わると予測された。当初の解析集団として、果皮の厚さ、寛皮性、果皮及び果肉の色、糖度、酸含量、種子数などの形質の分離 (口絵2-1-1) が期待される‘清見’×‘宮川早生’後代135個体を主な解析集団として養成した。また、DNAマーカーの遺伝分離 (genetic segregation) を確認するための補助集団 (‘清見’×‘トロビタ’オレンジ, ‘寿太郎’ウンシュウ×‘トロビタ’オレンジなど) も対象とした。その後、育種

世代の進展に伴い、高糖度の期待される‘興津46号’ (A255) と優性の無核性 (seedlessness) 遺伝子をもつ ‘中間母本農5号’ (G434) との交雑集団 (hybrid population) (以下、この系間の交雑集団をAG集団と略す) などを加えての遺伝地図作製が行われた。

カンキツの品種は、一般にゲノム中の多くの領域でヘテロ接合型 (heterozygous) である。そのため、品種間の単交雑を主体とする育種実生集団を対象とした遺伝地図作製において、安定した分子マーカーを大量作製することとともに、予測される多様な分離パターンに適用できる連鎖解析の方法を確立する課題もあった。カンキツでは単交配集団においても、各STS座について最大4つの対立遺伝子 (アレル, allele) の組み合わせで後代の遺伝子型が構成される。両親に由来する4つすべてのアレル (母親由来のアレルをa, b, 父親由来をc, dとする) が識別できれば、メンデル型遺伝では $ab \times cd$ 型の分離モードとなり、その子供の遺伝子型はac, ad, bc, bdのいずれかとなる。各遺伝子座がすべてこの分離モードで解析できるマーカーが得られれば、遺伝子座間の連鎖解析は単純なテスト交配型の計算となり、各座位は容易に遺伝地図に座乗される。しかしながら、EST配列に基づいて作製するCAPSマーカーでは、STS座位当たりの多型検出頻度が低いことから、多くの座位では一方の親に由来するヘテロ接合型だけが識別できる。後代集団での分離モードのほとんどはSTS座位ごとに $aa \times ab$ または $ab \times aa$ の BC_1 型であり、一部が $ab \times ab$ の F_2 型分離モードとなる。3アレルおよび4アレル型のSTS座の検出頻度は低い。したがって、STS座位ごとの分離型が均一化されたテスト交配の連鎖解析モデルではなく、優性型のマーカーの場合と同じように、擬似テスト交配 (pseudo-testcross, Carlson *et al.* 1991) 法を適用することが多い。この方法では、分離型によりマーカーを区分し、 BC_1 型分離を示すマーカーにより片親ごとの連鎖地図を作製する。そして、 F_2 型など両親に共通するマーカー座位を参照して、各種のマーカー座位を統合連鎖地図に集約する。

これらの分離モードの出現程度を、‘清見’ × ‘宮川早生’ 及び ‘寿太郎’ ウンシュウ × ‘トロビタ’ オレンジの2集団によって、検定した (表2-1-3)。

表2-1-3 交雑集団におけるCAPS遺伝子型分離様式

親品種遺伝子型(CAPS)			CAPS分離型				CAPS 座数
トロビタ オレンジ (Tr)	清見(Km)	ウンシュウ (Mg ² /Jt ³)	Km × Mg 集団		Jt × Tr ³ 集団		
			分離型	分離親	分離型	分離親	
ab	ab	aa	ab × aa	Km	aa × ab	Tr	65
ab	ab	ab	ab × ab	両親	ab × ab	両親	8
ab	aa	ab	aa × ab	Mg	ab × ab	両親	11
ab	aa	aa	aa × aa	非分離	aa × ab	Tr	26
aa	ab	ab	ab × ab	両親	ab × aa	Jt	20
aa	aa	ab	aa × ab	Mg	ab × aa	Jt	17
bb	ab	aa	ab × aa	Km	aa × bb	非分離	9
bb	ab	ac	ab × ac	両親	ac × bb	Jt	1
ac	ab	ab	ab × ab	両親	ab × ac	両親	1
ac	aa	ab	aa × ab	Mg	ab × ac	両親	2
cc	ac	ab	ab × ac	両親	ab × cc	Jt	2

²宮川早生 (Mg), ³寿太郎 (Jt), ³トロビタオレンジ (Tr)

162のCAPS中、ウンシュウミカン（‘宮川早生’，‘寿太郎’，ウンシュウ），スイートオレンジ（‘トロビタ’ オレンジ），‘清見’の間では基本的に2つのアレルだけで構成されるものが96.2%を占め，3アレルの検出されたものは僅かに6点で，4アレルのCAPSは検出されなかった．‘清見’ × ‘宮川早生’では，両親の遺伝子型に従って，aa × ab，ab × aa，ab × abの2アレル型分離モード及び少数の3アレル型各種の分離モードとなった．その結果，この後代集団で種子親品種‘清見’から分離するCAPSはBC₁型およびF₂型のモードあわせて106（65.4%），花粉親品種‘宮川早生’からは62（38.2%）で，それぞれの親品種の連鎖地図にマップされた．このうち両者にマップされるab × ab及び3アレルの分離モードを示す32を共通マーカーとして，それぞれの遺伝地図が比較，統合された．また同じCAPSを‘寿太郎’ウンシュウ × ‘トロビタ’

オレンジの集団に適用した場合、各親において分離するCAPSは、それぞれ62 (38.2%) 及び113 (69.8%) で、うち共通マーカーは22となった (Omura *et al.* 2003a). また、2集団のうち一方にだけマップされ、他の集団では分離しないものも出現し、これらは組み合わせの遺伝的特性との関係で、遺伝地図の完成度を点検する際の情報となった。

(3) 転写因子様遺伝子候補配列に基づくDNAマーカーの作出と統合連鎖地図の作製

遺伝分離様式が明瞭な質的形質 (qualitative trait) はDNAマーカーと同じ方法で連鎖地図上にマップすることができる。量的形質 (quantitative trait) についても、QTL (quantitative trait loci) 解析を適用して、その量的形質を制御する遺伝子の座位を推定することができ、量的形質に関わる遺伝子の単離や育種に関して選抜マーカーの作出が可能になっている。また、遺伝子発現強度を量的形質としてQTL解析を行うeQTL (expression QTL) 解析 (Jansen and Nap 2001) では、検出されたQTL領域にある転写制御 (transcriptional regulation) に関わる遺伝子を選抜することができる。この目的のためには、遺伝地図上の全領域をカバーするように多数の転写因子様遺伝子を配置することで、eQTL解析が示すゲノム領域から転写因子遺伝子の同定、単離できるように備える (Keurentjes *et al.* 2007)。カンキツの成分代謝酵素遺伝子の転写制御に関わる要因の解析を目的に、転写因子様遺伝子配列を対象にしたDNAマーカー作製を事例にして、多型検出の詳細化を検討する。

シロイヌナズナゲノムの転写因子様遺伝子として注釈される配列を質問 (クエリー, query) 配列として、カンキツESTから456の代表配列を選抜した。このうち、380 (88%) でカンキツゲノムDNAに対するPCR増幅が達成され、STS化に成功した。この時、ファミリー遺伝子の存在が予測される転写因子遺伝子群に対してもシロイヌナズナゲノム構造を参照するMAEZATOの適用がSTS化に有効であった。これらのSTSの平均増幅断片長は988 bで、16種の制限酵素による総認識サイトはSTS当たり平均8.3カ所であった。そのうち199

STS (54.3%) ではA255とG434との系統間に多型が検出された。これらのSTS座には、複数の制限酵素で多型を示すものが52存在し、全380STSでのCAPS検出率は、STS当たり0.83であった。しかし、複数の制限酵素で多型を示す52STSのうち、両親の連鎖地図のそれぞれにマッピングできるCAPSをもつものは6つに過ぎなかった。A255及びG434間で検出された多型のうち、後代で非分離型となるものや安定性を欠いたSTSを除いて、最終的に116の転写因子様遺伝子を対象としたSTSがCAPSマーカーとして遺伝地図に位置づけられた (Shimada *et al.* 2014)。

転写因子遺伝子を対象としたSTS座のアレル識別性を高めるため、CAPS領域における塩基配列上の全変異の検出し、SNPマーカー化を試みた。CAPS解析に用いた207のSTSについて、A255及びG434間の配列比較を行った結果では、132のSTSに少なくとも1カ所のSNPが検出され、解析塩基中のSNP数は557 (100塩基当たり0.578) であった。CAPSが全く検出されなかった37のSTSでも、SNPが新たな多型として検出された。また、すでに一方の親だけから分離するCAPSが得られていたSTS座位についても、他方の親から分離が期待されるSNPも検出されるSTSもあった。これらのSNPについてGoldenGate法により、AG集団各個体の遺伝子型決定(ジェノタイピング, genotyping)を行った (Fuji *et al.* 2013)。転写因子様遺伝子を対象としてマーカーにおいては、新たに転写因子様遺伝子と推察された座位を加えたCAPS125座にSNP37座を加え (両者の重複10)、合計152の座位をマップした。この転写因子様遺伝子マーカーは、708座による基本遺伝地図 (Shimada *et al.* 2014) に対してマーカー数として全座位の25%を占め、各連鎖群の47%-91%の領域をカバーした。このSTSは、クレメンチン (*Citrus clementina* hort. ex Tanaka) ゲノム (URL1-4-6) の主要な9つの骨格配列の53.7% (scaffold_1) ~98.7% (scaffold_5) の範囲に配置され、ゲノム全体では84.5%をカバーした (表2-1-4)。

交雑集団を用いた連鎖解析では、遺伝地図上に質的及び量的育種形質をマッピングすることができる。AG集団では、花粉親‘中間母本農5号’のもつ‘無核紀州’に由来する無核性 (Nesumi and Matsumoto 2000) は、‘中間母本農

表2-1-4 カンキツ転写因子様遺伝子候補EST配列に基づいて作製したSTSにおけるCAPS及びSNPマーカーの獲得頻度

統合連鎖群	Ccl scaffold配列に対するカバナー				CAPSマーカー				SNPマーカー		
	転写因子マーカー数 (対全マーカー座数%)	連鎖群カバナーcM (対全長%)	Scaffold ^z 両端マーカー間bp	Scaffold対塩基数(%)	STS長(平均bp)	制限サイト数 ^y	多型サイト数	CAPS座数 ^x	解析塩基数(平均bp)	SNP数/解析配列数	SNP座 ^x /100bp
AGI-01	11 (18.3)	56.0 (47.1)	scaf1 15,550,250	53.7	1,513	11.9	1.0	8	634	7.50	1.18
AGI-02	14 (18.1)	77.2 (76.5)	scaf7 16,014,568	75.7	1,421	12.7	1.5	12	534	2.88	0.54
AGI-03	13 (17.8)	71.0 (89.2)	scaf9 30,913,818	98.4	992	10.0	1.3	11	608	2.71	0.45
AGI-04	10 (20.0)	67.1 (54.9)	scaf6 17,216,258	67.2	1,145	10.2	1.1	8	571	6.60	1.16
AGI-05	31 (25.8)	80.5 (74.7)	scaf3 47,977,184	93.9	972	9.9	1.4	23	606	5.22	0.86
AGI-06	9 (16.0)	105.6 (70.3)	scaf2 26,873,310	73.8	1,105	8.6	1.3	7	621	2.33	0.38
AGI-07	22 (25.8)	78.4 (86.9)	scaf4 24,606,471	95.9	1,122	9.5	1.5	15	533	3.67	0.69
AGI-08	24 (21.2)	96.8 (91.5)	scaf8 22,123,059	88.0	937	9.2	1.5	20	534	2.92	0.55
AGI-09	18 (25.0)	92.3 (79.6)	scaf5 42,754,582	98.7	972	9.3	0.7	12	544	3.42	0.63
全体	152 (21.5)	725.0 (73.1)	(合計) 244,029,500	(84.5)	1,090	10.0	1.3	116	570	4.17	0.73

*STS配列に対応するクレンチングノム(V1.0)Scaffold番号

*16種制限酵素による切断全切断部位数

*AG集団連鎖地図上に座乗したCAPSおよびSNPマーカー数

5号' (G434) の第9連鎖群にマップされた。同時に、胚 (embryo) の色 (緑/白) も質的形質としてマップされた。また、異なる集団で行われたマッピング結果は、EST座位の相互比較により、多胚性 (polyembryony, Nakano *et al.* 2008a) は第1連鎖群、雄性不稔 (Nakano *et al.* 2003) は第2連鎖群、カラタチ由来のCTV抵抗性 (Ohta *et al.* 2014) は第2連鎖群にマップされた。

カロテノイド総含有量及び各成分組成などの量的形質についても、AG集団各個体のカロテノイド各成分の含有量を測定することで、QTL解析が行われ、カロテノイド総含有量に対するQTLなどが検出された (Sugiyama *et al.* 2011)。また、カロテノイドの代謝酵素各遺伝子の発現強度のeQTL解析 (Sugiyama *et al.* 2014) により、代謝酵素遺伝子の発現制御領域が推定された。また、その他の量的な形質についても、マーカー座位を比較することで、統合化表示する (口絵2-1-2) ことができる。

カンキツの連鎖地図作製と形質マッピングが開始されて以来、世界の各チームでは、多くの質的及び量的形質のマッピングが行われてきた。これらについても多くの育種形質の参照可能なマーカーにより統合遺伝地図上に俯瞰できるように表示し、育種計画を立案に参照できる情報となることが期待される。

(4) EST座遺伝子型タイピングの育種的利用

DNAマーカーの開発には、育種における選抜を効率化することへの効果が期待されてきた。量的形質の改良を主な育種目標とする選抜の場面において、必要かつ十分なマーカーの開発はまだ途上にある。一方、主動遺伝子に制御されるカラタチ (*Poncirus trifoliata* (L.) Ref.) 由来のカンキツトリステザウイルス (citrus tristeza virus, CTV) 抵抗性に関しては、育種実生へのマーカーの適用が可能になっている。カラタチを抵抗性素材とした交雑育種においては、ゲノム全体のマーカー遺伝子型を把握することにより目的個体を効果的に選抜できる。カラタチ由来のCTV抵抗性遺伝子座が詳細にマッピングされ (Gmitter *et al.* 1996)、その領域のゲノム配列が解析された (Young *et al.* 2003) ことから、抵抗性遺伝子座領域を対象とした選抜マーカーが作出され、カラタチ由

来の交雑後代においてマーカーによる遺伝子型検定が可能になった (Ohta *et al.* 2011). さらに, 抵抗性遺伝子をもつカラタチ雑種にカンキツを戻し交雑し, カンキツの遺伝的背景に置換された個体を選抜することで, カラタチのもつ落葉性 (deciduous), 刺 (thorn), 三出葉 (trifoliate), 果実の苦味などの好ましくない特性を除くことができる. 戻し交雑世代で, ゲノム全体にわたるグラフ遺伝子型 (graphical genotype, Young and Tanksley 1989) を決定し, CTV抵抗性遺伝子座近傍での組み換え型個体を検索することで, 次世代の育種母本候補を効率よく選抜できることを示した (Ohta *et al.* 2014).

EST座位の多型に基づくマーカーは, 適切な制限酵素を選択すれば多くのカンキツ品種のジェノタイピングに適用できるため, カンキツ品種・系統, 育種母本, 交雑後代についてゲノム全体を通じたマーカー遺伝子型を知ることができる (Omura *et al.* 2003b; Fujii 2013). カンキツ育成品種について第7連鎖群のCAPS遺伝子型を示すように (表2-1-5), グラフ遺伝子型で表示することで品種の遺伝的構成を明瞭に把握できるようになる. このようなジェノタイピングは, 品種・系統識別や親子鑑定などに簡易に応用できるほか, 交雑後代系統の遺伝子型構成と育種特性との関連解析や, 育成経過に伴う遺伝子型頻度の遷移などの育種情報を読み取ることができる. 品種のマーカー遺伝子型と育種形質との関連解析の試み (Omura *et al.* 2005) では, 果皮の厚さ, 果皮色, 収穫期などの果実形質と関連を示すマーカー座位のいくつかが分離集団を用いたQTL解析と重なることを示した. 現段階では, このような情報は限られた程度でしかないが, 品種のもつ育種情報を遺伝子型から理解するために, 品種のゲノム全体に渡る遺伝子型データに遺伝統計学的な育種形質の解析情報を重ね合わせた情報整理が重要になると思われる.

(5) 機能マーカーの整備・探索

CAPSあるいはSNPマーカーによる遺伝地図では, 座位間の連鎖関係からゲノム上における遺伝子座の相対的位置を示す. 同時に, その遺伝子座位におけるDNA配列の多型分化の情報であることから, 遺伝子座を構成するアレルに

表2-1-5 カンキツ品種のCAPSによるグラフ遺伝子型 (第7連鎖群)

STS名	制限 酵素	Ccl. scaffold_4	AGI-07 座位 (cM)	品種 (CAPS遺伝子型)																
				宮川早生	トロビタ	清見	あまが、 陽香	不知火	はるみ	西之香	清峰	朱見	津之香	ありあけ	南香					
A10524	<i>Hha</i> 1	115362	6.0	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	
T10410	<i>Msp</i> 1	213798	6.1	a_a	a_b	a_a	-	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	b_b	b_b	a_b
Ov0507	<i>Xba</i> 1	642157	8.3	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	b_b	b_b	a_b
A10617	<i>Eco</i> R5	1021179	8.6	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	b_b	b_b	a_b
Ov0002	<i>Hinf</i> 1	2422861	13.7	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	b_b	b_b	a_b
Lp0112	<i>Nde</i> 2	3749809	19.6	a_a	a_b	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	b_b	b_b	a_b
A10206	<i>Hae</i> 3	12309480	38.6	a_a	a_b	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
Cp1279	<i>Hind</i> 3	12900777	40.2	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
T10318	<i>Hinc</i> 2	13751433	37.5	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
B10136	<i>Hha</i> 1	15056834	45.1	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
Gn0040	<i>Hae</i> 3	16232756	43.5	a_a	a_b	a_a	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
M10097	<i>Dra</i> 1	16712914	44.7	a_a	a_b	a_a	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
T10165	<i>Dra</i> 1	19605473	55.5	a_a	a_b	a_a	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
T10337	<i>Hae</i> 3	19635415	54.9	a_a	a_b	a_a	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
Cp0675	<i>Msp</i> 1	20294240	59.8	a_a	a_b	a_a	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a
Cp0452	<i>Eco</i> R1	23575399	75.4	a_b	a_a	a_a	a_a	a_b	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_a	a_b	a_b	a_a

機能分化を生じている可能性を示唆する。質的形質のマッピング、QTL解析や全ゲノム関連解析 (genome wide association study, GWAS) では、形質との連鎖関係を示すゲノム中の塩基配列多型を形質に対するマーカーとして検出する。しかしながら、形質と連鎖するDNAマーカーとの間には組換え型の出現は避けがたく、そのことによる選抜エラーをさけるためには、遺伝子の発現差や形質差と関連する塩基配列の差異に着目する機能マーカーの検索が重要である。しかしながら、特定の交雑集団による解析で得られる多型情報では、カンキツ育種素材品種群のもつ全てのアレルを検出しているわけではない。より広範な育種実生群に対して有効な選抜マーカーとするためには、育種形質に関与する遺伝子の塩基配列上の変異を網羅的に検索し、広範な育種素材系統における多様なハプロタイプを識別するタグ (haplotype tag) SNPを選抜することで、アレルの機能差と関連を示す機能マーカーを作製することが今後より重要になる。

カロテノイド (carotenoid) 代謝に関わるゼアキサントニンエポキシダーゼ (zeaxanthin epoxidase, ZEP) 遺伝子の発現強度に関わるeQTL領域がZEP1/2座位と一致したことから、アレル間に発現差を引き起こす配列差ZEPがこの遺伝子のシス制御領域であると推察された (Sugiyama *et al.* 2010, 2014)。生食用カンキツ品種の間にもZEP1/2座には塩基配列上多様に分化したハプロタイプがあり、イントロンの長さや5'非翻訳領域 (5' untranslated region, 5' UTR) 配列などに大きな差異により区別されるZEP1アレル群とZEP2アレル群とが認められた。このアレル群の間には、明らかな発現差も認められた。この事例では、カンキツでの品種における発現の多様性解析とCAPSマーカー遺伝子型との関連解析 (association analysis) では、単に大きくアレル群を区別するに過ぎず、多様に分化したアレル群のなかで、どの2つのアレル組み合わせが適正な機能を示すのかの選抜指標としてはまだ十分ではない。今後、より有効な選抜マーカーの作出が求められる。

蓄積された多数品種のEST配列あるいはSTS配列からのSNPの検出も進められている (Chen and Gmitter 2013, Fujii *et al.*; 2013)。今後急速に解析が進

むことが期待される品種ごとのゲノムシークエンスの多型データを用いることで、ハプロタイプを識別しつつ、その配列をタグとしてより機能的なマーカーを作製することに役立つものと思われる。

5) カンキツ特有な遺伝子解析におけるゲノム情報解析への期待

スループットの高いゲノム塩基配列決定技術が開発されたことで、オレンジ (Xu *et al.* 2013) やクレメンティン (Wu *et al.* 2014) のゲノム配列が決定・公表された。これらのゲノム配列に対しては、それぞれ380万及び156万の発現配列をマッピングし、25,376及び24,533の蛋白質コード座位が予測されている。ゲノム配列と注釈情報の公開により、マップベースクローニング (map base cloning) 法などによる遺伝子同定の道筋が整備されてきたといえる。

カンキツの多胚性制御遺伝子の解析においては、詳細な連鎖解析により第1連鎖群にマッピングされ、宮川早生のBACを用いてその領域380 kbゲノム配列が解析された。この領域は、近傍に座乗するEST配列由来のSTSマーカー Mf0086 (Nakano *et al.* 2008a, b) によりクレメンティン scaffold_1上にマップされることが分かり、多胚性・単胚性各品種間の配列比較も容易となった。しかしながら、一般的にはカンキツの単交雑集団で目的遺伝子座の近傍領域を高解像度でマッピングするのは難しく、マップベース法だけで遺伝子を同定するのは容易ではない。そのため、高密度なマーカーの準備、高精度な遺伝統計学的解析の利用や変異系統の比較解析などの解析技術の高度化とともに、他種類のゲノム情報の集約できるシステムが求められている。カンキツの多胚性制御遺伝子近傍領域では、ブドウゲノムとの間に高い遺伝子配列順の相同性 (シンテニー, synteny) が検出された (Nakano *et al.* 2012)。このことは、異なる果樹の種類間の部分的なシンテニー解析だけでなく、高密度なマーカー情報をも利用して、近縁種・遠縁種におけるゲノム構造多型の解析も有効になると期待される。今後は、構造比較によるゲノム配列に対する注釈づけ (annotation) もそのための重要な情報整備の領域と思われる。

ミカン科 (Rutaceae) に特有な二次代謝 (secondary metabolism) や形態的・

生理的特性に関与する遺伝子については、他の生物種の遺伝子解析からの参照できるのはモチーフ構造などに限られている。このため、ゲノム配列上に予測される読み枠 (open reading frame, ORF) 配列の上に、カンキツに特有な遺伝子の配列としての可能性を示唆する情報が注釈されることはほとんどない。例えば、ミカン科に特有なリモノイド (limonoid) の合成経路に関わる遺伝子は、代謝酵素群プロテオーム解析がほとんど未着手でもあるため、ゲノム配列からはほとんど予測されていない。また、カンキツに特徴的なアルカロイド (alkaloid) であるアクリドン (acridone) 類の骨格生成に関与するアクリドン合成酵素遺伝子については、ミカン科のヘンルーダ (*Ruta graveolens*) などから報告されているが、その構造多様化に関わる遺伝子は、ほとんど解析されていない。このようなカンキツに特異性の高い代謝系の酵素遺伝子などについて、配列推察の基盤となるような解析研究も今後の課題となろう。形質発現の面からは、果実のメタボロミクス (metabolomics) と連携した発現プロファイル (Lin *et al.* 2015) の解析なども開始された。果実の発育、成熟などに関わる一連の遺伝子群の発現プロファイルなど情報が、品質や形態的特性の遺伝的理解に新しい手がかりとなるため、カンキツの各種網羅的解析情報の整備、公開利用が今後一層重要になると思われる。

特定の生物種にだけ見いだされる固有の遺伝子が、その生物種に特有な形質に関与する可能性があることから、他の生物には相同性遺伝子がみつからない孤児遺伝子 (オーファンジーン, orphan gene, Arendsee *et al.* 2014) の抽出が行われている。カンキツのオーファンジーンを抽出し、遺伝子・ESTの発現に基づいた樹種比較などの情報を整備することで、カンキツに特有な形質と関連づけ手がかりになるかもしれない。樹種間・生物間で直系相同遺伝子ごとに対応付けしたセット (オーソログセット, ortholog set) として整備しておくことで、異なる果樹の間で遺伝子の有無や発現プロファイルを系統的に一覧比較できるようになり、新たな遺伝子機能情報やより詳細な相互関係の抽出に貢献するものと期待される。また、生物種特有な遺伝子の解明は、生物における非コード配列からの遺伝子創出 (Li *et al.* 2015) を含めた遺伝子分化の研

究に貢献する可能性もあり、これまでに蓄積されてきたESTの配列情報や発現情報、多型情報など統合・整備・公開されることで、カンキツに特異的な遺伝子の同定・解析に寄与する情報を提示してくれることを期待したい。

引用文献

- Adams, M. D. *et al.* (1991) Complementary DNA sequencing : Expressed sequence tags and human genome project. *Science*. 252 : 1651-1656.
- Arendsee, Z. W. *et al.* (2014) Coming of age: orphan genes in plants. *Trends in Plant Science*. 19 : 698-708.
- Baltimore, D. (ed) (1988) Report of the Ad Hoc Program Advisory Committee on Complex Genomes. Reston, VA, February 1988 (National Institute of Health, Bethesda, MD, 1988) .
- Botstein, D. *et al.* (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum Genet.* 32 : 314-331.
- Brenner, S. (1990) *Ciba Found Symp.* 149 : 6.
- Carlson, J. E. *et al.* (1991) Segregation of random amplified DNA markers in F1 progeny of conifers. *Theoretical and Applied Genetics*. 83 : 194-200.
- Chen, C. and F. G. Gmitter, Jr. (2013) Mining of haplotype - based expressed sequence tag single nucleotide polymorphisms in citrus. *BMC Genomics*. 14 : 746.
- Durham, R. E. *et al.* (1992) Linkage of restriction fragment length polymorphisms and isozymes in *Citrus*. *Theoretical and Applied Genetics*. 84 : 39-48.
- Endo, T. *et al.* (2007) Promoter analysis of a type 3 metallothionein - like gene abundant in Satsuma mandarin (*Citrus unshiu* Marc.) fruit. *Scientia Horticulture*. 112 : 207-214.
- 藤井浩ら (2003) カンキツESTデータベースシステムの開発. 果樹研究所報告. 2 : 91-99.
- Fujii, H *et al.* (2010) Development of software to improve the design efficiency of the intron - spanning PCR primer from EST utilizing genomic information of *Arabidopsis thaliana*. *Proceedings of Plant and Animal Genomes XVIII Conference*. P860.

- Fujii, H. (2013) Studies on development and application of high-throughput genomic and bioinformatics tools for citrus fruit physiology and breeding. Ph. D. thesis. Gifu University, Gifu.
- Fujii, H. *et al.* (2013) High-throughput genotyping in citrus accessions using an SNP genotyping array. *Tree Genetics and Genomes*. 9 : 145-153.
- Gmitter, Jr. F. G. *et al.* (1996) A localized linkage map of the citrus tristeza virus resistance gene region. *Theoretical and Applied Genetics*. 92 : 688-695.
- Hisada, S. *et al.* (1996) Random sequencing of sweet orange (*Citrus sinensis* Osbeck) cDNA library derived from young seeds. *Journal of Japanese Society for Horticultural Science*. 65 : 487-495.
- Hisada, S. *et al.* (1997) Expressed sequence tags of *Citrus* fruit during rapid cell development. *Journal of American Society for Horticultural Science*. 122 : 808-812.
- Höfte, H. *et al.* (1993) An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. *Plant J.* 4 : 1051-1061.
- Höög, C. (1991) Isolation of a large number of novel mammalian genes by a differential cDNA library screening strategy. *Nucleic Acids Res.* 19 : 6123-6127.
- Jansen, R. C. and J. P. Nap. (2001) Genetical genomics : the added value from segregation. *Trends in Genetics*. 17 : 388-391.
- Jarrell, D. C. *et al.* (1992) A genetic map of citrus based on the segregation of isozymes and RFLPs in an intergeneric cross. *Theoretical and Applied Genetics*. 84 : 49-56.
- Keurentjes, J. J. *et al.* (2007) Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proceedings of the National Academy of Science*. 30 : 1708-1713.
- Koltunow, A. M. *et al.* (1996) Polyembryony in *Citrus*. Accumulation of seed storage proteins in seeds and embryos cultured *in vitro*. *Plant Physiology*. 110 : 599-609.
- Konieczny, A. and F. M. Ausubel (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant Journal*. 4 : 403-410.
- Li, L. and E. S. Wurtele (2015) The QQS orphan gene of *Arabidopsis* modulates carbon and nitrogen allocation in soybean. *Plant Biotechnology Journal*. 13 : 177-187

- Lin, Q. C. *et al.* (2015) Transcriptome and metabolome analyses of sugar and organic acid metabolism in Ponkan (*Citrus reticulata*) fruit during fruit maturation. *Gene*. 554 : 64-74
- McCombie, W. R. *et al.* (1992) *Caenorhabditis elegans* expressed sequence tags identify gene families and potential disease gene homologues. *Nature Genet.* 1 : 124-131
- Moriguchi, T. *et al.* (1998) Characterization of gene repertoires at mature stage of citrus fruits through random sequencing and analysis of redundant metallothionein-like genes expressed during fruit development. *Gene*. 211 : 221-227.
- Nakano, M. *et al.* (2003) Linkage analysis between male sterility of citrus and STS markers. *Proceedings of the International Society of Citriculture IX Congress 2000*. 179-180.
- Nakano, M. *et al.* (2008a) Mapping and haplotyping of the flanking region of the polyembryony locus in *Citrus unshiu* Marcow. *Journal of Japanese Society for Horticultural Science*. 77 : 109-114.
- Nakano, M. *et al.* (2008b.) Marker enrichment and construction of haplotype - specific BAC contigs for the polyembryony genomic region in *Citrus*. *Breeding Science*. 58 : 375-383.
- Nakano, M. *et al.* (2012) Characterization of genomic sequence showing strong association with polyembryony among diverse *Citrus* species and cultivars, and its synteny with *Vitis* and *Populus*. *Plant Science*. 183 : 131-142.
- Nesumi, H. and Matsumoto, R. (2000) Improvement of citrus scion cultivars by crossbreeding in Japan. *Proceedings of the International Society of Citriculture IX Congress 2000*. 46-47.
- Ohta, S. *et al.* (2011) PCR primers for marker assisted backcrossing to introduce a CTV resistance gene from *Poncirus trifoliata* (L.) Raf. into Citrus. *Journal of Japanese Society for Horticultural Science*. 80 : 295-307.
- Ohta, S. *et al.* (2014) Construction of genetic linkage map and graphical genotyping of pseudo - backcrossed F₂ (BC'₂) progeny to introduce a CTV resistance from *Poncirus trifoliata* (L.) Raf. into *Citrus* by introgression breeding. *Tree Genetics and Genomes*. 11 : 797
- Omura, M. *et al.* (1993) PCR markers in *Citrus* identification and mapping. p.66-73. Hayashi, T. *et al.* (eds.) *Techniques on gene diagnosis and breeding in fruit trees*. Fruit Tree Research Station, MAFF, Japan.

- Omura, M. *et al.* (2000) Analysis and application of cDNA cataloguing from *Citrus* fruit for breeding. *Acta Horticulturae*. 535 : 175-181.
- 大村三男 (2001) カンキツのゲノム解析 -CAPSマーカーによる品種のタイピングとマッピング-. p. 216-220. 佐々木卓治ら監修. 植物のゲノム研究プロトコール 秀潤社 東京.
- Omura, M. *et al.* (2003a) EST mapping of *Citrus*. *Proceedings of the International Society of Citriculture IX Congress 2000*. 71-74.
- Omura, M. *et al.* (2003b) Graphical genotype of citrus cultivars by co-dominant CAPS marker. *Proceedings of Plant and Animal Genomes XI Conference*. p. 206.
- Omura, M. *et al.* (2005) Genome typing of citrus cultivars by CAPS markers to detect QTL candidates on fruit and tree growth characteristics. *Proceedings of Plant and Animal Genomes XIII Conference*. p. 504.
- Park, Y. S. Lee and H. G. Nam. (1993) Generation of expressed sequence tags of random root cDNA clones of *Brassica napus* by single-run partial sequencing. *Plant Physiol*. 103 : 359-370.
- Sasaki, T. *et al.* (1994) Toward cataloguing all rice genes : large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. *Plant Journal*. 6 : 615-624.
- Shimada, T. *et al.* (2014) Construction of a citrus framework genetic map anchored by 708 gene-based markers. *Tree Genetics and Genomes*. 10 : 1001-1013.
- Sugiyama, A. *et al.* (2010) Structure and expression levels of alleles of *Citrus* zeaxanthin epoxidase genes. *Journal of the Japanese Society for Horticultural Science*. 79 : 263-274.
- Sugiyama, A. *et al.* (2011) Quantitative trait loci (QTL) analysis of carotenoid content in *Citrus* fruit. *Journal of the Japanese Society for Horticultural Science*. 80 : 136-144.
- Sugiyama, A. *et al.* (2014) Expression quantitative trait loci analysis of carotenoid metabolism-related genes in *Citrus*. *Journal of the Japanese Society for Horticultural Science*. 83 : 32-43.
- Uchimiyama, H. *et al.* (1992) Random sequencing of cDNA libraries reveals a variety of expressed genes in cultured cells of rice (*Oryza sativa* L.). *Plant Journal*. 2 : 1005-1009.
- Watson, J. D. (1990) The human genome project : past, present, and future. *Science*. 248 : 44-49.

- Wu, G. A. *et al.* (2014) Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature biotechnology*. 32 : 656-662
- Xu, Q. *et al.* (2013) The draft genome of sweet orange (*Citrus sinensis*) . *Nature Genetics*. 45 : 59-66.
- Young, N. D. and S. D. Tanksley (1989) Restriction fragment length polymorphism maps and the concept of graphical genotypes. *Theoretical and Applied Genetics*. 77 : 95-101.
- Young, Z. N. *et al.* (2003) Sequence analysis of a 282-kilobase region surrounding the citrus tristeza virus resistance gene (*ctv*) locus in *Poncirus trifoliata* L. *Raf. Plant Physiology*. 131 : 482-492.

