



Science Plaza

「染井吉野」のゲノムを解読して分かったこと、できるようになること



公益財団法人 かずさ DNA 研究所
 先端研究開発部 植物ゲノム・遺伝学研究室
 主任研究員
 白澤 健太

1. 桜の代表品種「染井吉野」

冬の寒さに耐え忍び、春先に花を咲かせ、そして一斉に散ってしまう桜は、私たち日本人の風情によく似合うとされ、日本を象徴する花として広く国民に親しまれています。春が近づくと多くの人が桜の開花を心待ちにし、開花予測報道は「桜前線」などの言葉とともに世間を大きく賑わせます。残念ながら2020年は新型コロナウイルス（COVID-19）の感染拡大防止対策のために全国各地で花見が自粛されました。来る2021年こそは、全国から花見の便りが届くことを願って止みません。

なぜ桜は一斉に花を咲かせ、一斉に散ってしまうのでしょうか。じつは本州、四国、および九州にお住まいの方が普段の花見で目にする桜のほとんどは「染井吉野」という品種です（図1）。全国の「染井吉野」は遺伝的に均一なコピー（クローン）なので性質がお互いによく似ています。そのために私たちは「染井吉野」ばかりを見ては、「桜は一斉に咲いて一斉に

散ってしまう」と感じているのです。事実、「染井吉野」とは異なる遺伝情報を持つ「河津桜」は「染井吉野」よりも早い時期に開花しますし、かずさDNA研究所のある千葉県木更津市矢那地区に伝わる「鎌足桜」は逆に遅い時期に開花します。

このように、一口に桜と言っても様々な種（しゅ）や品種があります。日本には10種の桜が自生し、品種の数は200を超えます。これらの桜は、森林総合研究所多摩森林科学園（東京都八王子市）、



▲図1 見頃を迎えた「染井吉野」の花

Contents

Science Plaza

「染井吉野」のゲノムを解読して分かったこと、
 できるようになること 1

新課程特集 Part1~4

物理教育 東京学芸大学 教授 新田 英雄 6

化学教育 開成学園 教諭 小松 寛 10

生物教育 都立八王子高等学校 教諭 長尾 嘉崇 16

地学教育 埼玉県立小鹿野高等学校 教諭 小幡 喜一 20

連載

研究で君が光り輝くために 第5回 24

高校生へ私が選んだ1冊の本
 カガク力を強くする！ 28

国立遺伝学研究所（静岡県三島市）、京都府立植物園（京都市）、および島根大学本庄総合農場（島根県松江市）などで保存・管理されています。「染井吉野」は「江戸彼岸」と「大島桜」の系統が交雑することで誕生した種間雑種の品種で、江戸時代の末期に染井村（今の豊島区駒込あたり）で発祥したとされています。しかし、2つの種が自然交雑したのか（いつ、どこで、どの木とどの木が？）、または人工的に交雑されたのか（いつ、どこで、誰が、どの木とどの木を？）など、成り立ちの詳細は分かっていません。

2. ゲノムは生命の設計図

生物がもつ全ての遺伝情報をゲノムと言います。生物の多くの性質は遺伝情報によって決められているので、ゲノムは生物の設計図に例えられます。その遺伝情報は、染色体を構成するDNAに4種類の塩基〔アデニン（A）、グアニン（G）、シトシン（C）、およびチミン（T）〕の配列として書き込まれています。これらの塩基は相補的な対を構成し、これによりDNAは二重らせんを形作ります。長いDNAのところどころには、遺伝情報が書かれた遺伝子があります。例えばヒトのゲノムには約30億の塩基対があり、その中に約2万個の遺伝子が見つっています。遺伝子部分のDNAの塩基配列はRNAに転写され、RNAの塩基配列に従ってアミノ酸の配列が決定されてタンパク質に翻訳されます。転写や翻訳を遺伝子発現と呼びます。ゲノムDNAの塩基配列は個人間で異なり、それによってタンパク質の性質が変わる場合があります。このため、生物は同種であっても個体間で性質が異なるのです。さらに、発現する遺伝子もからだの組織や器官、また外部の環境に依存して変化します。クローンである「染井吉野」が全国で同時に開花するのではなく、気温の高い南から順に開花するのはこのためです。生物の設計図であるゲノムDNAの塩基配列を解読することは、遺伝子のはたらきを知り生物の性質を理解するためのはじめの一歩となるのです。

3. 「染井吉野」のゲノム解読に挑む

「染井吉野」は、どのように誕生したのか？いつ咲くのか？その手がかりを得るために、約6億9千万塩基対と推定される「染井吉野」のゲノムDNAの配列を決定しよう、というのが私たちの「染井吉野」

ゲノム解読プロジェクトの始まりでした。

ゲノムDNAの塩基配列はどのように解読するのでしょうか？DNAの塩基配列分析にはDNAシーケンサーという分析装置を使用します。ヒトゲノム解読プロジェクトが完了した2000年台前半まではサンガー法を利用した分析装置が主力でしたが、2000年代後半から次世代シーケンス（Next-Generation Sequencing, NGS）技術が実用化されると、得られるデータの量が爆発的に増えました。私たちは、上野恩賜公園に植栽されている「染井吉野」（図2）から公園の許可を得て葉を採取し（図3）、ゲノムDNAを抽出し、DNAシーケンサーを使用してDNAの塩基配列データを得ました（図4）。一つの塩基配列データには約1万7千対の塩基が含まれ、取得したデータの全体量は373億塩基に達しました。この塩基配列の量は、「染井吉野」のゲノムDNAの大きさである6億9千万塩基の50倍以上の量に匹敵します。

なぜ、ゲノム解読にはゲノムDNAの大きさの50倍以上ものデータが必要なのでしょう？その理由は、



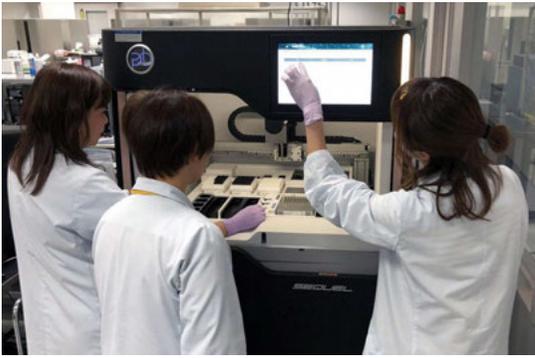
▲図2 ゲノム解読に使用した上野恩賜公園に植栽されている「染井吉野」

上野動物園の正門近くに植栽されている「染井吉野」をゲノム解読に使用しました。



▲図3 上野恩賜公園に植栽されている「染井吉野」からの葉の採取作業

共同研究者である江角智也准教授（島根大）が上野恩賜公園の許可を得て「染井吉野」の葉を採取している様子です。



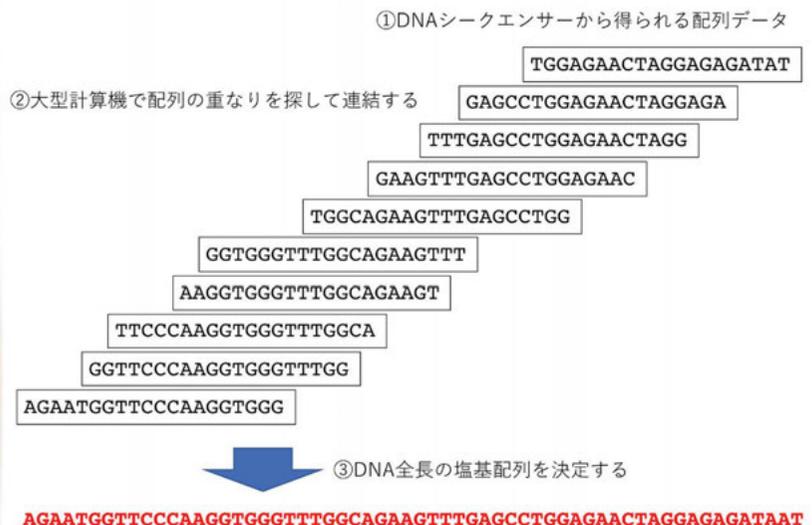
▲図4 DNAの塩基配列データの収集に使用したDNAシーケンサー

かずき DNA 研究所の研究室メンバーがDNAシーケンサーを操作している様子です。この機種からは、一度の分析で約100億塩基分に匹敵する膨大な塩基配列データ（一つのデータに含まれる塩基配列は約2万塩基分です）が得られます。

ゲノム解読方法の技術的な限界にあります。現在の技術では一度に数万塩基程度の長さの配列しか読み取ることができません。従来は数百塩基程度の長さでしたので、その頃に比べればずいぶん長くなりましたが、それでも数億塩基あるゲノムDNAの端から端までを一度に分析することはできないのです。そこで、数万塩基ほどの長さで読み取った膨大な塩基配列データの中から、お互いに同じ配列を持つデータを探して出して、それを重ね合わせることで、元のゲノムDNAの塩基配列を再構築します。この作業には高性能計算機

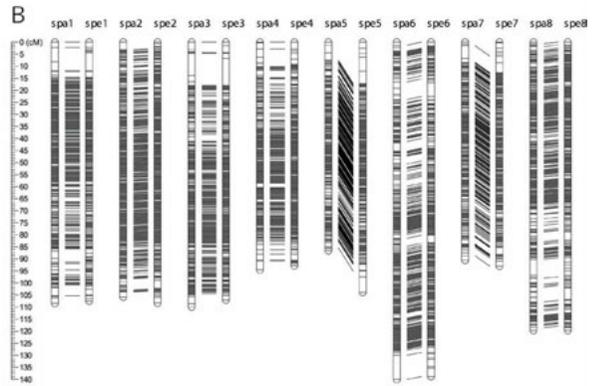
の利用が不可欠です（図5）。このようにして、373億塩基分のデータを連結して、「染井吉野」のゲノム配列を決定しました。決定した配列の長さは6億9千万塩基で、これは事前に推定した通りの大きさでした。

そして、決定したゲノム配列を染色体に沿って整列化しました。メンデルはエンドウの交雑実験から遺伝の法則を発見し、モーガンはショウジョウバエの交配実験から遺伝子の連鎖と染色体の組み換えを発見し染色体地図を作成しました。これらの現象はDNAの塩基配列にも当てはめることができます。個体間に見つかるゲノムDNAの配列の変異はDNA多型と呼ばれ、その中でも一塩基多型（SNP）は最も高頻度で見つかるDNA多型です。このSNPがメンデルの法則に従って遺伝します。同一染色体上のSNPは連鎖しますが、染色体の組み換えが起きれば連鎖が解消されます。私たちは「染井吉野」を「山桜」と交雑して後代を得て、交雑後代のSNP遺伝子型を解析することで「染井吉野」の染色体地図を作成しました（図6）。この染色体地図に沿って「染井吉野」のゲノム配列を並べることで「染井吉野」の染色体のDNA配列を再現しました。そして、その中から約9万5千個の遺伝子を見出しました。このような解析を経て、私たちは「染井吉野」のゲノム解読を完了したのです。



▲図5 ゲノムDNAの塩基配列を決定する方法

膨大な配列データから大型計算機を使用してゲノムDNAの塩基配列を決定する方法を示します。DNAシーケンサーから得た配列データ①から、お互いに同じ配列を持つデータを探して出して重ね合わせることで配列データを連結し②、ゲノムDNAの塩基配列を決定します③。この図では、200塩基分の配列データ（長さ20塩基のデータを10個分）を使用して60塩基の配列を決定する様子を模しています。



▲図6 交雑実験から作成した「染井吉野」の染色体地図
 かずさ DNA 研究所の研究室メンバーが「山桜」に「染井吉野」を交雑している様子です (A)。交雑して得られた 237 粒の種子から DNA を抽出し SNP 遺伝子型を解析することで「染井吉野」の染色体地図を作成しました (B)。

4. 「染井吉野」はどこからきたのか？

今回のゲノム解読により「染井吉野」が持つ両親に由来する 2 つのゲノムの塩基配列を決定し、2 つのゲノムの間に DNA の塩基配列を見出しました。DNA の塩基配列には時間とともに突然変異が蓄積しますから、逆算すると「染井吉野」の両親は約 552 万年前に別の種に分岐したと推定されました。そして、これら 2 種のゲノムが江戸時代の末期に交雑により再会を果たしたことで「染井吉野」が誕生したと考えられました。

では、「染井吉野」の両親は何だったのでしょうか？ 今回の解析では、島根大学本庄総合農場が保有する 139 種類の桜の解析も行いました。SNP 解析から 139 種類の桜は 8 つのグループに分かれることが分かりました (図 7)。それぞれの桜のゲノムの組成を調べたところ、「染井吉野」のゲノムには「江戸彼岸」と「大島桜」の系統のゲノムが約半分ずつ含まれていることが推定されました。これまでの交雑実験や分子遺伝学実験から、「染井吉野」は「江戸彼岸」と「大島桜」の系統の交雑に由来するとされていた仮説が、ゲノム解析によっても裏付けられたのです。今後、「染井吉野」のゲノム情報を利用した DNA 親子鑑定により、「染井吉野」の誕生に関わった本当の親の木が見つかるようになるかもしれません。

なお、韓国の済州島の「瀛州 (えいしゅう) 桜 (韓国名・王桜)」が「染井吉野」の祖先であるとの推定がありましたが、「染井吉野」のゲノム解析からも両者は別物であることが確認されました。



◀図7 島根大学本庄総合農場で保存・管理されている 139 種類の桜の分類
 SNP 遺伝子型の解析により 139 種類の桜を分類しました。右のカラーパターンは、それぞれの品種を構成するゲノムの種類と割合を示します。「染井吉野」は「江戸彼岸」と「大島桜」の系統のゲノムを半分ずつ持っていました。

5. 「染井吉野」はいつ咲くのか？

「染井吉野」の開花は遺伝子によってどのように制御されているのでしょうか？これが分かれば、「染井吉野」の開花日を遺伝子情報から予測できるようになるかもしれません。そこで私たちは、開花 1 年前から 1 ヶ月ごとに、また、開花 1 ヶ月前からは数日ごとに「染井吉野」の花芽を採取して、そこではたらく遺伝子を調べました。

まず、「染井吉野」が春先に花を咲かせるために 5 種類の遺伝子グループがはたらいていることが分かりました (図 8)。開花前年の夏までに、花芽形成に関わる *CEN* 遺伝子の発現量が上がっていました。続いて夏から冬にかけて *CBF*、*EBB* および *DAM* 遺伝子の発現が順番に上昇しており、開花 2~3 ヶ月前には *FT* 遺伝子が、そして開花 1 ヶ月前には *DELLA* 遺伝子の発現が上昇することが分かりました。桜と同じバラ科樹木である梅や桃では、花芽形成が終わったつぼみのなかで *CBF* や *EBB* の発現量が上がり、これが自発休眠に関わる *DAM* の発現を促します。この *DAM* は開花を誘導する *FT* や *DELLA* の発現を抑制するのですが、冬の寒さで *DAM* の発現が減少することでその抑制が解除され、*FT* や *DELLA* の発現が上昇することで開花に向かいます。「染井吉野」も梅や桃と同じ遺伝子機

構によって春先に開花を迎えることが分かりました。

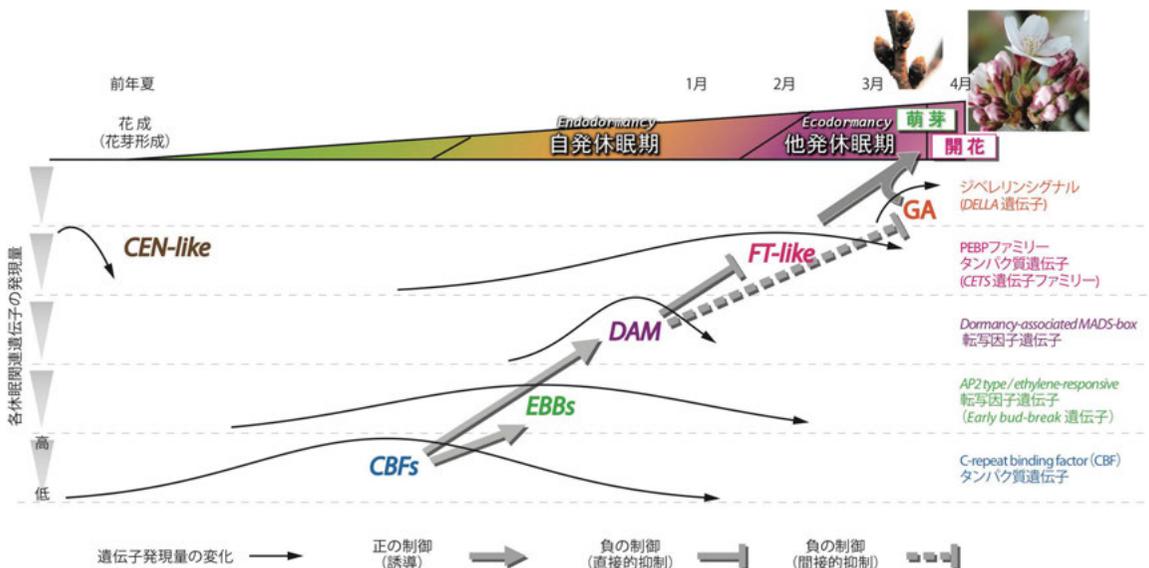
一方、「染井吉野」が開花する「日」を遺伝子で予測できないのでしょうか？開花1ヶ月前から数日ごとに採取したつぼみの遺伝子発現データを調べることで、その答えが見えてきました。遺伝子を発現の 패턴ごとに分類すると、31種類のグループに分類できました。このうち10種類のグループが興味深い発現パターンを示しました。あるグループの遺伝子は開花の2日前に特異的に発現していました。別のグループでは4日前、さらに別のグループでは6日前など、開花の20日前までの特異的な日に発現が高くなる遺伝子のグループが見つかったのです。「染井吉野」のつぼみの中で、開花に向けた準備のための遺伝子の発現が日毎に変化している様子がうかがえました。逆に言えば、これらの遺伝子の発現量を調べれば、「染井吉野」があと何日で開花するのかを予測できるようになるでしょう。現在、年次や場所を変えて、また「染井吉野」とは異なる品種でも遺伝子の発現解析を行っています。精度の高い開花予測遺伝子が見つければ、いずれ遺伝子情報で様々な桜の開花予想ができるようになるかもしれません。

6. おわりに

医療、エネルギー、食糧などの分野でのゲノム研究は、10年前では想像ができなかったほどのスピー

ドで進展しています。私たちの研究グループでは主に農作物のゲノム解読を行い、生産性や機能性の高い品種開発につなげる研究を行っています。世界中でも様々な生物のゲノム解読が進行し、毎日のように新しいゲノム解読の成果が報じられていますので、多くの方にとってゲノムがより身近に感じられるようになったのではないのでしょうか。しかし、ゲノムがどのように私たちの生活に役に立っているのかを実感する機会が少ないかもしれません。ここでは、皆さんが親しみを感じる「染井吉野」を題材として、ゲノムを解読して分かったことと、これからゲノム情報を使ってできるようになること、を紹介しました。

いまは、不思議だなと思ったことはいつでもスマートフォンを使って手軽に調べることができる時代です。これも10年前には想像ができなかった世界です。不思議だなと思ったことはどんどん調べてください。いくら調べても、いつまで経っても不思議だなと思うことはなくなるはずですよ。不思議だなと思うことがなくなるのは、自分自身が不思議だなと思わなくなってしまったときです。この資料を最後まで読んでくださった方には、いつまでも不思議だなと思う気持ちを忘れないでほしいと願っています。



▲図8 遺伝子の転写発現解析に基づく「染井吉野」の開花制御モデル
開花の1年前から直前までのつぼみを1ヶ月ごとに採取して遺伝子の転写発現を調べました。「染井吉野」の開花は梅や桃などのバラ科樹木と同じ遺伝子機構で制御されていることが分かりました。