

令和 6 年 6 月 7 日現在

機関番号：16301

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K06344

研究課題名（和文）Azure Custom Vision Serviceを用いたダンチクの植生調査

研究課題名（英文）Distributional Survey of Arundo donax by using Azure Custom Vision Service

研究代表者

佐久間 洋（SAKUMA, Yoh）

愛媛大学・理工学研究科（理学系）・准教授

研究者番号：70452688

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：MS Azure Custom Visionに四国全域に渡るコロニーの航空写真を学習画像に適応し再現率、適合率ともに90%程度の画像判別システムが得られた。さらに、愛媛県宇和島市の3地域について画像データを用い、認識モデルのブラッシュアップをおこなった。
原産地である東アジアのダンチクについてのゲノム情報は非常に少ない。愛媛県の個体についてゲノムシーケンスを行い、SSR-HRM用のプライマーを設定した。これらをもちいて14系統のダンチクについてHRM解析を行ったところ、7つのグループに分けることができた。SSR-HRM解析で十分検出可能な遺伝的多様性が存在することが明らかとなった。

研究成果の学術的意義や社会的意義

本研究により、ダンチクの生息地を特定できる画像識別システムが構築された。四国内のダンチクのコロニーは津波浸水被害予想域との照合が可能となっている。ダンチクの生息地が津波浸水被害域とオーバーラップすることが多く、大型で独特の形態から見分けることも容易である。ダンチクを津波浸水被害域の指標とする情報基盤が整備できた。ダンチクの分布には人為的な影響も多いが、ジェノタイピングのツールが準備できたことで、人為的な影響の評価も可能となった。ダンチクは東アジア以外では移入種として問題を引き起こしている。本認識システムを、ヨーロッパ系統の画像でさらに改良し、地球レベルでの認識システムの構築も期待できる。

研究成果の概要（英文）：Aerial photographs of colonies across Shikoku were utilized to MS Azure Custom Vision as training images, and an image discrimination system with a reproduction rate and fit rate of around 90% was obtained. Moreover, image data from three areas in Uwajima, Ehime Prefecture, were used to brush up the recognition model.
There is very little genomic information on *A. donax* in East Asia, where it originated. Genome sequencing was carried out on individuals taken in Ehime, and primers for SSR-HRM were set. Using these primers, HRM analysis was conducted on 14 lines of *A. donax*, which could be divided into seven groups, indicating that there is enough genetic diversity to be detected by SSR-HRM analysis.

研究分野：植物生理学

キーワード：画像認識 植生調査 津波 ジェノタイピング HRM

様式 C-19、F-19-1、Z-19(共通)

1. 研究開始当初の背景

ヨーロッパでの研究ではダンチクは稔性を持った種子を作らないとされており、種子の散布により分布域を広げることはない。ダンチクの稈には各節に芽があり、台風の際などに離脱した稈が流された先で定着し分布域を広げていると思われる。従って、横方向や上から下への移動は容易であるが、下から上へどのように分布を広げたのかは不明である。また、ダンチクが優占して生息している地域でも、ある程度の海岸からの距離、高さを境にダンチクが生息しなくなる領域がある。これらの事実から、申請者はダンチクの分布に津波が大きな影響を与えているのではないかと考えている。また、ダンチクは大型で独特の株型を呈するため、植物に詳しくない一般市民でも容易に見分けられ、安価で、迅速に四国全域に展開できる津波避難の目印にすることが可能である。

近年、ヨーロッパを中心にダンチクの研究が盛んになってきている。しかし、バイオマス生産、バイオレメディエーション、移入先でのコントロールに関する報告が多い。移入先であるヨーロッパ、アメリカ、オーストラリアにおけるダンチクの遺伝的多様性は低いことが報告されている。一方、Hardion et al. (Annals of Botany 114, 2014)は東アジア近づくほどダンチクの遺伝的多様性が高まることを明らかにしている。ダンチクは多倍数性のゲノムを持ち、6~9倍体との報告がある。ヨーロッパ、北米などに生息するグループは9倍体とされている。また、日本と欧米のダンチクの表現型は大きく異なっている。申請者はダンチクの分布に津波が影響しているのではないかと予測しているが、この点に関してもこれまでに全く報告はなく、学術的な解析は行われていない。

2. 研究の目的

本研究ではダンチクの植生調査を行うため、効率的な遂行のためインターネット上に公開されている画像データを用いる。ダンチクは大型で独特の株形をしているため、Google マップなどの航空写真でもコロニーを判別できる。また、Google ストリートビューでは目視に近い信頼度が得られる。本研究では、このようなインターネット上の画像リソースを活用し、最小限の実地調査で効率的にダンチクの分布状況の調査を行う。また、本研究では地理空間数値情報(GIS データ)を用いて、標高や海岸線からの距離、主要道路からの距離などの地理データ、気温、降水量などの気候データ、津波到達予測域データ、などをダンチクコロニーと照合し数理解析を行う。これによりダンチクの分布に影響を及ぼす環境要因を詳細に明らかにすることができる。ダンチクの分布に関わる要因を明らかにすることは、移入先でのダンチクの分布をコントロールする上でも重要である。

、ダンチクのコロニーは津波到達予測域の先端部に多く存在している。ダンチクは大型で独特の樹形から植物に詳しくない人でも判別しやすく、津波避難の際の有効なランドマークとなる。津波の際は、ダンチクの生えている地点よりも高いところに逃げれば良い。ダンチクは関東以西に自生しており、本研究の成果は南海トラフ巨大地震の津波の影響を受ける広い地域に展開できる。前述の通り、ダンチクの応用として荒廃地でのバイオマス生産が注目されており、より生産性の高いダンチクの育種も考えられている。しかしながら移入先のダンチクは遺伝的多様性が低く育種の可能性が制限されている。原産地での遺伝的多様性が明らかになれば、大きな遺伝子プールを提供できる可能性がある。また、もし有性生殖を行うダンチクが発見されれば育種において非常に有効なツールとなる。

3. 研究の方法

Microsoft Azure Custom Vision Service を用いて、Google マップなどから得る航空写真を入力しダンチクのコロニーを自動的に特定するシステムを構築する。

海浜や河原に棲息するダンチクは特に土木工事の影響を受けやすい。実地調査で試料を採取してゲノム解析を行い、津波の影響による近隣コロニーからの分散か、人間活動により遠隔地から移入したものを推定する。また、近畿、中国、九州のダンチクを採取し、DNA を抽出、系統解析を行う。ゲノムシーケンスを行い、SSR マーカーを設定し Real-Time PCR High-Resolution Melting curve を用いて解析する。これにより、多サンプルを、安価で高速、高精度で解析することが出来る。得られた結果と、ダンチクの分布を統合し、特にダンチクの分散に海流の及ぼす影響を評価し、ダンチクの起源と分散の過程を推定する。



図1 画像認識システムの学習に用いた画像

4. 研究成果

MS Azure Custom Vision を用いたダンチクコロニーの自動検出システム

我々はすでに、Google Map の航空写真を用いて四国におけるダンチクのコロニーを特定しているが、この画像を学習モデル用いて Azure Custom Vision service にインスタンスを作成した。画像はダンチクポジティブ、ダンチクネガティブ、それぞれ 500 枚使用した。ネガティブ画像は、アシ、タケ、ササ、ヤシなど、上空からの写真でダンチクと類似の形態を取るものを選択した。ダンチクは道路脇に生息しているコロニーが多く、道路がダンチクコロニーの認識に影響してしまうことが多かった。このため、人工物をなるべく排除し、ダンチクのコロニーだけが写っているポジティブ画像を準備し、学習を行った。構築されたインスタンスは、学習画像を用いた判定ではポジティブ画像に対して適合率 91.3%、再現率 88.4%、ネガティブ画像に対して適合率 88.9%、再現率 91.4%であった。このインスタンスを用いて高知県香南市の海岸付近の画像を用いてコロニーの推定を行った。



図2 高知県香南市の海岸付近におけるコロニーの推定
赤がダンチクのコロニーと推定されたポイント、大きさは信頼性を表している。青は、航空写真から肉眼で特定されたコロニーを示している。

33.521684, 133.754399 - 33.518630, 133.764162、256 x 256px(32 x 32m 相当)の画像 319 枚を Google Maps Satellite よりダウンロード、Azure Custom Vision Service で prediction を行った。大きなコロニーは正しく予測できているが、小さなコロニーに見落としがあり、海のスコアが高く出してしまう(図2)。この問題を解決するために、さらに学習を行った。上記の判定画像について、判定結果(probability) 50%以上を正解とし、QGIS 上で、我々が衛星写真より目視で認識済みのダンチクのコロニーとの照合を行った。これにより、probability 50% 以上かつダンチクのコロニーと重複した画像をポジティブ教材、probability 50% 以上でダンチクのコロニーと重

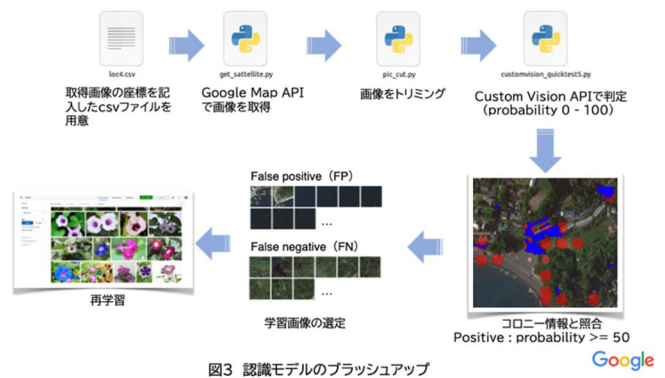


図3 認識モデルのブラッシュアップ

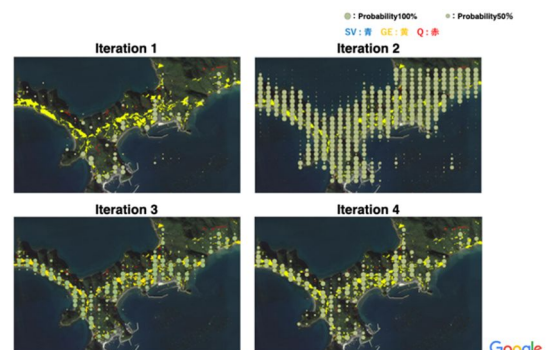


図4 認識モデルのブラッシュアップ結果

複しない画像をネガティブ教材として、再学習を行った。画像の取得、加工、Azul Custom Vision API を用いた判定には、Python のスクリプトを作成し自動化した(図3)。再学習を4階繰り返した結果、良好な正解率の認識システムが得られた(図4)。このシステムのネガティブ画像における適合率を確認するために、ダンクコロニーの存在しない画像において、判定テストを行った。植物で覆われている香川県高松市女木町の画像と、都市部である広島県福山市三之丸町の画像、それぞれ319米を用いて判別した結果、probability 50% と判別された画像は738枚中1枚だけであった(図5)。さらに、我々の構築したシステムがヨーロッパ系統の認識にも適用できるか検証を行った。Google Street View でダンクのコロニーが確認できる信頼性の高い地点での航空写真を用いて判別を行った。画像はイタリア ローマ市のものを使用した。ダンクのコロニーが確認された地点で、1ヶ所のみが probability 50% と判別された。ヨーロッパ系統は日本の系統より大型で、青みがかったため今回構築されたシステムでは判別できなかったと思われる。しかしながら、我々の今回の手法と同様に、ヨーロッパ系統の画像を用いて学習すれば、Azul Custom Vision で判別システムを構築することは十分可能であると考えられる。

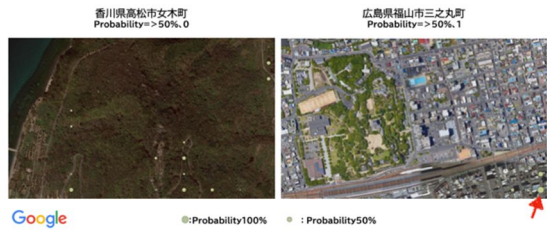


図5 ダンクコロニーのない画像での判別



図6 ヨーロッパの画像を用いた判別

ダンクのコロニーが確認された地点で、1ヶ所のみが probability 50% と判別された。ヨーロッパ系統は日本の系統より大型で、青みがかったため今回構築されたシステムでは判別できなかったと思われる。しかしながら、我々の今回の手法と同様に、ヨーロッパ系統の画像を用いて学習すれば、Azul Custom Vision で判別システムを構築することは十分可能であると考えられる。

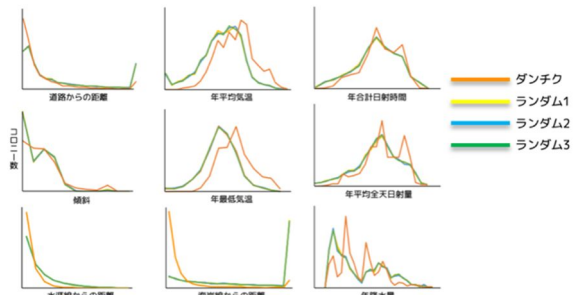


図7 ダンクコロニーの出現頻度曲線

ダンク生息地の地理的要因、環境要因

四国におけるダンクコロニーについて、地理的および環境要因における特徴を解析した。解析に用いた地理的要因、環境要因は、標高、海岸線からの距離、水害線からの距離(以上は、国土地理院の基盤情報サイトより入手した)、傾斜、地上開度、地下開度、道路からの距離、(コンサベーションコンソーシアム GIS ジャパンより入手した)、年平均気温、年最低気温、年降水量、年合計日照時間、年平均全日射量(国土制作局・国土情報課 GIS ホームページ国土数値情報ダウンロードサービスから入手した)、津波フロントラインからの距離(G 空間情報センターのデータをもとに計算した。用いたモデルは「浸水メッシュ_ケース 05_堤防 03 分破壊_04 系_太平洋(「四国沖 九州沖」に「大すべり域+超大すべり域」を設定し、地震発生から 3 分後に堤防が破壊する)である。比較対象として、調査地域にランダムに15000ヶ

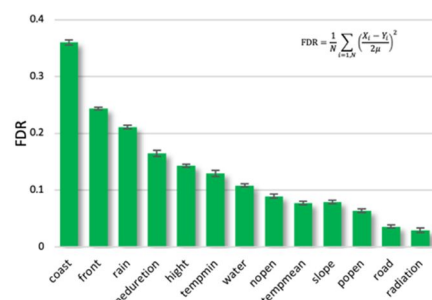


図8 ダンクコロニーの FDR

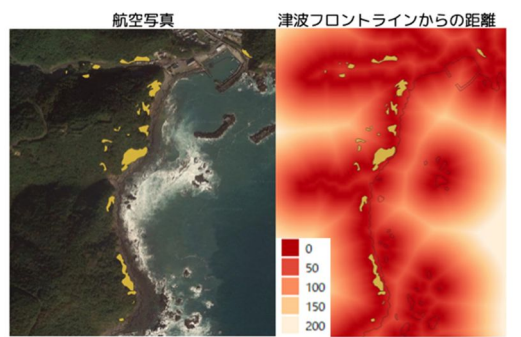


図9 ダンクコロニーと津波フロントラインの照合

所のスポットを設定し、上記の要因データを入力した。これを3回繰り返し、ダンチクコロニーのデータと比較した。評価基準には、出現頻度の密度曲線作成し(図7)、ダンチクコロニーとランダムスポットで曲線の相違度(FDR, Frequency Difference Rate)を数値化し用いた。FDR を比較したグラフを図8に示す。ランダムスポットともっとも出現パターンが異なっていた要因は海岸線からの距離であった。2番目が津波フロントラインからの距離、3番目が降水量であった。津波フロントラインからの距離とダンチクのコロニーを、自然海岸の多い高知県黒潮町の画像で比較したのが図9である。津波が入り込む谷筋に沿ってダンチクのコロニーが内陸側に広がっているのがわかる。これらのデータを画像認識システムの結果と統合することで、予測の信頼性を上げることができる。

Repeat unit size	Minimum number of repeats	Number of SSRs
2bp	15	3,675
3bp	10	809
4bp	7	280
5bp	6	103

Total number of sequences examined	6763323
Average length of sequences examined (bp)	150
Total size of examined sequences (bp)	10144984
Total number of identified SSRs	4867

図10 SSRの検出数

ゲノム解析とジェノタイピング

愛媛県松山市で採取したダンチクについて、ゲノムシーケンスを行った。シーケンサーは MGI Tech 社 DNBSEQ-G400、平均長 400bp, 150 bp x 2 paired end で行った。836,262,962 reads, 1.25675 x 10¹¹ base が得られた。K-mer 分布解析(jellyfish, 17mer)の結果、カバレッジ 11、推定ゲノムサイズ 10.2G となった。しかしながら、フローサイトメトリーを用いた報告ではダンチクのゲノムサイズは 3G 程度と予測されているので、カバレッジは 30x 程度であると予測される。のリードデータから MISA を用いて SSR の検索を行った。結果を図 10 に示す。得られた配列を Primer3 に投入し、増幅物の長さが 70 150bp になるプライマーを合成した。また、RNA seq の結果から作成されたヨーロッパ系統の既報の SSR プライマーも解析に用いた。これらのプライマーを用いて、中四国で採取したダンチクの PCR-HRM ジェノタイピングを試みた。多形が得られたマーカーの結果をまとめたのが図11である。各系統を分離できるマーカーが得られた。KOC009は我々がファイリダンチクと呼んでいる個体で、葉に斑が入り、他の系統より大型である。ヨーロッパ系統の栽培種と予測している。この系統は他の系統と異なるマーカーが多かった。その他の系統の採取時の写真を図13に示す。EHM002 と YMG002 が同じグループに分離されたが、EHM002 河川の堤防より内側、整備工事があったと思われる場所で、YMG002 は港近くの整地された場所に生息していた。人為的な影響が疑われる。

SSR marker	KOC 009	EHM 002	EHM 003	YMG 001	YMG 002	KGW 001	HRS 001
CPSSR_20, 24, 32, 34, 35, 39, 51	A	B	B	B	B	B	B
CPSSR_15	C	D	D	E	D	X	D
CPSSR_31	F	G	G	X	G	H	X
Adg 2, 4, 6, 7, 9, 11, 13, 14	I	J	J	J	J	J	J
Adg 10	K	L	M	L	L	L	L

図11 PCR-SSR-HRMジェノタイピングの結果

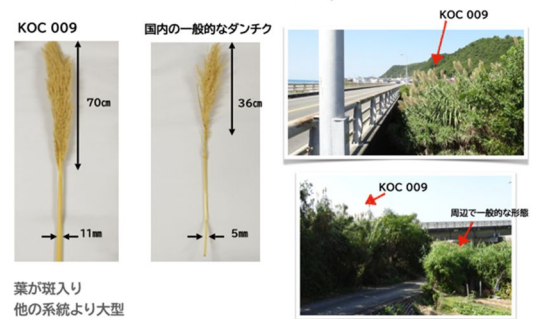


図12 KOC009の表現型

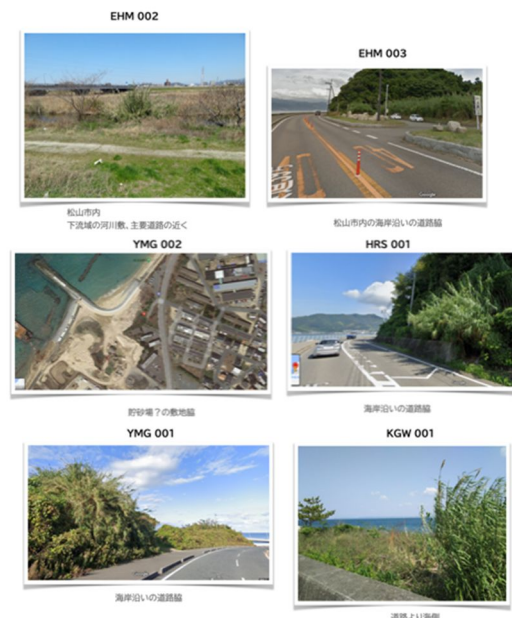


図13 各系統の採取地

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計6件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 永瀬 真, 吉川 朝彦, 堀澤 栄, 井上 雅裕, 佐久間 洋
2. 発表標題 四国における Google Map を用いたダンチクの分布調査 環境要因
3. 学会等名 日本植物学会 第86回大会
4. 発表年 2022年

1. 発表者名 福壽 慶太, 永瀬 真, 吉川 朝彦, 森脇 直人, 堀澤 栄, 井上 雅裕, 佐久間 洋
2. 発表標題 MS Azure Custom Visionを用いたダンチクコロニーの自動検出システム
3. 学会等名 中国四国植物学会 第78回大会
4. 発表年 2022年

1. 発表者名 佐久間 洋・井下 裕貴・堀澤 栄・井上 雅裕
2. 発表標題 ゲノムシーケンスに基づいたダンチク(Arundo donax)の HRM-SSR によるジエノ タイピング新規プライマーの検索
3. 学会等名 中国四国植物学会 第79回大会
4. 発表年 2023年

1. 発表者名 永瀬 真, 吉川 朝彦, 森脇 直人, 堀澤 栄, 井上 雅裕, 佐久間 洋
2. 発表標題 四国におけるGoogle Mapを用いたダンチクの分布調査 -画像認識サービスを用いた自動検出の試み-
3. 学会等名 第77回中国四国植物学会大会
4. 発表年 2021年

1. 発表者名 ダンチク (Arundo donax) のHRMを用いたジェノタイピング用SSRマーカーの検索
2. 発表標題 佐久間 洋、久保 雅秀、井上 雅裕
3. 学会等名 日本植物学会 第85回大会
4. 発表年 2021年

1. 発表者名 永瀬 真, 吉川 朝彦, 森脇 直人, 堀澤 栄, 井上 雅裕, 佐久間 洋
2. 発表標題 四国におけるGoogle Mapを用いたダンチクの分布調査 -Azure Custom Vision Serviceを用いた検出-
3. 学会等名 日本植物学会 第84回大会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------