(Formerly MENDELIANA)

July 2024
Volume XXXV
Issue 1
E-ISNN: 1852-6233

BAG

Journal of Basic & Applied Genetics

Journal of the Argentine Society of Genetics
Revista de la Sociedad Argentina de Genética

www.sag.org.ar/jbag
Buenos Aires, Argentina

DOI: 10.35407/bag.2023.35.01

Journal of the
Argentine Society
of Genetics

BAG | Journal of Basic & Applied Genetics

V. XXXV - No. 1

July 2024

BAG - Journal of Basic and Applied Genetics

Not yet assigned quartile

SJR 2022
0

powered by scimagojr.com

Editorial Board
Comité Editorial

Editor General:

Dra. Elsa L. Camadro
Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata. Balcarce, Argentina
bag.editor@sag.org.ar

Editores Asociados:

Citogenética Animal y Citogenética Vegetal

Dra. Liliana Mola
Depto. de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Buenos Aires, y Consejo Nacional de Investigaciones Científicas y Técnicas. Buenos Aires, Argentina
limola@ege.fcen.uba.ar

Dra. Mariel Schneider
Dep. de Ciências Biológicas, Universidade Federal de São Paulo. São Paulo, Brasil
maricb@rc.unesp.br

Citogenética Vegetal

Dr. Julio R. Daviña
Instituto de Biología Subtropical, Universidad Nacional de Misiones. Posadas, Argentina
juliordavina@fceqyn.unam.edu.ar

Genética de Poblaciones y Evolución

Dra. Mariana Pires de Campos Telles
Dep. de Genética, Laboratório de Genética & Biodiversidade, Escola de Ciências Médicas e Vida, Pontifícia Universidade Católica de Goiás e Universidade Federal de Goiás. Goiás, Brasil
tellsmpc@gmail.com

Dra. María Isabel Remis
Depto. de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, y Consejo Nacional de Investigaciones Científicas y Técnicas. Buenos Aires, Argentina
mariar@ege.fcen.uba.ar

Dr. Juan César Vilardi
Depto. de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Buenos Aires, y Consejo Nacional de Investigaciones Científicas y Técnicas. Buenos Aires, Argentina
vilardi@bg.fcen.uba.ar

Genética Humana, Genética Médica, y Citogenética

Dra. María Inés Echeverría
Instituto de Genética, Facultad de Ciencias Médicas, Universidad Nacional de Cuyo. Mendoza, Argentina
miecheve@fcm.uncu.edu.ar

Genética Humana

Dr. Carlos Bacino
Dept. of Molecular and Human Genetics, Baylor College of Medicine. Texas, USA
cbacino@bcm.edu

Genética Médica

Dr. José Arturo Prada Oliveira
Facultad de Medicina, Departamento de Anatomía Humana y Embriología, Universidad de Cádiz. Cádiz, España
arturo.prada@uca.es

Genética Médica y Molecular

Dr. Bernardo Bertoni Jara
Facultad de Medicina, Universidad de la República. Montevideo, República Oriental del Uruguay
bbertoni@fmed.edu.uy

Dra. Mev Domínguez Valentín
Oslo University Hospital. Oslo, Norway
mev.dominguez.valentin@rr-research.no

Genética Molecular Animal

Dr. Guillermo Giovambattista
Instituto de Genética Veterinaria, Facultad de Ciencias Veterinarias, Universidad Nacional de La Plata, y Consejo Nacional de Investigaciones Científicas y Técnicas. La Plata, Argentina
ggiovam@fcv.unlp.edu.ar

Genética Molecular Vegetal

Dr. Alberto Acevedo
Centro de Investigación de Recursos Naturales, Instituto Nacional de Tecnología Agropecuaria. Hurlingham, Argentina
acevedo.alberto@inta.gob.ar

Dr. Andrés Zambelli
Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata, y Consejo Nacional de Investigaciones Científicas y Técnicas. Balcarce, Argentina
andres.zambelli@mdp.edu.ar

Genética y Mejoramiento Genético Animal

Dra. María Inés Oyarzábal
Facultad de Ciencias Veterinarias, Universidad Nacional de Rosario. Rosario, Argentina
moyazabr@unr.edu.ar

Dr. Gustavo Rodríguez Reynoso
Universidad Agraria La Molina, y Consejo Nacional de Ciencia y Tecnología. Lima, Perú
gustavogr@lamolina.edu.pe

Genética y Mejoramiento Genético Vegetal

Dra. Natalia Bonamico
Facultad de Agronomía y Veterinaria, Universidad Nacional de Río Cuarto. Río Cuarto, Argentina
nbonamico@ayv.unrc.edu.ar

Dr. Ricardo W. Masuelli
Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo, y Consejo Nacional de Investigaciones Científicas y Técnicas. Mendoza, Argentina
rmasuelli@fca.uncu.edu.ar

Dr. Rodomiro Ortiz
Dept. of Plant Breeding, Swedish University of Agricultural Science. Uppsala, Suecia
rodomiro.ortiz@slu.se

Dra. Mónica Poverene
Depto. de Agronomía, Universidad Nacional del Sur. Bahía Blanca, Argentina
poverene@criba.edu.ar

Dr. Pedro Rimieri
Profesional asociado y asesor científico-técnico. INTA, Pergamino. Buenos Aires, Argentina
primieri730@gmail.com

Genética de Microorganismos

Dra. Mariel Sanso
Facultad de Ciencias. Veterinarias, Universidad Nacional del Centro de la Provincia de Buenos Aires. Tandil, Argentina
msanso@vet.unicen.edu.ar

Mutagénesis

Dr. Alejandro D. Bolzán
Lab. de Citogenética y Mutagénesis, Instituto Multidisciplinario de Biología Celular, y Consejo Nacional de Investigaciones Científicas y Técnicas. La Plata, Argentina
abolzan@imbice.gov.ar

Consultor Estadístico

Dr. David Almorza
Facultad de Ciencias del Trabajo, Depto. de Estadística e Investigación Operativa, Universidad de Cádiz. Cádiz, España
david.almorza@uca.es

Secretaria de Redacción

Dra. María de las Mercedes Echeverría
Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata. Balcarce, Argentina
mecheverria@mdp.edu.ar

Diseño y maquetación
Lic. Mauro Salerno
maurosalerno92@gmail.com

Corrección de estilo
Dra. Gabriela Leofanti
gabrielaleofanti@gmail.com

Comité de colaboradores

Dra. María Mercedes Ibañez
Facultad de Agronomía y Veterinaria,
Universidad Nacional de Río Cuarto. Río Cuarto, Córdoba, Argentina
mibanez@ayv.unrc.edu

Dr. Daniel Maizon
Estación Experimental Agropecuaria “Guillermo Covas”, Instituto Nacional de Tecnología Agropecuaria, y Universidad Nacional de La Pampa. Anguil, La Pampa, Argentina
maizon.daniel@inta.gob.ar

Dra. María Andrea Tomas
Estación Experimental Agropecuaria Rafaela, Instituto Nacional de Tecnología Agropecuaria; Consejo Nacional de Investigaciones Científicas y Técnicas; Universidad Nacional de Rafaela. Rafaela, Santa Fe, Argentina
tomas.maria@inta.gob.ar

Imágen de tapa:
Antesala de la fecundación. Cruzamiento compatible en papa silvestre: tubos polínicos creciendo hacia la base del pistilo.
Autora: **Gabriela Agustina Leofanti**

Contents
Contenidos

Jan.-July 2024
Enero-Julio 2024

ARTICLE 1
RESEARCH
9 - 19

REVISIÓN DE LA DISPLASIA ECTODÉRMICA HIPOHIDRÓTICA: DIAGNÓSTICO, TRATAMIENTOS Y TERAPIAS GÉNICAS

HYPOHIDROTIC ECTODERMAL DYSPLASIA: A REVIEW ON DIAGNOSIS, TREATMENT, AND GENE THERAPIES

Herrador T.A., Migotti J.S.

ARTICLE 2
RESEARCH
21 - 38

INFLUENCE OF DEMOGRAPHIC PARAMETERS AND DIETARY HABITS ON THE CYTOME ASSAY BIOMARKERS IN LYMPHOCYTES AND BUCCAL EPITHELIAL CELLS FROM A GROUP OF ARGENTINE ADOLESCENTS

INFLUENCIA DE PARÁMETROS DEMOGRÁFICOS Y HÁBITOS ALIMENTARIOS SOBRE BIOMARCADORES DEL ENSAYO CITOMA EN LINFOCITOS Y CÉLULAS EPITELIALES BUCALES DE ADOLESCENTES ARGENTINOS

Martínez Perafán F., Casciaro S., López Nigro M.M., Carballo M.A.

ARTICLE 3
REVIEW
39 - 51

***RANDOM FOREST* IN PLANT GENETICS AND BREEDING: AN APPLICATION IN TOMATO AS A MODEL CROP**

RANDOM FOREST EN GENÉTICA Y MEJORAMIENTO GENÉTICO DE PLANTAS: UNA APLICACIÓN EN TOMATE COMO CULTIVO MODELO

Faviere G., Vitelleschi M.S., Pratta G.R.

ARTICLE 4
REVIEW
53 - 64

DEVELOPMENT OF GENETIC STOCKS OF SUNFLOWER WITH RESISTANCE TO SUNFLOWER CHLOROTIC MOTTLE VIRUS

CONFORMACIÓN DE RECURSOS GENÉTICOS DE GIRASOL CON RESISTENCIA AL VIRUS DEL MOTEADO CLORÓTICO DEL GIRASOL

Fernández Moroni I., Lenardon S., Giolitti F., Álvarez D., Poverene M., Presotto A., Cantamutto M.

ARTICLE 5
REVIEW
65 - 72

REPRODUCTIVE BIOLOGY AND POLLEN-PISTIL COMPATIBILITY RELATIONSHIPS IN AN ARGENTINIAN COLLECTION OF *Stevia rebaudiana* BERTONI

BIOLOGÍA REPRODUCTIVA Y RELACIONES DE COMPATIBILIDAD POLEN-PISTILO EN UNA COLECCIÓN DE *Stevia rebaudiana* BERTONI EN ARGENTINA

Budeguer C.J., Camadro E.L., Erazzú L.E.

MEMOIR
73 - 74

DR. NÉSTOR OSCAR BIANCHI

1931-2023

BAG | Journal of Basic & Applied Genetics

Vol XXXV (1): 9-19; Jan.-July 2024

Journal of the Argentine Society of Genetics

REVISIÓN DE LA DISPLASIA ECTODÉRMICA HIPOHIDRÓTICA: DIAGNÓSTICO, TRATAMIENTOS Y TERAPIAS GÉNICAS

HYPOHIDROTIC ECTODERMAL DYSPLASIA: A REVIEW ON DIAGNOSIS, TREATMENT, AND GENE THERAPIES

Herrador T.A.[1,2], Migotti J.S.[1,3]

[1] Instituto Universitario de Ciencias Biomédicas de Córdoba (IUCBC). Naciones Unidas 420, Córdoba, Argentina.

[2] tiago.alejo.h@hotmail.com

[3] julimigottiscaglia@gmail.com

Corresponding author:
Tiago Alejo Herrador
tiago.alejo.h@hotmail.com

ORCID 0009-0007-3160-3751

Cite this article as:
Herrador T.A., Migotti J.S. 2024. REVISIÓN DE LA DISPLASIA ECTODÉRMICA HIPOHIDRÓTICA: DIAGNÓSTICO, TRATAMIENTOS Y TERAPIAS GÉNICAS. BAG. Journal of Basic and Applied Genetics XXXV (1): 9-19.

Received: 10/22/2023
Revised version received: 02/02/2024
Accepted: 03/08/2024
General Editor: Elsa Camadro
DOI: 10.35407/bag.2024.35.01.01
ISSN online version: 1852-6233

ABSTRACT

Hypohidrotic ectodermal dysplasia (HED) is a genodermatosis caused by a monogenic defect that affects the protein ectodysplasin in tissues derived from the ectoderm. Depending on the transmission mechanism, we can distinguish three types of HED: autosomal recessive, autosomal dominant and linked to the X chromosome (this being the most frequent). Although the phenotypic expression varies according to which gene is affected, all forms of HED share three clinical characteristics: hypodontia, hypohidrosis and hypothricosis. The conventional treatments have the objective of improving the patient's quality of life, and include dental interventions such as orthodontics and dental implants, dermatological routines for dry skin, and the use of artificial sweat. In recent years, scientific research has focused on genetic engineering to treat HED, developing new promising strategies such as gene therapy with recombinant EDA1, or ER004, developed by Schneider, H. and colleagues. In someclinical trials this strategy effectively reversed clinical manifestations of the disease. Throughout the review we will address HED's symptoms, tools for the correct diagnosis, available treatments and considerations for the differential diagnosis.

Key words: ectodysplasin, ER004, genetic engineering, recombinant protein

RESUMEN

La displasia ectodérmica hipohidrótica (DEH) es una genodermatosis producida por un defecto monogenético que afecta la proteína ectodisplasina en los tejidos derivados del ectodermo. De acuerdo al mecanismo de transmisión, podemos distinguir tres tipos de DEH: autosómica recesiva, autosómica dominante, o ligada al cromosoma X (siendo ésta la más frecuente). Si bien la expresión fenotípica presenta diferencias de acuerdo al gen afectado, las tres características clínicas de la DEH son: hipodoncia, hipotricosis e hipohidrosis. El tratamiento convencional busca mejorar la calidad de vida del paciente, abarcando intervenciones dentales como ortodoncia e implantes dentales, rutinas dermatológicas para piel seca y escamosa, y el uso de sudor artificial. En los últimos años, la investigación científica desplazó el interés desde estos tratamientos hacia la ingeniería genética, surgiendo alternativas prometedoras que en ciertos ensayos clínicos lograron revertir la enfermedad, como lo es la terapia génica con EDA1 recombinante, o ER004, desarrollada por Schneider, H. A lo largo de la revisión se abordarán los síntomas de la enfermedad, las herramientas para el correcto diagnóstico y los tratamientos disponibles, así como otras condiciones a tener en cuenta en el diagnóstico diferencial.

Palabras clave: ectodisplasina, ER004, ingeniería genética, proteína recombinante

Available online at www.sag.org.ar/jbag

INTRODUCCIÓN

La displasia ectodérmica hipohidrótica (DEH) es una enfermedad genética del tipo genodermatosis, que produce alteraciones en la proteína transmembrana ectodisplasina, lo que conlleva a defectos en dos o más tejidos derivados del ectodermo embrionario. De esta forma, la DEH se caracteriza por la manifestación clínica de hipotricosis (pérdida o ausencia del cabello y vello corporal), hipohidrosis o anhidrosis (disminución o ausencia de la sudoración) que conlleva a hipertermia crónica, y anomalías dentarias de número, forma y tamaño (Araníbar *et al.*, 2005; Marín Botero *et al.*, 2013). No sólo es una enfermedad poco común, presentándose en aproximadamente uno de cada diez mil recién nacidos, sino que comúnmente pasa desapercibida por el personal médico que no tiene las herramientas necesarias para la identificación y correcto diagnóstico de la misma (por lo que es probable que su incidencia en la población sea aún mayor).

Al ser un defecto monogenético, los tratamientos convencionales no buscan curar la enfermedad sino mejorar la calidad de vida de quien la padece, desde un enfoque interdisciplinario que articula la odontología, dermatología y psicología, entre otras ramas de las ciencias de la salud. Sin embargo, en los últimos años, gracias a los avances en el campo de la ingeniería genética, han surgido alternativas terapéuticas con un prospecto prometedor que buscan revertir la enfermedad, como lo es la administración prenatal de proteínas artificiales que suplanten la carencia de ectodisplasina en embriones con DEH.

Esta revisión está orientada hacia el personal médico, teniendo como objetivo brindar herramientas para la identificación y diagnóstico de la DEH, así como recopilar las distintas vías terapéuticas existentes en la actualidad. Se aborda la sintomatología clínica, los mecanismos genéticos y patrones de herencia, las técnicas diagnósticas y los tratamientos convencionales y genéticos disponibles, así como también otros diagnósticos diferenciales a tener en cuenta en la consulta clínica.

MECANISMOS DE TRANSMISIÓN

Se conocen tres mecanismos distintos de transmisión hereditaria de la DEH: ligada al cromosoma X (XDEH o LXDEH), autosómica recesiva, y autosómica dominante. (Araníbar *et al.*, 2005).La XDEH es la forma más común de DEH. Se produce por una mutación en el gen *EDA*, encontrado en el cromosoma X, en el locus Xq13, encargado de codificar la proteína transmembrana ectodisplasina (Araníbar *et al.*, 2005). Al ser ligada al cromosoma X, afecta principalmente al hombre mientras que las mujeres suelen ser portadoras. Una mujer heterocigota para una variante *EDA* patogénica transmite la enfermedad a su descendencia en un 50%, mientras que un hombre la misma variante variante la transmite a todas sus hijas y a ningún hijo (Wright *et al.*, 2022).

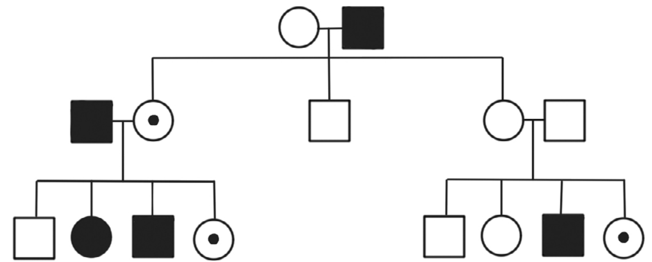


Figura 1. Árbol genealógico de la herencia de la displasia ectodérmica hipohidrótica ligada al cromosoma X.

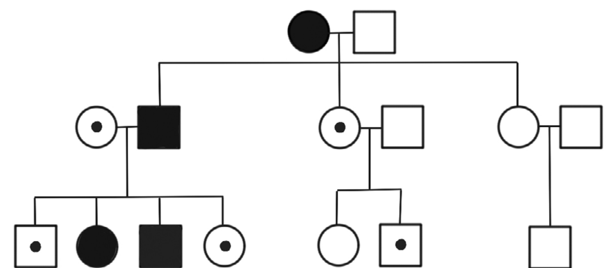


Figura 2. Árbol genealógico de la herencia de la displasia ectodérmica autosómica recesiva.

Por otro lado, la DEH autosómica puede ser recesiva o dominante. En la DEH autosómica recesiva, los padres de un infante afectado son heterocigotos para una variante patogénica de los genes *EDAR*, encontrado en el cromosoma dos, en el locus 2q13, *EDARADD*, encontrado en el cromosoma uno, en el locus 1q42.3-q43, o *WNT10A*, encontrado en el cromosoma dos, en el locus 2q35. La descendencia de dos padres heterocigotos tiene un 25% de posibilidades de expresar la enfermedad, un 50% de posibilidades de ser heterocigoto, y un 25% de posibilidades de no heredar ninguna variante patogénica. En la DEH autosómica dominante, la descendencia de un individuo con la enfermedad tiene 50% de posibilidades de heredar la variante patogénica EDAR, EDARADD o WNT10A, y por ende expresar la enfermedad. (Wright et al., 2022)

La Figura 4 corresponde a un ideograma del cariotipo humano, donde se ven representados esquemáticamente los 46 cromosomas con sus patrones de bandas G. A su vez, se encuentran señalizados los genes EDA, EDAR, EDARADD y WNT10A en sus respectivos locus Xq13, 2q13, 1q42.2-q43 y 2q35.

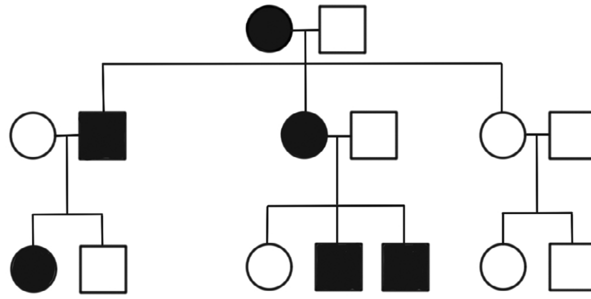


Figura 3. Árbol genealógico de la herencia de la displasia ectodérmica autosómica dominante.

Figura 4. Ideograma del cariotipo humano. Los genes *EDA, EDAR, EDARADD* y *WNT10A* se encuentran señalizados.

Tabla 1. Patrones y mecanismos de transmisión de la DEH. Adaptado de Wright *et al.*, 2022.

Transmisión	Varón afectado	Mujer afectada	Mutación	Fenotipo
Ligada al X	Transmite únicamente a hijas mujeres	50% de probabilidad de transmisión	Gen *EDA*	Desde DEH clásico hasta hipodoncia no sindrómica
Autosómica recesiva	La descendencia de dos individuos heterocigotos tiene un 25% de expresar la enfermedad, un 50% de ser heterocigoto y un 25% de no heredar la variante patogénica	La descendencia de dos individuos heterocigotos tiene un 25% de expresar la enfermedad, un 50% de ser heterocigoto y un 25% de no heredar el patógeno	Gen *EDAR*, *EDARADD* o *WNT10A*	Fenotipo EDAR de leve a severo; reporte de casos de amastia e hiperqueratosis palmoplantar Fenotipo WNT10A variable, relacionado con dentición anormal junto a síntomas ectodérmicos leves o agenesia dental selectiva
Autosómica dominante	50% de probabilidad de transmisión	50% de probabilidades de transmisión	Gen *EDARADD*, *WNT10A* o *novoEDAR*	Fenotipo WNT10A variable, relacionado con dentición anormal junto a síntomas ectodérmicos leves o agenesia dental selectiva

CRITERIO DIAGNÓSTICO Y MANIFESTACIONES CLÍNICAS

La DEH debe tenerse en cuenta para el diagnóstico diferencial en individuos que presenten:

- **Hipotricosis**: disminución y ligera pigmentación de cabello y vello corporal, así como tallos delgados, quebradizos y/o retorcidos; sin embargo, estas últimas dos características no son suficientemente sensibles para considerarse de beneficio a diagnóstico (Rouse *et al.*, 2004). El vello sexual secundario puede presentarse de manera normal.

- **Hipohidrosis**: sudoración reducida o ausente que conlleva a hipertermia crónica. Aunque se pueden realizar biopsias de la piel para poder determinar tanto la morfología como la distribución de las glándulas sudoríparas en el paciente, las técnicas no invasivas presentan la misma tasa de efectividad. Se puede evaluar el funcionamiento de las glándulas sudoríparas al poner en contacto con solución de yodo a la piel del paciente y elevar la temperatura ambiente para inducir la sudoración: si se expone al sudor, el yodo reacciona cambiando de color, posibilitando evaluar la cantidad y ubicación de la sudoración. A su vez, el número y la distribución de poros pueden ser determinados usando materiales de impresión para cubrir partes del cuerpo del paciente (por lo general, se cubren las eminencias hipotenares de las palmas), o utilizando microscopio confocal en vivo para visualizar los conductos de sudor en las palmas (Jones *et al.*, 2013). Estas técnicas son particularmente útiles para el diagnóstico de mujeres heterocigotas, ya que éstas presentan patrones de mosaico de función y distribución de las glándulas sudoríparas.

- **Hipodoncia**:anomalía dentaria caracterizada por la ausencia de hasta seis piezas dentales. Si bien en la DEH clásica se estima que se desarrollan alrededor de nueve dientes permanentes en la persona afectada, esto es altamente variable (Lexner *et al.*, 2007). Por lo general, los dientes suelen ser más pequeños de lo normal y presentar alteraciones morfológicas como coronas cónicas. El uso de imágenes radiológicas permite determinar el grado de hipodoncia en el paciente, siendo de gran utilidad en el diagnóstico de DEH leve. El alargamiento de la cámara pulpar (taurodontismo) se presenta con mayor regularidad en los molares de individuos con DEH que en la población no afectada.

PREVALENCIA

La prevalencia exacta de la enfermedad es desconocida. Dentro de las displasias ectodérmicas hay más de veinte tipos clínicos distintos, siendo la DEH la más común. Se estima que uno, entre cinco mil y diez mil recién nacidos presenta DEH, aunque muchos infantes portadores de la condición pueden pasar desapercibidos hasta edades más adultas donde las características de la DEH se vuelven más evidentes (Priolo y Laganà, 2001; Nguyen-Nielsen *et al.*, 2013; Wright *et al.*, 2022).

ASESORAMIENTO GENÉTICO

Como se muestra en la tabla 1,a DEH producida por mutación en el gen EDA se hereda de forma ligada al cromosoma X, mientras la DEH producida por mutación

en los genes EDAR, EDARADD o WNT10A se hereda de forma autosómica (recesiva o dominante).

- **DEH ligada al cromosoma X.** La posibilidad de transmisión de una madre heterocigota para la variante patogénica *EDA* a la descendencia es del 50%, mientras que un padre con el la variante patogénica *EDA* la transmitirá a todas sus hijas y a ninguno de los hijos. Los hombres que hereden la enfermedad la expresan, mientras que las mujeres son heterocigotas y pueden expresar manifestaciones de displasia ectodérmica. Para poder detectar heterocigotos femeninos se necesita haber identificado previamente la variante patogénica *EDA* en la familia, ya que mujeres con XHED presentan manifestaciones leves que pueden superponerse con características de la población general (por ejemplo, la hipodoncia es relativamente común en la población general, pudiendo pasar desapercibida), por lo que la detección basada en la exploración clínica es imprecisa.

- **Parental de un probando masculino.** Padres afectados no transmiten la enfermedad a sus hijos, por lo que el padre de un varón afectado no porta la enfermedad. Cuando una familia presenta más de un individuo afectado, la madre de un varón afectado es heterocigota. Si en una familia un único varón presenta la enfermedad, la madre puede ser heterocigota o no (en este caso, el varón tendrá la variante *novoEDA*), o presentar mosaicismo somático/germinal. Pruebas genéticas moleculares en la madre son recomendadas para confirmar el diagnóstico y así evaluar el riesgo de recurrencia de la enfermedad en la familia, aunque en algunos casos las pruebas de ADN de leucocitos maternos no logran detectar casos de mosaicismo somático ni variantes patogénicas presentes únicamente en células germinales.

- **Parental de un probando femenino.** Mujeres afectadas por la enfermedad pueden heredar la variante patogénica *EDA* del padre o de la madre, o puede presentar *novoEDA*. Se recomienda llevar a cabo pruebas genéticas moleculares en ambos padres.

- **Hermanos de un probando masculino.** Si la madre es heterocigota para una variante patogénica *EDA*, los varones que hereden la variante expresan la enfermedad y las mujeres serán heterocigotas. Si no se puede detectar la variante patogénica en la madre, se presume riesgo bajo, pero mayor que el de la población general, en los hermanos del probando debido a la posibilidad de mosaicismo.

- **Hermanos de un probando femenino.** Si la madre es heterocigota para una variante patogénica EDA, los varones que hereden la variante expresan la enfermedad y las mujeres serán heterocigotas. Si el padre presenta la variante patogénica, la transmite a todas sus hijas, pero no a sus hijos. Si no se puede detectar la variante patogénica ni en la madre ni en el padre, se presume riesgo bajo, pero mayor que el de la población general, en los hermanos del probando debido a la posibilidad de mosaicismo.

- **DEH autosómica recesiva.** Para que una persona presente DEH autosómica recesiva, los padres obligadamente son heterocigotos para una variante patogénica *EDAR*, *EDARADD* o *WNT10A*. En la descendencia de dos individuos heterocigotos, se tiene un 25% de probabilidad de presentar la enfermedad, un 50% de ser heterocigoto y un 25% de no heredar la variante patogénica.

- **Parental de un probando.** Tanto el padre como la madre del probando son heterocigotos para una variante patogénica *EDAR*, *EDARADD* o *WNT10A*. Luego de haberse realizado el diagnóstico en el probando, es recomendado que se lleven a cabo pruebas genéticas moleculares en ambos padres y así tener una evaluación del riesgo de recurrencia. Si se detecta la presencia de la variante patogénica en sólo uno de los padres, y pruebas de maternidad y paternidad biológica atestiguan la identidad de los padres, existe la posibilidad de que una de las variantes patogénicas del probando haya ocurrido como evento *de novo* en el probando o como evento postcigótico *de novo* en un padre con mosaicismo; adicionalmente si el probando es homocigoto debe considerarse una deleción de uno o más exones no detectado en el análisis de secuencia, o isodisomía uniparental.

- **Hermanos de un probando.** En la descendencia de dos individuos heterocigotos, se tiene un 25% de probabilidad de presentar la enfermedad, un 50% de ser heterocigoto y un 25% de no heredar la variante patogénica.

- **DEH autosómica dominante.** Las personas que portan la enfermedad tienen un 50% de probabilidades de transmitirla a la descendencia, por lo que algunos individuos afectados sólo tienen un padre afectado. Esta enfermedad puede ser resultado de una variante patogénica *EDARADD*, *WNT10A* o *novoEDAR*.

- **Padres de un probando.** Se recomienda realizar examen físico y pruebas genéticas moleculares para evaluar los padres de un probando con *novoEDAR*. Si no se identifica la variante patogénica en ninguno de los padres y se ha comprobado la identidad de los mismos, se debe considerar que el probando tenga la variante *de novo*, o que haya heredado una variante de un padre con mosaicismo.

Tabla 2. Sintomatología de enfermedades a tener en cuenta en el diagnóstico diferencial. Adaptado de Wright *et al.*, 2022.

Desorden	Online Mendelian Inheritance in Man (OMIM)	Manifestaciones dentales	Manifestaciones capilares	Manifestaciones dérmicas	Otras consideraciones
Síndrome EEC (gen *TP63*)	604292	Anomalías del crecimiento, número y forma de dientes	Cabello escaso	Piel seca y escamosa	Ectrodactilia y paladar o labio hundido No presenta alteraciones en glándulas sudoríparas
Ictiosis (gen *ABCA12*)	242500	No presenta	No presenta	Piel seca y escamosa	No presenta alteraciones en glándulas sudoríparas
Síndrome KID (gen *GJB2*)	148210	No presenta	Cabello escaso	Ictiosis, hiperqueratosis palmoplantar progresiva	Queratitis, fotofobia, sordera
Síndrome de Clouston (gen *GJB6*)	129500	No presenta	Cabello escaso o alopecia	Hiperqueratosis palmoplantar	Fotofobia, distrofia ungueal No presenta alteraciones en glándulas sudoríparas
Síndrome Trico-dento-óseo (gen *DLX3*)	190320	Defectos en el esmalte dental y taurodontismo	Cabello quebradizo	No presenta	Uñas quebradizas y anormalidades en el esqueleto
Incontinentia Pigmenti (gen *IKBKG*)	308300	Hipodoncia	Atrofia en el cabello	Ampollas, verrugas, hiperpigmentación	Anomalías oculares y en el sistema nervioso central
Síndrome de Witkop (gen *MSX1*)	189500	Hipodoncia	Cabello fino	No presenta	No presenta alteraciones en glándulas sudoríparas Disgenesia en las uñas

- **Hermanos de un probando.** El riesgo de recurrencia de la enfermedad cuando se identifica que uno de los padres del probando presenta la misma variante patogénica es del 50%, mientras que cuando no se puede detectar el gen patogénico en los padres se estima que la probabilidad de recurrencia es del 1% (existe la posibilidad teórica de mosaicismo en la línea parental).

CORRELACIÓN FENOTIPO-GENOTIPO

Las manifestaciones clínicas del paciente con DEH están directamente relacionadas con su genotipo. Pacientes con variantes patogénicas del gen *EDA* pueden expresar fenotipos que varían desde DEH clásico hasta hipodoncia no sindrómica. En este último, se cree que la variante patogénica es una variante sin sentido ubicada en la región que codifica el dominio del factor de necrosis tumoral. En la XDEH, los *EDA* patogénicos son variantes de pérdida de función que incluyen variantes sin sentido, inserciones y deleciones (Zhang *et al.*, 2011). Tanto los pacientes varones como mujeres que presentaron la variante *EDA* patogénica expresaron fenotipos asociados a la XDEH. Se ha reportado que análisis de inactivación del cromosoma X en mujeres portadoras del *EDA* patogénico ha revelado un patrón aleatorio que favorecía la inactivación del cromosoma X, pero el fenotipo expresado en estas pacientes no se correlacionó con el patrón (Martínez-Romero *et al.*, 2019). En pacientes con variantes patogénicas del gen *EDAR* se pueden manifestar fenotipos leves a severos (Chassaing *et al.*, 2006), y se han reportado dos casos de asociación del gen patogénico y de características de DEH junto con amastia e hiperqueratosis palmoplantar (Mégarbané *et al.*, 2008). Sin embargo, no es posible distinguir clínicamente a pacientes con genotipo *EDA* patogénico de aquellos con la variante *EDAR* patogénica (Martínez-Romero *et al.*, 2019). Se han reportado fenotipos variables en pacientes con variantes *WNT10A* patogénicas, siendo la homocigosidad de la variante sin sentido c.321C>A (p.Cys107Ter), aquella identificada con más frecuencia en pacientes con manifestaciones graves (Krøigård *et al.*, 2016). El *WNT10A* patogénico también se encuentra relacionado con la expresión de dentición anormal junto a síntomas ectodérmicos leves o agenesia dental selectiva (Mues *et al.*, 2014; Bergendal *et al.*, 2016).

DIAGNÓSTICO DE DISPLASIA ECTODÉRMICA HIPOHIDRÓTICA

La identificación de una variante patogénica de *EDA* hemicigota en un hombre afectado, o una variante patogénica *EDAR*, *EDARADD* o *WNT10A* bialélica en un hombre o una mujer afectada confirma el diagnóstico. Los síntomas que caracterizan a la enfermedad se evidencian en la infancia: el cabello se vuelve delgado y de crecimiento lento, la persona sufre de episodios de hipertermia, y sólo unos pocos brotes dentarios emergen y de manera tardía. La presentación leve de la enfermedad puede diagnosticarse identificando un EDA, *EDAR*, *EDARADD* o *WNT10A* heterocigoto en una mujer, o identificando un *EDAR*, *EDARADD* o *WNT10A* heterocigoto en un hombre. (Wright *et al.*, 2022).

Una de las herramientas para establecer el diagnóstico es el Registro de Testeo Genético (GTR), que provee información acerca de los test genéticos utilizados para confirmar la DEH. (Srivastava *et al.*, 2023). El gran número de displasias ectodérmicas dificulta a los investigadores y al personal médico diagnosticar correctamente al paciente. A su vez, la intensidad de la expresión fenotípica de las displasias ectodérmicas varía en gran medida, lo que resulta en una gran varianza en los grados de intensidad en las manifestaciones clínicas de una persona afectada a la otra. El correcto diagnóstico temprano ayuda no sólo a predecir la severidad de la enfermedad sino a reducir el impacto socioeconómico en el paciente y en sus allegados (Peschel *et al.*, 2022).

El diagnóstico prenatal de la enfermedad es posible. La presencia de menos de seis brotes dentarios en embriones de entre 20 semanas y 23 semanas y seis días, documentado con exploración ultrasonográfica, es característico de la enfermedad. (Schneider *et al.*, 2018). A su vez, mediante amniocentesis, se puede diagnosticar de manera directa la presencia de la variante *EDA* patogénica en el embrión (Schneider *et al.*, 2023a).

DIAGNÓSTICO DIFERENCIAL

Como se muestra en la tabla 2, se han registrado más de 200 tipos de displasias ectodérmicas, muchas de las cuales presentan síntomas en común, particularmente la hipodoncia, la intolerancia al calor y la ligera escasez de cabello (Wright *et al.*, 2019). En el diagnóstico diferencial del paciente, es necesario tener presente no sólo las distintas formas de herencia de DEH sino también las otras formas de displasia ectodérmica. Por ejemplo, la manifestación de anomalías en el desarrollo de las uñas (onicodisplasia) sugiere un diagnóstico distinto a la DEH (Wright *et al.*, 2022).

El síndrome de ectrodactilia-displasia ectodérmica-fisura (EEC) comparte características con la XDEH, pero se diferencia de la misma por presentar falta de dedos en las manos o pies (ectrodactilia) y un labio o paladar hendido. En pacientes con piel seca y escamosa, se puede considerar ictiosis, que a diferencia de la DEH generalmente no afecta a las glándulas sudoríparas ni manifiesta hipotricosis e hipodoncia. También es necesario tener en cuenta que ciertos medicamentos, daño a los nervios y enfermedades autoinmunes pueden generar hipohidrosis adquirida (Aftab *et al.*, 2023).

El síndrome de queratitis, ictiosis y sordera (KID) se caracteriza por escasez de cabello y anomalías en la piel, pudiendo presentar ictiosis, pero a diferencia de la DEH también se manifiesta queratitis (inflamación en la córnea), pudiendo provocar dolor, neovascularización, fotofobia y cicatrización corneal (Raghavon *et al.*, 2022).

La DEH 2, o síndrome de Clouston, a diferencia de la DEH, no manifiesta anomalías de la sudoración ni hipodoncia. A su vez se caracteriza por presentar alopecia de parcial a completa, distrofia ungueal e hiperqueratosis palmoplantar. En esta DEH, se ve afectado el gen *GJB6* (Xu *et al.*, 2022).

El síndrome trico-dento-óseo, si bien presenta manifestaciones dentales, como hipoplasia y taurodontismo, y capilares, no produce anomalías en la piel y en la transpiración. A su vez, los pacientes con este síndrome pueden presentar engrosamiento del hueso del cráneo, esclerosis de la base del cráneo, esclerosis tubular y obliteración de los senos frontal y mastoideo (Fazel *et al.*, 2021).

Otras enfermedades a considerar en el diagnóstico diferencial incluyen: el síndrome Incontinentia pigmenti, que a diferencia de la DEH presenta anomalías en el sistema oculares y del nervioso central (How *et al.*, 2022); el síndrome de Witkop, caracterizado por anormalidades dentales y en las uñas (Najmuddin *et al.*, 2020); y las displasias ectodérmicas 4 (gen *KRT85*), 7 (gen *KRT74*) y 9 (gen *HOXC13*) (Wright *et al.*, 2022).

ESTABLECIENDO EL DIAGNÓSTICO

Para confirmar el diagnóstico de DEH en el probando, pueden realizarse pruebas genéticas moleculares que identifiquen mutaciones en los genes responsables de la condición: *EDA*, *EDAR*, *EDARADD* o *WNT10A*. Estas pruebas, además de confirmar el diagnóstico, permiten identificar portadores dentro de la familia, sirviendo como herramienta para el asesoramiento genético y para evaluar el riesgo de recurrencia de la enfermedad (Aftab *et al.*, 2023).

Para la **DEH clásica**, es frecuente establecer el diagnóstico después de la infancia en aquellas personas afectadas con hipotricosis, hipohidrosis e hipodoncia. En un probando masculino, se debe identificar una variante *EDA* patogénica o probablemente patogénica

hemicigotica o *EDAR*, *EDARADD* o *WNT10A* bialélicas para confirmar el diagnóstico. En un probando femenino, se confirma el diagnóstico identificando *EDAR*, *EDARADD* o *WNT10A* bialélico (Wright *et al.*, 2022).

En un probando masculino con manifestaciones clínicas leves de la enfermedad, el diagnóstico de **DEH leve** es confirmado con la identificación de una variante patogénica *EDAR*, *EDARADD* o *WNT10A* heterocigoto. En un probando femenino, puede confirmarse el diagnóstico al identificar un patrón de mosaico en la función y distribución de los poros, junto con hipodoncia e historial familiar de XDEH; identificar *EDA* patogénico heterocigoto confirma el diagnóstico (Wright *et al.*, 2022).

Cuando los hallazgos clínicos del probando son clásicos y consistentes con la herencia ligada al cromosoma X, se puede realizar pruebas seriadas de un solo gen para confirmar el diagnóstico. Primero se analiza la secuencia de *EDA*, y si esto no identifica una variante patogénica se sigue con un análisis de deleción/duplicación de los genes *EDA*. En cambio, si la exploración clínica del probando sugiere diagnóstico de DEH autosómico o leve, se debe analizar la secuencia de *EDAR*, *EDARADD* y *WNT10A*, y si no se logra identificar una variante patogénica se sigue con un análisis de deleción/duplicación de los genes *EDA*. Si luego del análisis de las secuencias *EDA*, *EDAR*, *EDARADD* y *WNT10A* no se logra la identificación de una variante patogénica, se debe considerar la posibilidad de otra forma de displasia ectodérmica. (Wright *et al.*, 2022)

Puede considerarse realizar un panel multigénico que incluya *EDA*, *EDAR*, *EDARADD*, *WNT10A* y otro/s gen/es de interés para poder identificar la causa genética de la condición del probando. Es necesario tener en cuenta que tanto los genes que incluye el panel como la sensibilidad diagnóstica de la prueba varían según el laboratorio donde se realice (Wright *et al.*, 2022).

TRATAMIENTOS CONVENCIONALES Y TERAPIAS GÉNICAS

El tratamiento convencional para la DEH, al ser una enfermedad causada por una mutación genética, está enfocado en mejorar la calidad de vida del paciente.

- Para la **hipotricosis**, se recomienda el uso de pelucas y rutinas de cuidado para cabello seco y escaso. Se ha utilizado minoxidilo tópico en infantes, con resultados exitosos para el crecimiento del cabello (Lee *et al.*, 2013).Para la **hipohidrosis o anhidrosis**, se recomiendan rutinas de cuidado de la piel con el uso de productos específicos para eczemas y piel seca, así como constante hidratación y disminuciones físicas de la temperatura corporal (uso de ventilador, aire acondicionado, paños húmedos). Rutinas regulares del cuidado de la piel y el uso de protector solar son recomendados, así como la derivación a un dermatólogo para mayor beneficio del paciente. La aplicación de sustitutos artificiales del sudor puede compensar la ausencia del sudor natural, asistiendo a regular la temperatura corporal (Aftab *et al.*, 2023).

- Para la **hipodoncia**, se recomienda que el tratamiento dental se inicie en edades tempranas. El tratamiento incluye restauraciones simples, unión de dientes cónicos, ortodoncia, implantes dentales en la porción anterior del arco mandibular en niños mayores, implantes dentales en adultos, y reemplazo de las prótesis dentales según sea necesario. Algunas consideraciones a tener en cuenta son que, tanto la unión de dientes cónicos como los implantes dentales mejoran la dentición y la habilidad de masticar, así como también la estética oral; los implantes dentales en la porción anterior del arco mandibular han resultado exitosos únicamente en infantes menores de siete años; los infantes con prótesis dentales típicamente necesitan reemplazarlas cada 2,5 años. Estas intervenciones dentales se deben acompañar de controles regulares en el desarrollo de los dientes, junto a un consumo frecuente de agua y suplementos de flúor para evitar formaciones de caries dentales. Cuando las anomalías dentales impidan la correcta masticación y deglución, se puede buscar asesoramiento dietético y terapia del habla y lenguaje (Aftab *et al.*, 2023).

Se recomiendan evaluaciones dentales cada seis a doce meses, y evaluaciones de las manifestaciones cutáneas, capilares, oftalmológicas y respiratorias anualmente y/o según la necesidad de la persona afectada. Las pacientes embarazadas que presenten DEH son consideradas de alto riesgo. En las visitas prenatales se debe discutir la necesidad de evitar aquellas situaciones con la potencialidad de aumentar el riesgo de hipertermia, así como una nutrición prenatal óptima y el uso de vitaminas prenatales (Aftab *et al.*, 2023)

Es esencial proporcionar asesoramiento genético y psicológico a los pacientes con DEH y a sus familias. Sesiones con profesionales de la salud mental pueden servir para navegar las dificultades emocionales y psicológicas de vivir con un trastorno genético raro, así como para brindar un espacio para poder compartir experiencias y consejos. A su vez, el asesoramiento genético permite conocer los patrones de herencia, riesgo de recurrencia de la enfermedad en la familia y las opciones reproductivas (Aftab *et al.*, 2023).

Si bien las terapias tradicionales no proporcionan una cura para la enfermedad, estos tratamientos buscan mejorar la calidad de vida con un enfoque interdisciplinario, involucrando diferentes áreas como odontología, dermatología, oftalmología, psicología e inmunodeficiencia (Marín Botero *et al.*, 2013). En los

últimos años, debido al avance tecnológico y en el estudio del genoma humano, así como al auge de la ingeniería genética, han surgido alternativas terapéuticas para defectos monogénicos que involucran a la genética. Dentro de éstas, se encuentran la edición genética, el reemplazo de genes y la inserción de proteínas artificiales en el paciente. Sin embargo, con estas nuevas vías de tratamiento se suscita un fuerte debate sobre la ética y legalización de estos procedimientos, sobre el cual la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) ha establecido una serie de recomendaciones mundiales para regular la correcta práctica e implementación de las terapias génicas. Estos lineamientos buscan respetar la individualidad e integridad del paciente, y establecen que sólo se deben llevar a cabo estos tratamientos luego de una pre evaluación rigurosa sobre los beneficios. Estas terapias génicas buscan la corrección de genes con expresiones fenotípicas incompatibles con el normal funcionamiento del ser humano, pero tienen la potencialidad de expresar efectos secundarios indeseados a largo plazo, como la mutagénesis de genes supresores de tumores y cáncer (Mikkelsen *et al.*, 2019; Yung *et al.*, 2021).

Recientemente se han llevado a cabo estudios que lograron resultados prometedores en la prevención y reversión de la DEH. La investigación de Margolis *et al.*, 2019, en modelos de perros tratados prenatalmente con EDA1 recombinante demostró un mejoramiento en los resultados fenotípicos, incluyendo tasas de crecimiento, lagrimeo, crecimiento del cabello, formación de glándulas de Meibomio, sudoración, dentición y aclaramiento mucociliar (Margolis *et al.*, 2019). Hermès *et al.*, 2014, administraron por vía intraamniótica EDA1 recombinante en ratones prenatales, y los resultados mostraron que los ratones inyectados presentaban cambios fenotípicos visibles en contraste con el grupo de control, incluyendo pelaje denso, distribución normal del cabello alrededor de los ojos, colas de forma normal, manchas oscuras en las patas, apertura normal de los ojos, y la presencia de glándulas ecrinas confirmada mediante evaluaciones histológicas (Hermes *et al.*, 2014).

También se han llevado a cabo ensayos clínicos en pacientes humanos con EDA1 recombinante. Schneider *et al.*, 2023b, administraron la proteína prenatalmente vía líquido amniótico en seis niños con la enfermedad, y lograron inducir el desarrollo de glándulas sudoríparas funcionales de manera consistente. Los niños recuperaron la habilidad de sudar con normalidad hasta ahora, habiendo pasado más de cinco años, y se comprobó que no hubo respuesta inmunitaria detectable al fármaco administrado (Schneider *et al.*, 2023b). Esto convirtió a esta terapia en una estrategia prometedora para la DEH. Sin embargo, la eficiencia de este tratamiento para la hipodoncia es subóptima, y no se han observado mejoras en el desarrollo del cabello. Es necesario tener presente que el desarrollo de esta intervención farmacológica todavía está en curso, por lo que aún no se encuentra aprobado ni ampliamente disponible.

ER004, previamente conocida como EDA1 recombinante o EDI200, es una proteína de fusión recombinante completamente humanizada que consta del dominio de unión al receptor de EDA1 y la región cristalizable (Fc) del fragmento de la inmunoglobulina humana G1 (IgG1), que permite mejorar la formación de hexámeros y la señalización del receptor, y también cumple la función de estabilizador de la proteína *in vivo*. Se ha logrado demostrar que al administrar ER004 en murinos, caninos y humanos, ésta se une con alta afinidad y especificidad a EDAR pero no a otros receptores de la familia del factor de necrosis tumoral (TNF).

Actualmente, Schneider *et al.*, 2023a, han comenzado el ensayo EDELIFE, un estudio clínico prospectivo, abierto, controlado por genotipo compatible y multicéntrico, para corroborar la eficacia y seguridad de la administración intraamniótica de ER004 como tratamiento prenatal en varones que padecen XDEH. Los participantes del estudio son mujeres embarazadas cuyos fetos masculinos hayan sido previamente diagnosticados de XDEH. Los criterios de aceptación son: madres mayores de 18 años al momento de dar su consentimiento, con embarazo confirmado a más tardar en la semana gestacional completa 23 + 6 días, y portadora conocida de una variante *EDA* patogénica; feto masculino con diagnóstico genético molecular de una mutación en uno de los alelos *EDA* maternos y documentación ultrasonográfica de menos de seis gérmenes dentales en el maxilar o la mandíbula entre la semana gestacional 20 o más, o con un diagnóstico directo de XDEH causado por *EDA* hemicigota mediante amniocentesis. A su vez, se buscó una cohorte de 20 familiares varones de entre seis meses y 75 años de edad con diagnóstico de XDEH para formar parte del grupo de control, que den consentimiento para someterse a pruebas genéticas durante la visita de estudio. (Schneider *et al.*, 2023a).

CONSIDERACIONES A FUTURO

Si bien las terapias génicas con EDA1 recombinante resultan prometedoras para prevenir e incluso revertir la DEH, la aprobación y legalización de las mismas sigue pendiente. En el debate bioético surgen preguntas como hasta qué punto estas intervenciones modifican a la herencia, y por ende a la especie humana, o quién puede decidir qué tratamiento está amparado por la ley y bajo qué régimen lo hace. Por otro lado, la distribución y el acceso a estas terapias tiene limitaciones económicas y geográficas: la administración de la proteína recombinante ER004, al ser una patente privada, se realiza en pocas clínicas en todo el mundo; y se estima

que el costo promedio de las terapias génicas ronda los dos millones de dólares por paciente(Marsden *et al.*, 2017) debido al gran costo de producción para las empresas farmacéuticas. A su vez, la documentación en los efectos secundarios a largo plazo es escasa, necesitando aún más tiempo para poder definir si son beneficiosas o ponen en riesgo la integridad del paciente.

CONCLUSIÓN

La DEH es una genodermatosis con tres posibles mecanismos de transmisión, siendo el más común la XDEH. La carencia de ectodisplasina deriva en la manifestación de la tríada hipotricosis, hipohidrosis e hipodoncia que caracteriza a la enfermedad. A lo largo de la revisión se expusieron las técnicas diagnósticas que confirman la enfermedad en pacientes con sospecha de DEH, como la identificación de mutaciones en los genes responsables de la enfermedad (*EDA*, *EDAR*, *EDARADD*, *WNT10A*) mediante técnicas de diagnóstico molecular y los paneles multigénicos que incluyan los genes de interés.

A su vez, se hizo una recopilación de los tratamientos actuales para la DEH, que incluyen terapias génicas y terapias convencionales interdisciplinarias. Si bien en la actualidad se está avanzando con la investigación en el área de la ingeniería genética para el desarrollo de las terapias génicas, estos tratamientos aún son poco accesibles para la población general y existe poca documentación de sus efectos secundarios a largo plazo, siendo potencialmente perjudiciales a la salud (por ejemplo, generando mutagénesis de genes supresores de tumores). Por otro lado, las vías terapéuticas convencionales consisten en el manejo de los síntomas del paciente con un enfoque interdisciplinario que abarca a la dermatología, odontología, psicología, entre otras.

A la hora de elegir qué tratamiento llevar a cabo es necesario que la prioridad sea la calidad de vida de la persona. Debe considerarse bien cuándo poner en juego la seguridad e integridad del paciente en pos de intentar mejorar el estado de su enfermedad, y cuándo acudir a tratamientos convencionales cuyos efectos a largo plazo ya son conocidos. A pesar de que las terapias génicas resultan prometedoras, aún son necesarios trabajos de investigación desde la genética y las ciencias de la salud que detallen sus posibles efectos secundarios, y si son realmente beneficiosas o si potencialmente agraven la salud integral de la persona.

BIBLIOGRAFÍA

Aftab H., Escudero I.A., Sahhar F. 2023. «X-Linked Hypohidrotic Ectodermal Dysplasia (XLHED): A Case Report and Overview of the Diagnosis and Multidisciplinary Modality Treatments». *Cureus*, junio. https://doi.org/10.7759/cureus.40383.

Aranîbar D.L., Lay-Son R.G., Sanz C.P., Castillo T.S. 2005. «Displasia Ectodérmica hipohidrótica, caso clínico y revisión de la literatura». *Revista chilena de pediatría* 76 (2): 166-72. https://doi.org/10.4067/S0370-41062005000200007.

Bergendal B., Norderyd J., Zhou X., Klar J., Dahl N. 2016. «Abnormal Primary and Permanent Dentitions with Ectodermal Symptoms Predict WNT10A Deficiency». *BMC Medical Genetics* 17 (1): 88. https://doi.org/10.1186/s12881-016-0349-4.

Chassaing N., Bourthoumieu S., Cossee M., Calvas P., Vincent M.C. 2006. «Mutations in EDAR Account for One-Quarter of Non-ED1-Related Hypohidrotic Ectodermal Dysplasia». *Human Mutation* 27 (3): 255-59. https://doi.org/10.1002/humu.20295.

Fazel M., Afshari E., Jarrahi N. 2021. Dental management of tricho-dento-osseous syndrome in adolescent patients: Literature review and case presentation. *Dental Research Journal* 18 (noviembre): 98. https://doi.org/10.4103/1735-3327.330879.

Hermes K., Schneider P., Krieg P., Dang A., Huttner K., Schneider H.. 2014. Prenatal Therapy in Developmental Disorders: Drug Targeting via Intra-Amniotic Injection to Treat X-Linked Hypohidrotic Ectodermal Dysplasia. *Journal of Investigative Dermatology* 134 (12): 2985-87. https://doi.org/10.1038/jid.2014.264.

How K.N., Leong H.J.Y., Pramono Z.A.D., Leong K.F., Lai Z.W., Yap WH. 2022. Uncovering incontinentia pigmenti: From DNA sequence to pathophysiology. *Frontiers in Pediatrics* 10 (septiembre): 900606. https://doi.org/10.3389/fped.2022.900606.

Jones K.B., Goodwin A.F., Landan M., Seidel K., Tran D.K., Hogue J., Chavez M., Fete M., Yu W., Husseon T., Johnson R., Huttner K., Jheon A.H., Klein O.D. 2013. Characterization of X-Linked Hypohidrotic Ectodermal Dysplasia (XL-HED) Hair and Sweat Gland Phenotypes Using Phototrichogram Analysis and Live Confocal Imaging. *American Journal of Medical Genetics. Part A* 161A (7): 1585-93. https://doi.org/10.1002/ajmg.a.35959.

Krøigård A.B., Clemmensen O., Gjørup H., Hertz J.M., Bygum A. 2016. Odonto-Onycho-Dermal Dysplasia in a Patient Homozygous for a WNT10A Nonsense Mutation and Mild Manifestations of Ectodermal Dysplasia in Carriers of the Mutation. *BMC Dermatology* 16 (marzo): 3. https://doi.org/10.1186/s12895-016-0040-7.

Lee, H.E., Chang I.K, Im M., Seo Y.J., Lee J.H., Lee Y. 2013. Topical Minoxidil Treatment for Congenital Alopecia in Hypohidrotic Ectodermal Dysplasia. *Journal of the American Academy of Dermatology* 68 (4): e139-40. https://doi.org/10.1016/j.jaad.2012.10.019.

Lexner M.O., Bardow A., Hertz J.M., Nielsen L.A., Kreiborg S. 2007. Anomalies of Tooth Formation in Hypohidrotic Ectodermal Dysplasia. *International Journal of Paediatric Dentistry* 17 (1): 10-18. https://doi.org/10.1111/j.1365-263X.2006.00801.x.

Margolis C.A., Schneider P., Huttner K., Kirby N., Houser T.P., Wildman L., Grove G.L, Schneider H., Casal M.L. 2019. Prenatal Treatment of X-Linked Hypohidrotic Ectodermal Dysplasia Using Recombinant Ectodysplasin in a Canine Model. *Journal of Pharmacology and Experimental Therapeutics* 370 (3): 806-13. https://doi.org/10.1124/jpet.118.256040.

Marín Botero M.L., Espinal Botero G., Arroyo Fuentes T.M., Posso Zapata M.V., Pérez M.D., Castañeda Peláez D.A., Sierra Pérez J.I. 2013. Displasia ectodérmica hipohidrótica: Reporte de casos. *Avances en Odontoestomatología* 29 (1): 11-23.

Marsden G., Towse A., Pearson S.D., Dreitlein B., Henshall C. 2017. Gene therapy: understanding the science, assessing the evidence, and paying for value: A Report from the 2016 ICER Membership Policy Summit. https://icer.org/wp-content/uploads/2020/10/ICER-Gene-

Therapy-White-Paper-030317.pdf.

Martínez-Romero M.C., Ballesta-Martínez M.J., López-González V., Sánchez-Soler M.J., Serrano-Antón A.T., Barreda-Sánchez M., Rodriguez-Peña L. 2019. EDA, EDAR, EDARADD and WNT10A Allelic Variants in Patients with Ectodermal Derivative Impairment in the Spanish Population. *Orphanet Journal of Rare Diseases* 14 (1): 281. https://doi.org/10.1186/s13023-019-1251-x.

Mégarbané H., Cluzeau C., Bodemer C., Fraïtag S., Chababi-Atallah M., Mégarbané A., Smahi A. 2008. Unusual Presentation of a Severe Autosomal Recessive Anhydrotic Ectodermal Dysplasia with a Novel Mutation in the EDAR Gene. *American Journal of Medical Genetics. Part A* 146A (20): 2657-62. https://doi.org/10.1002/ajmg.a.32509.

Mikkelsen R.B., Frederiksen H.R.S., Gjerris M., Holst B., Hyttel P., Luo Y., Freude K., Sandøe, P. 2019. Genetic protection modifications: moving beyond the binary distinction between therapy and enhancement for human genome editing. *The CRISPR Journal* 2(6): 362-369. https://doi.org/10.1089/crispr.2019.0024

MueS G., BondsJ., Xiang L., Vieira A.R., Seymen F., Klein O., D'Souza R.N. 2014. The WNT10A Gene in Ectodermal Dysplasias and Selective Tooth Agenesis. *American Journal of Medical Genetics. Part A* 164A (10): 2455-60. https://doi.org/10.1002/ajmg.a.36520.

Najmuddin M., Saheb S.A.K., Alharbi A.N., Alsobil F.M., Jhugroo C., Khan A.A., Divakar D.D., Naik S., Khanagar S.B. 2020. Autosomal Dominant Mutation of MSX1 Gene Causing Tooth and Nail Syndrome. *The Pan African Medical Journal* 36: 229. https://doi.org/10.11604/pamj.2020.36.229.23019.

Nguyen-Nielsen M., Skovbo S., Svaneby D., Pedersen L., Fryzek J.. 2013. The Prevalence of X-Linked Hypohidrotic Ectodermal Dysplasia (XLHED) in Denmark, 1995-2010. *European Journal of Medical Genetics* 56 (5): 236-42. https://doi.org/10.1016/j.ejmg.2013.01.012.

Peschel N., Wright J.T., Koster M.I., Clarke A.J., Tadini G., Fete M., Hadj-Rabia S., Sybert V.P., Norderyd J., Maier-Wohlfart S., Fete T.J., Pagnan N., Visinoni A.F., Schneider H. 2022. Molecular Pathway-Based Classification of Ectodermal Dysplasias: First Five-Yearly Update. *Genes* 13 (12): 2327. https://doi.org/10.3390/genes13122327.

Priolo M., Laganà C. 2001. Ectodermal dysplasias: a new clinical-genetic classification. *Journal of Medical Genetics.* https://jmg.bmj.com/content/38/9/579.

Raghavon U.N., Bhuptani N.V., Patel B.K. 2022. KID Syndrome: A Rare Congenital Ichthyosiform Disorder. *Indian Dermatology Online Journal* 13 (4): 505-7. https://doi.org/10.4103/idoj.idoj_651_21.

Rouse C., Siegfried E., Breer W., Nahass G. 2004. Hair and Sweat Glands in Families with Hypohidrotic Ectodermal Dysplasia: Further Characterization. *Archives of Dermatology* 140 (7): 850-55. https://doi.org/10.1001/archderm.140.7.850.

Schneider H., Faschingbauer F., Schuepbach-Mallepell S., Körber I., Wohlfart S., Dick A., Wahlbuhl M., Kowalczyk-Quintas C., Vigolo M., Kirby N., Tannert C., Rompel O., Rascher W., Beckmann M.W., Schneider P. 2018. Prenatal Correction of X-Linked Hypohidrotic Ectodermal Dysplasia. *New England Journal of Medicine* 378 (17): 1604-10. https://doi.org/10.1056/NEJMoa1714322.

Schneider H., Hadj-Rabia S., Faschingbauer F., Bodemer C., Grange D.K., Norton M.E., Cavalli R., Tadini G., Stepan H., Clarke A., Guillén-Navarro E., Maier-Wohlfart S., Bouroubi A., Porte F. 2023a. Protocol for the Phase 2 EDELIFE Trial Investigating the Efficacy and Safety of Intra-Amniotic ER004 Administration to Male Subjects with X-Linked Hypohidrotic Ectodermal Dysplasia. *Genes* 14 (1): 153. https://doi.org/10.3390/genes14010153.

Schneider H., Schweikl C., Faschingbauer F., Hadj-Rabia S., Schneider P. 2023b. A Causal Treatment for X-Linked Hypohidrotic Ectodermal Dysplasia: Long-Term Results of Short-Term Perinatal Ectodysplasin A1 Replacement. *International Journal of Molecular Sciences* 24 (8): 7155. https://doi.org/10.3390/ijms24087155.

Srivastava H., Singh C.K., Qureshi S.M.R., Mastud C.S. 2023. Hypohidrotic ectodermal dysplasia: A rare entity. *Journal of Oral and Maxillofacial Pathology: JOMFP* 27 (Suppl 1): S75-79. https://doi.org/10.4103/jomfp.jomfp_72_21.

Wright J.T., Grange D.K., Fete M. 2022. «Hypohidrotic Ectodermal Dysplasia». En *GeneReviews®*, editado por Margaret P. Adam, Ghayda M. Mirzaa, Roberta A. Pagon, Stephanie E. Wallace, Lora JH Bean, Karen W. Gripp, y Anne Amemiya. Seattle (WA): University of Washington, Seattle. http://www.ncbi.nlm.nih.gov/books/NBK1112/.

Wright J.T., Fete M., Schneider H., Zinse M., Koster M.I., Clarke A.J., Hadj-Rabia S., Tadini G., Pagnan N., Visinoni A.F., Bergendal B., Abbott B., Fete T., Stanford C., Butcher C. D´Souza R.N., Sybert V.P. Morasso M.I. 2019. Ectodermal Dysplasias: Classification and Organization by Phenotype, Genotype and Molecular Pathway. *American Journal of Medical Genetics. Part A* 179 (3): 442-47. https://doi.org/10.1002/ajmg.a.61045.

Xu Y., Wang M., Huang L., Hu J. 2022. The Clinical Manifestation of p.Asp50Asn Heterozygous Mutation of GJB2 Gene in 3 Members of a Family Is Similar to That of Clouston Syndrome. *Annals of Dermatology* 34 (5): 382-86. https://doi.org/10.5021/ad.20.278.

Yung N. K., Maassel N.L., Ullrich S.J., Ricciardi A.S., Stitelman D.H. 2021. A Narrative Review of in Utero Gene Therapy: Advances, Challenges, and Future Considerations. *Translational Pediatrics* 10 (5): 1486-96. https://doi.org/10.21037/tp-20-89.

Zhang J., Han D., Song S., Wang Y., Zhao H., Pan S., Bai B., Feng H. 2011. Correlation between the Phenotypes and Genotypes of X-Linked Hypohidrotic Ectodermal Dysplasia and Non-Syndromic Hypodontia Caused by Ectodysplasin-A Mutations. *European Journal of Medical Genetics* 54 (4): e377-382. https://doi.org/10.1016/j.ejmg.2011.03.005.

BAG | Journal of Basic & Applied Genetics

Vol XXXV (1): 21-38; Jan.-July 2024

Journal of the Argentine Society of Genetics

INFLUENCE OF DEMOGRAPHIC PARAMETERS AND DIETARY HABITS ON THE CYTOME ASSAY BIOMARKERS IN LYMPHOCYTES AND BUCCAL EPITHELIAL CELLS FROM A GROUP OF ARGENTINE ADOLESCENTS

INFLUENCIA DE PARÁMETROS DEMOGRÁFICOS Y HÁBITOS ALIMENTARIOS SOBRE BIOMARCADORES DEL ENSAYO CITOMA EN LINFOCITOS Y CÉLULAS EPITELIALES BUCALES DE ADOLESCENTES ARGENTINOS

Martínez Perafán F.[1], Casciaro S.[1,2], López Nigro M.M.[1,2], Carballo M.A.[1,2]

[1] Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Departamento de Bioquímica Clínica, Citogenética Humana y Genética Toxicológica (CIGETOX) Buenos Aires, Argentina.

[2] Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Instituto de Fisiopatología y Bioquímica Clínica (INFIBIOC), Buenos Aires, Argentina

Corresponding author:
Marta Ana Carballo
macarballo11@gmail.com

ORCID 0000-0001-6497-2544

Cite this article as:
Martínez Perafán F., Casciaro S., López Nigro M.M., Carballo M.A. 2024. INFLUENCE OF DEMOGRAPHIC PARAMETERS AND DIETARY HABITS ON THE CYTOME ASSAY BIOMARKERS IN LYMPHOCYTES AND BUCCAL EPITHELIAL CELLS FROM A GROUP OF ARGENTINE ADOLESCENTS. BAG. Journal of Basic and Applied Genetics XXXV (1): 21-38.

Received: 01/12/2024

Revised version received: 03/08/2024

Accepted: 03/11/2024

General Editor: Elsa Camadro

DOI: 10.35407/bag.2024.35.01.02

ISSN online version: 1852-6233

Available online at
www.sag.org.ar/jbag

ABSTRACT

The use of the cytome assay in monitoring studies on children has increased in recent years. For this reason, it is necessary to know the role of possible confounding factors that could affect its outcomes. The objective was to evaluate the influence of some demographic variables and diet on the baseline values of the cytome assay biomarkers in lymphocytes and buccal mucosa cells from a group of Argentine adolescents. Following the calculation of the biomarkers, a multivariate regression analysis including confounders was performed. In lymphocytes it was observed that micronuclei (MNi) had a negative association with a moderate consumption of roots and tubers, while the number of nuclear buds (NBUDs) was higher in minors not exposed to second-hand smoke (SHS). Regarding epithelial cells, MNi had a negative relationship with the intake of tropical fruits and red meat; on the contrary, this parameter increased with the moderate ingestion of legumes. In addition, oral NBUDs had a positive association with citrus and red meat consumption, whereas cereals and oil decreased its frequency. Furthermore, an increased number of binucleated cells was observed for adolescents who ate white meat and an increase in pyknotic cells for those exposed to SHS. These results revealed that in adolescents the baseline level of the cytome assay biomarkers, especially of those related to genotoxicity, can be influenced by exogenous variables, for instance, dietary habits. Thus, such factors need to be considered when carrying out biomonitoring studies on child populations.

Key words: baseline values, CBMN-cyt, confounding factors, individual food preferences, young population

RESUMEN

El uso del ensayo de citoma en estudios de seguimiento en niños se ha incrementado en los últimos años y resulta necesario conocer el papel de posibles factores de confusión que podrían afectar sus resultados. El objetivo fue evaluar la influencia de algunas variables demográficas y de la dieta en los valores basales de los biomarcadores de este ensayo en linfocitos y células de mucosa bucal de un grupo de adolescentes argentinos. Luego del cálculo de los biomarcadores, se realizó un análisis de regresión multivariada incluyendo factores de confusión. En linfocitos se observó que los micronúcleos (MN) se asociaron negativamente con un consumo moderado de raíces y tubérculos, mientras que los brotes nucleares (BrN) aumentaron en los menores no expuestos al humo de segunda mano (HSM). En células epiteliales, los MN tuvieron una relación negativa con el consumo de frutas tropicales y carnes rojas, aunque aumentaron con el consumo moderado de legumbres. Los BrN orales tuvieron una asociación positiva con el consumo de cítricos y carnes rojas, mientras que los cereales y el aceite disminuyeron su frecuencia. Además, se encontró un mayor número de células binucleadas para quienes comieron carne blanca y un aumento de células picnóticas para los expuestos a HSM. Estos resultados revelaron que en los adolescentes el nivel basal de los biomarcadores del ensayo de citoma, especialmente de aquellos relacionados con la genotoxicidad, puede verse influenciado por variables exógenas como los hábitos alimentarios. Por lo tanto, dichos factores deben considerarse al realizar estudios de biomonitoreo en poblaciones infantiles.

Palabras clave: valores basales, MNCB-cit, factores de confusión, preferencias alimentarias individual, población joven.

INTRODUCTION

In recent years, the number of human biomonitoring studies has risen following a substantial increase of interest in different environmental problems that affect people's health (WHO, 2015; Louro *et al.*, 2019). Within this group of studies those related to the child and adolescent population have gained importance since minors are more susceptible than adults to pollutants (Perlroth and Branco, 2017). In addition, there is evidence that associates early exposure to xenobiotics with the development of various chronic diseases in adulthood, including some related to genetic instability such as cancer (Landrigan *et al.*, 2004; WHO, 2011; Sly *et al.*, 2014).

Taking into account the relationship between environmental exposure to xenobiotics and its effect on DNA, different tests have been developed which measure biomarkers associated with this outcome (Beedanagari, 2014). One of the most popular approaches is the micronucleus test, especially the more comprehensive version known as cytome assay. This method evaluates a variety of biomarkers related to genotoxicity, chromosomal instability and cell death depending on the matrix in which it is carried out (Fenech, 2006; Sommer *et al.*, 2020).

The assay on a culture of lymphocytes with blocked cytokinesis, for example, evaluates micronuclei (MNi), nuclear buds (NBUDs) and nucleoplasmic bridges (NPBs). MNi are the result of chromosome fragments or whole chromosomes that are left behind during cell division, whereas NBUDs are related to the elimination of amplified DNA and/or DNA repair complexes, and NPBs are caused by dicentric chromosomes (telomere fusion) or by sister chromatids that cannot be separated due to defects in cohesins and/or separases (Fenech *et al.*, 2016). On the other hand, in buccal cells, apart from MNi and NBUDs, it is also possible to measure the frequency of binucleated (BN) cells as a result of cytokinesis failures and cells with nuclear anomalies such as condensed chromatin (CC), karyorrhexis (KHC), karyolysis (KYL) and pyknosis (PYK), which are associated with cell death processes (Thomas *et al.*, 2009).

Most studies in genetic toxicology report the use of the cytome assay in lymphocytes and buccal mucosa cells (BMC), with the cytokinesis-block micronucleus cytome (CBMN) assay in lymphocyte cultures being the version most used in genotoxicity assays and biomonitoring studies (Kirsch-Volders *et al.*, 2018; Nersesyan *et al.*, 2016). Indeed, understanding the molecular bases for formation of the biomarkers assessed by this approach have turned it into a robust test to such an extent that the frequency of micronuclei in this type of cells is considered as a predictor of cancer risk (Bonassi *et al.*, 2011b; Fenech *et al.*, 2016). On the other hand, the use of the cytome assay in BMC has recently increased, mainly due to the methodological advantages it provides. Firstly, sampling for this test is minimally invasive, which is ideal for working with children and people reluctant to venipuncture. Furthermore, this assay does not require cell culture, so it is easily carried out in laboratories without the necessary equipment for cell maintenance (Thomas *et al.*, 2009; Holland *et al.*, 2008).

Different observational studies on children and adolescents have used the cytome assay in lymphocytes or BMC in order to assess the effect of exposure to environmental pollutants (Neri *et al.*, 2006; Castañeda-Yslas *et al.*, 2016; Martínez-Perafán *et al.*, 2018; Lemos *et al.*, 2020; Panico *et al.*, 2020), and many others report the basal frequency of the biomarkers of this test in healthy young people (Neri *et al.*, 2005; Gajski *et al.*, 2013; Silva da Silva *et al.*, 2015). Following this, the relevance of the cytome assay in the biomonitoring of child population is evident; however, there is still a lack of information about the influence that some demographic, genetic and lifestyle factors could have on the levels of the biomarkers measured, especially in some of those evaluated in the buccal test (Bonassi *et al.*, 2011b). Indeed, some characteristics such as age, sex and diet could be confounding factors or moderating variables, which has already been analysed in the lymphocyte test (Fenech and Bonassi, 2011; Holland *et al.*, 2011). Therefore, it is crucial to carry out specific studies with a biomonitoring that evaluates the association between these factors and the frequency of the cytome assay markers in a multivariate model.

Following the aforementioned, this study had two objectives: first, to determine the baseline levels of the biomarkers of the CBMN assay in lymphocytes and of the buccal micronucleus cytome (BMCyt) assay in a group of Argentine adolescents not environmentally exposed to chemical agents except for second-hand smoke (SHS). Second, to evaluate the possible association between the main demographic variables of the group and their dietary habits, and the levels of the biomarkers of both tests in a multivariate model.

MATERIALS AND METHODS

Study population and ethical considerations

The present study is a descriptive cross-sectional analysis carried out in a group of 54 school-age adolescents of both sexes. All participants live in Exaltación de la Cruz, a rural zone in the Province of Buenos Aires, Argentina. This community is part of a social integration program at the University of Buenos Aires. The adolescents' parents or legal guardians signed the corresponding informed consent, and were later interviewed in order to fill an anamnesis survey

and dietary questionnaire. Both the surveys and the informed consent followed the guidelines established in the Declaration of Helsinki (WMA, 2013). Furthermore, the study was approved by the Ethics Committee of the Faculty of Pharmacy and Biochemistry of the University of Buenos Aires (EXP-FYB N° 0087945/2016).

Adolescents with any severe disease or familiar cancer history were not part of this study. Additionally, adolescents exposed to environmental or therapeutic xenobiotics such as antibiotics and other medication, as well as those who had undergone radiation therapy or X-rays six months prior the sample collection, were excluded.

Food consumption frequency questionnaire (FFQ)

Taking into account the demographic characteristics of the population and their possible dietary habits, a questionnaire constructed by the authors and previously tested with people from the same region was used (Martínez-Perafán *et al.*, 2018). The survey collected information about the consumption per week of a list of foods from the traditional Argentine diet; these foods were later sorted into groups based on their nutritional characteristics. In this way, 13 food groups were formed: citrus, tropical fruits, fruit vegetables, leaf vegetables, roots and tubers, legumes, dairy products and eggs, red meat, white meat, lunch meat, cereals, fat and oil, and candies. Finally, a frequency-variety index was calculated for each group by multiplying the average frequency of consumption per week of all foods within the same group by the number of foods that had a frequency of consumption greater than or equal to once per week.

Collection of biological samples

Peripheral blood samples (n=43) were taken by venipuncture and collected in heparinised tubes, while buccal epithelium samples (n=54) were obtained by scraping the inside of each cheek with different cytological brushes, which were later put into a centrifuge tube containing Saccomannos' fixative (Biopur, Rosario, Argentina). All samples were taken in the morning, then they were stored in ice, protected from light and transported to the laboratory. Blood samples were usually processed two hours after taking the last sample, whereas those from buccal mucosa were stored at 4 °C until processing (no more than two weeks after sampling).

Cytokinesis-block micronucleus cytome (CBMN) assay

The CBMN assay in lymphocytes was performed following the indications proposed by Fenech (2007) with minimal modifications. According to this, lymphocytes were isolated from the peripheral blood samples using Ficoll-Paque (GE Healthcare Bio-Sciences, Uppsala, Sweden). Subsequently, 106 cells/well were cultivated in a 12-well culture plate with 1 mL/well of RPMI 1640 medium (Gibco®, Life Technologies Corporation, New York, US) supplemented with fetal bovine serum (15%; Internegocios, Buenos Aires, Argentina) and phytohemagglutinin (10 µg/mL; Gibco®, Life Technologies Corporation, New York, US). Cultures were incubated at 37 °C and after 44 h, cytochalasin B (4.5 µg/mL; Sigma-Aldrich, Steinheim, Germany) was added to each well in order to block cytokinesis. Later, 72 h after the establishment of the cultures, the lymphocytes were harvested, placed onto a slide and fixed with methanol (Merck, Darmstadt, Germany). Finally, the cells were stained with Giemsa (10%; Merck, Darmstadt, Germany) and 2000 BN cells were scored per individual at 1000x magnification, using a transmitted light microscope (Olympus CX31). The frequencies of the genotoxicity biomarkers MNi, NBUDs and NPBs were reported.

Buccal micronucleus cytome (BMCyt) assay

This test was carried out following the protocol of Thomas *et al.* (2009) with some adjustments. Firstly, buccal mucosa samples were washed three times with a buffer solution (0.1 M EDTA, Sigma-Aldrich, Steinheim, Germany; 0.01 M Tris–HCl, Sigma-Aldrich, Steinheim, Germany, and 0.02 M NaCl, Merck, Darmstadt, Germany). In order to achieve optimal cell disaggregation, 50 µL of DMSO (Merck, Darmstadt, Germany) per mL of cell suspension was added during the first two washes. After the third wash, the cells were transferred onto a slide, fixed with methanol (Merck, Darmstadt, Germany) and stained with propidium iodide (30 µg/mL; Sigma-Aldrich, Steinheim, Germany). Finally, 2000 cells per individual were scored by a fluorescence microscope at 400x magnification (Olympus BX-40F4). The frequencies of MNi and NBUDs were considered as genotoxicity biomarkers, whereas, CC, KHC, PYK and KYL were reported as cell death parameters, and BN as a result of failures in cytokinesis.

Statistical analysis

The data analysis consisted of two parts, first it was performed a univariate analysis of the demographic characteristics of the adolescents, their dietary information and the frequency of biomarkers in both lymphocytes and BMC. Secondly, a multivariate regression analysis between each cytogenetic biomarker and the main demographic characteristics of the adolescents was carried out. Similarly, a multivariate test was applied to find out the relationship between the biomarkers and the dietary habits of the study participants.

A negative binomial distribution was used for the multivariate analysis, this approach is especially effective with count data such as those obtained from the scoring of the cytome assay biomarkers (Ceppi *et al.*, 2010). Regarding the multivariate analysis with demographic characteristics, the previous step was to categorize the independent variables, avoiding categories with low frequencies when possible. The demographic characteristics taken into account were age, sex, and weight status (endogenous factors), and the exposure to SHS (exogenous variable).

On the other hand, prior to the multivariate analysis with the food groups, a bivariate analysis was carried out to select the variables that would fit in the model. Following this purpose, the values corresponding to the frequency-variety index of each food group were categorized into tertiles of the observed distribution, with the indices within the first tertile being the lowest, and those in the third tertile being the highest. In addition, the mean frequency of each biomarker in each of the three tertiles was calculated. Subsequently, the effect of consuming a certain food group was described as the percentage variation between the average frequency of each biomarker in the second and third tertiles compared to the mean of the first tertile. After the bivariate approach, the variables that presented a p-value less than 0.2 in relation to each biomarker were selected to be included in the multivariate regression model. Additionally, the variance inflation test was used as a second criterion for the exclusion of independent variables that presented multicollinearity.

Univariate analysis was carried out using IBM SPSS Statistics for Windows (IBM Corp, 2015), whereas negative binomial regression models, both bivariate and multivariate, were conducted in R (R Core Team, 2019) using the MASS (Venables and Ripley, 2002), jtools (Long, 2020) and car (Fox and Weisberg, 2019) packages.

Ethics approval and Informed consent

The present study was reviewed and approved by the Ethics Committee of the Faculty of Pharmacy and Biochemistry of the Universidad de Buenos Aires (EXP-FYB N° 0087945/2016), following national and international guidelines for research on humans.

This study was carried out with the written consent of the participants' parents or legal guardians.

RESULTS

Demographic characteristics and frequency of biomarkers

The main demographic characteristics and frequency of cytome assay biomarkers in peripheral blood lymphocytes (PBL) and BMC from the adolescents are listed in Table 1. The average age of the volunteers was 11.19 years (range from 10 to 14 years old). In addition, it was observed that more men than women participated in the study and that exposure to SHS was frequent. It should also be noted that there was no significant difference between the body mass index of male and female individuals (p=0.346), and most of the participants had a normal weight according to the criteria of the WHO (2009). Furthermore, considering the anamnesis data, it was also possible to verify that despite living in a rural area, the adolescents were not directly or indirectly exposed to agrochemicals or other types of environmental xenobiotics (data not shown), with the exception of cigarette smoke from parents who smoked.

In relation to the cytogenetic biomarkers, the CBMN assay revealed that the mean baseline frequencies of genotoxicity markers in the adolescent lymphocytes were low (in PBL MNi= 1.87‰, and in BMC MNi= 3.23‰ and NBUDs= 1.29‰). Similarly, the micronucleus cytome assay in BMC showed that the average of baseline values associated with genotoxic risk and cytokinesis failure were also close to zero, while the parameters related to cell death had a diverse mean frequency (Table 1).

Relationship between biomarkers in PBL and demographic characteristics

Table 2 shows the results of the multivariate analysis between the different biomarkers in PBL and the main demographic characteristics. There is a statistically significant relationship between exposure to SHS and NBUDs, with a significantly higher frequency in adolescents whose parents did not smoke (factor of 116%, p=0.028).

Relationship between biomarkers in BMC and demographic characteristics

A significant association (p=0.044) was observed between age and the frequency of NBUDs in BMC. Specifically, 11-year-old adolescents had a 51% decrease in this marker compared to 10-year-olds (Table 3). Similarly, Table 4 displays that the average number of PYK cells in the buccal mucosa was significantly lower in adolescents not exposed to SHS than in those exposed (percentage variation 57%, p=0.046).

Relationship between biomarkers in PBL and dietary habits

After the bivariate analysis between cytogenetic markers and different food groups (Table 5), a multivariate regression model was built to illustrate the relationship between biomarkers and dietary habits. Following this, Table 6 shows the results of the multivariate model between the genetic damage markers in lymphocytes and the short-listed food groups. Only a significant reduction (32%) in the frequency of MNi (p=0.032) was

Table 1. Mean value and standard error (SE) of demographic characteristics and frequencies of cytome assay biomarkers in peripheral blood lymphocytes (PBL) and buccal mucosa cells (BMC) in a group of Argentinean adolescents.

	n	Mean	SE
Age (years)	54	11.19	1.30
Sex			
Female	25	—	—
Male	29	—	—
BMI per sex*			
Female	25	20.98	1.15
Male	29	19.77	0.63
Weight status (♀:♂)			
Normal weight	35 (16:19)	—	—
Overweight	9 (4:5)	—	—
Obesity	10 (5:5)	—	—
Exposure to second-hand smoke			
Exposed	31	—	—
Unexposed	23	—	—
Biomarkers in PBL (‰)			
MNi	43	1.87	0.19
NBUDs	43	0.28	0.07
NPBs	43	0.15	0.04
Biomarkers in BMC (‰)			
MNi	54	3.23	0.77
NBUDs	54	1.29	0.25
BN	54	0.76	0.14
CC	54	45.44	4.53
KHC	54	15.07	1.30
PYK	54	0.61	0.14
KYL	54	19.89	2.49

*No significant difference between body mass index (BMI) averages of females and males (p=0.346).

MNi: micronuclei; NBUDs: nuclear buds; NPBs: nucleoplasmic bridges; BN: binucleated cells; CC: cells with condensed chromatin; KHC: karyorrhectic cells; PYK: pyknotic cells; KYL: karyolytic cells.

Table 2. Data obtained from a multivariate analysis between the different biomarkers in peripheral blood lymphocytes (PBL) and the main demographic characteristics

	MNi			NBUDs			NPBs		
	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*
Age (years)									
10	0	—	—	0	—	—	0	—	—
11	-25.06	0.35/1.12	117	-44.19	0.10/1.05	61	-31.03	0.14/4.53	811
12	4.51	0.47/1.73	747	-41.86	0.07/1.27	102	89.66	0.56/21.19	183
13 - 14[a]	-24.83	0.27/1.63	372	-22.09	0.07/2.27	292			
Sex									
Female	0	—	—	0	—	—	0	—	—
Male	0.27	0.60/1.39	680	34.04	0.60/3.80	388	-50.00	0.09/1.20	92
Weight status									
Normal weight	0	—	—	0	—	—	0	—	—
Overweight	-21.53	0.47/1.64	690	-18.03	0.31/5.00	761	60.00	0.45/4.61	537
Obesity[b]	-20.30	0.40/1.19	185	-27.87	0.20/2.17	490			
SH smoke									
Exposed	0	—	—	0	—	—	0	—	—
Unexposed	6.30	0.72/1.64	696	115.79	1.12/7.62	0.028*	-6.45	0.14/2.11	373

Percentage variation and *p*-values refer to comparison with the first category of each variable.

[a] The last category for NPBs was formed with adolescents from 12 to 14 years old

[b] The last category for NPBs was formed with adolescents with overweight and obesity

* Significant difference

MNi: micronuclei; NBUDs: nuclear buds; NPBs: nucleoplasmic bridges; CI: confidence interval

found for young people whose intake of roots and tubers was located in the second tertile compared to those who presented a consumption belonging to the first tertile.

Relationship between biomarkers in BMC and dietary habits

The association between the consumption of different types of food and markers of genotoxicity and cytokinesis failure is shown in Table 7. It can be observed that the highest level of intake of tropical fruits and red meat was related to a decrease in the frequency of MNi (83%, p=0.007 and 69%, p=0.006, respectively) when compared to the first tertile of consumption. On the contrary, a higher intake of legumes (tertile 2) was associated with an increase of 133% in this biomarker (p=0.009); however, this effect was not significant for the third tertile of consumption. Furthermore, the highest ingestion of citrus and red meat had an influence on the rise in the number of NBUDs (122%, p=0.022 and 175%, p=0.006, respectively). While the frequency of this parameter fell with the intake of cereals both in tertile 2 and 3 (39%, p=0.007 and 49%, p=0.0002), this effect was similar with the consumption of fat and oil but only up to the second tertile (95%, p=0.011). In addition, an increase in BN cells associated with the highest level of white meat consumption was observed (126%, p=0.004). In contrast to the genotoxicity results of the BMCyt assay, the cell death biomarkers of this test showed no association with any of the food groups included in the multivariate model (Table 8).

DISCUSSION

This study shows the baseline values of the cytome assay biomarkers in two different cell types from Argentine adolescents (10–14 years old) and their association with possible confounding variables. With regard to the frequency of genotoxicity parameters in lymphocytes, it was found that values of MNi, NBUDs and NPBs were comparable with those described by Gajski *et al.* (2013)

Table 3. Data obtained from a multivariate analysis between demographic characteristics and baseline frequency of genotoxicity and cytokinesis failure biomarkers in buccal mucosa cells (BMC)

	MNi			NBUDs			BN		
	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*
Age (years)									
10	0	—	—	0	—	—	0	—	—
11	-42.80	0.20/3.38	788	-51.22	0.09/0.97	0.044*	-35.00	0.20/1.83	371
12	45.20	0.35/9.99	466	-21.46	0.11/1.80	251	66.25	0.57/5.48	320
13 - 14	33.33	0.24/15.67	537	-57.32	0.04/1.62	147	56.25	0.48/7.08	377
Sex									
Female	0	—	—	0	—	—	0	—	—
Male	13.93	0.41/2.90	859	10.25	0.42/2.26	963	34.37	0.49/2.13	962
Weight status									
Normal weight	0	—	—	0	—	—	0	—	—
Overweight	-66.28	0.11/1.64	216	23.90	0.51/4.74	431	85.00	0.99/6.06	52
Obesity	20.12	0.31/4.00	879	-8.37	0.18/1.81	345	66.67	0.66/4.03	284
SH smoke									
Exposed	0	—	—	0	—	—	0	—	—
Unexposed	-11.57	0.24/1.82	426	8.87	0.64/3.59	350	-3.90	0.45/2.10	942

Percentage variation and *p*-values refer to comparison with the first category of each variable

* Significant difference

MNi: micronuclei; NBUDs: nuclear buds; BN: binucleated cells; CI: confidence interval

in their study on healthy Croatian children (4–14 years old). In addition, the average frequency of MNi found in the present study (1.87‰) was lower than that informed by Neri *et al.* (2005) for minors aged 10 to 14 years (6.02‰) in their pooled analysis. Furthermore, the levels of markers found in Argentine adolescents were similar to those reported for reference groups, or individuals not exposed, in other studies on child population (Kapka *et al.*, 2007; Milne *et al.*, 2015; Mørck *et al.*, 2016). Therefore, due to the similarities between the levels of the tested biomarkers and the results of other reports carried out on healthy children, it can be stated that the adolescents participating in the present study did not present genomic damage evaluated with the cytome assay in PBL.

Comparison of baseline levels of BMCyt assay parameters in the Argentine adolescents with those found in other studies revealed that the mean frequency of MNi and NBUDs were slightly higher in the Argentine minors than in healthy children selected as reference groups in biomonitoring studies from Italy (Villarini *et al.*, 2018; Panico *et al.*, 2020) Brazil (Sisenando *et al.*, 2012; Silva da Silva *et al.*, 2015), Mexico (Neri *et al.*, 2006; Gómez-Arroyo *et al.*, 2013) and Malaysia (Sopian *et al.*, 2020). Furthermore, the mean frequencies of the biomarkers of cell death and cytokinesis failure in the participants of the present study were similar to those of the reference group described by Villarini *et al.* (2018). In contrast, the mean frequency of PYK and KHC cells was higher in Brazilian children (Silva da Silva *et al.*, 2015), whereas the values of all markers of this type were much lower in the studies carried out on Mexican minors (Gómez-Arroyo *et al.*, 2013; Castañeda-Yslas *et al.*, 2016). It was also observed that some studies only mentioned the frequencies of parameters related to genotoxic risk (Sisenando *et al.*, 2012; Panico *et al.*, 2020; Sopian *et al.*, 2020). Despite the differences between the reports, it should be noted that the values of the biomarkers in this study were within the ranges proposed by Thomas *et al.* (2009) in healthy young people, except for the frequency of KYL cells, which was lower, as in most of the other publications.

Table 4. Data obtained from a multivariate analysis of demographic characteristics of the adolescents and the baseline frequency of cell death biomarkers in buccal mucosa cells (BMC)

	CC			KHC			PYK			KYL		
	Percentage variation	95% CI	p	Percentage variation	95% CI	p	Percentage variation	95% CI	p	Percentage variation	95% CI	p
Age (years)												
10	0	—	—	0	—	—	0	—	—	0	—	—
11	47.43	0.82/2.38	213	04.06	0.54/1.50	694	16.00	0.56/10.28	242	-6.72	0.35/1.63	466
12	98.90	0.89/3.16	107	-14.00	0.40/1.37	341	100.00	1.00/26.93	51	10.59	0.28/1.80	477
13 - 14	105.00	0.83/3.98	137	-37.07	0.27/1.30	188	-50.00	0.13/18.33	743	-5.64	0.18/1.87	368
Sex												
Female	0	—	—	0	—	—	0	—	—	0	—	—
Male	34.17	0.88/1.81	211	-7.63	0.69/1.39	889	-7.81	0.31/2.10	650	32.25	0.84/2.44	185
Weight status												
Normal weight	0	—	—	0	—	—	0	—	—	0	—	—
Overweight	1.27	0.64/1.68	889	4.38	0.62/1.58	963	-6.67	0.21/2.80	686	-30.03	0.39/1.62	522
Obesity	-30.50	0.50/1.32	399	-20.21	0.43/1.12	132	16.67	0.59/6.45	275	-40.11	0.30/1.24	169
SH smoke												
Exposed	0	—	—	0	—	—	0	—	—	0	—	—
Unexposed	10.22	0.68/1.44	956	-14.04	0.60/1.24	426	-56.79	0.12/0.98	0.046*	28.41	0.75/2.25	354

Percentage variation and *p*-values refer to comparison with the first category of each variable

* Significant difference

CC: cells with condensed chromatin; KHC: karyorrhectic cells; PYK: pyknotic cells; KYL: karyolytic cells; CI: confidence interval.

The baseline level of the biomarkers evaluated by the cytome assay in both lymphocytes and BMC can vary between people according to some sociodemographic or lifestyle factors. For example, among the adult population, it has been observed that older people and women have higher frequencies of MNi in lymphocytes (Fenech and Bonassi, 2011). However, the multivariate model fitted in the present study showed no association between the genotoxicity biomarkers of the lymphocyte cytome assay and age or sex. These results can be attributed to the fact that the influence of these variables is not the same in the child population, in which the relationship between sex and the frequency of MNi in blood cells appears to be non-existent. Meanwhile, the effect of age would only be noticeable among minors with a considerable difference in age (Holland *et al.*, 2011).

The weight status of the participants showed no effect on the measured biomarkers in lymphocytes, which contrasts the results from Scarpato *et al.* (2011), who found that the number of MNi in children with overweight and obesity was significantly higher than in those with normal weight; however, the mean frequency of MNi in the three groups was relatively low, as they mentioned (<3‰). Other studies on adults found higher levels of these biomarkers in people with obesity and metabolic syndrome (Donmez-Altuntas *et al.*, 2014; Karaman *et al.*, 2015). Nevertheless, prospective studies are necessary, especially on children and adolescents, to establish if the DNA damage increase could be related to overnutrition or sedentary lifestyle. In this way, it could be determined whether the cytome assay outcomes are markers or mediators of health problems such as obesity and metabolic syndrome (Andreassi *et al.*, 2011).

On the other hand, no exposure to SHS was associated with a higher frequency of NBUDs. This result was unexpected since tobacco smoke is considered a genotoxic agent (DeMarini, 2004). However, a pooled analysis by Bonassi *et al.* (2003) including more than 5,710 participants showed a decrease in the MNi value of smokers and former smokers compared to non-smokers, what indicates that there does not appear to be a strong positive effect of smoking on the markers of the cytome test in PBL. In relation to exposure to indoor environmental tobacco smoke in children, only one study was found that reported an increase in

Table 5. Data obtained from a bivariate analysis between cytogenetic markers and different food groups in peripheral blood lymphocytes (PBL) and buccal mucosa cells (BMC)

	PBL			BMC						
	MNi	**NBUDs**	**NPBs**	**MNi**	**NBUDs**	**BN**	**CC**	**KHC**	**PYK**	**KYL**
Citrus										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	948	928	645	258	65	82	916	379	506	975
Tertile 3	912	817	612	142	104	709	248	766	607	786
Tropical fruits										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	585	803	118	793	710	962	859	162	234	324
Tertile 3	536	525	372	0.001*	200	759	383	776	710	310
Fruit vegetables										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	409	912	775	837	99	449	420	105	50	614
Tertile 3	769	182	395	833	0.045*	482	277	742	277	972
Leaf vegetables										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	354	0.046*	301	0.046*	488	773	681	684	414	666
Tertile 3	402	70	1.000	499	255	336	886	391	169	151
Roots and Tubers										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	125	683	151	968	319	362	739	276	97	990
Tertile 3	867	281	591	251	672	884	373	612	681	878
Legumes										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	199	376	158	166	413	898	0.008*	215	947	508
Tertile 3	725	440	86	316	367	723	772	760	373	303
Dairy products and eggs										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	515	966	540	455	485	774	749	279	615	646
Tertile 3	138	320	306	583	804	636	314	203	663	605
Red meat[a]										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	0.034*	489	427	876	885	56	616	492	164	145
Tertile 3	568	222		0.023*	0.021*	904	612	183	801	257
White meat										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	369	348	288	849	678	338	426	566	727	736
Tertile 3	199	334	530	306	421	0.048*	830	817	190	266
Lunch meat										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	1.000	0.024*	803	582	674	250	245	982	503	728
Tertile 3	179	0.004*	773	833	267	453	605	457	420	631

Table 5 (continues). Data obtained from a bivariate analysis between cytogenetic markers and different food groups in peripheral blood lymphocytes (PBL) and buccal mucosa cells (BMC)

	PBL			BMC						
	MNi	NBUDs	NPBs	MNi	NBUDs	BN	CC	KHC	PYK	KYL
Cereals										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	553	216	657	904	354	88	686	619	25	64
Tertile 3	604	845	136	808	167	309	148	780	74	233
Fat and oil										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	0.020*	198	695	311	0.009*	557	64	528	332	743
Tertile 3	677	83	785	613	346	108	250	411	173	939
Candies										
Tertile 1	—	—	—	—	—	—	—	—	—	—
Tertile 2	834	169	771	142	515	487	456	696	885	908
Tertile 3	77	355	685	678	302	325	317	785	540	728

[a] Only two categories of consumption were created for PBL NPBs according to the median

*Significant difference

MNi: micronuclei; NBUDs: nuclear buds; NPBs: nucleoplasmic bridges; BN: binucleated cells; CC: cells with condensed chromatin; KHC: karyorrhectic cells; PYK: pyknotic cells; KYL: karyolytic cells.

the frequency of MNi in exposed minors (Baier *et al.*, 2002). These contradictory results could be explained partly by the influence of genetic polymorphisms that confer protection or susceptibility to genotoxins (Chandirasekar *et al.*, 2011; Fenech and Bonassi, 2011).

Due to the increase in the number of human biomonitoring studies that use the BMCyt assay, it is necessary to know the influence of external variables on the biomarkers measured by the test. A pooled analysis of the results of this method in 5,424 individuals, including around 725 children, concluded that sex does not influence the values of MNi, NBUDs, BN or PYK cells; however, they found a higher number of KHC cells in males (Bonassi *et al.*, 2011a). It should be noted that some studies in this review reported only the frequency of MNi as a biomarker. Additionally, it was mentioned that the values of cells with CC and KYL were so heterogeneous that they were not taken into account in the statistical analysis. These same authors revealed a positive association between age and the frequency of MNi, KHC and PYK. In contrast, in our study only one relationship was observed with NBUDs; however, the statistical significance of this association is so close to the threshold of the set level of significance ($p<0.05$) that it may wear off if the sample size is increased. Furthermore, it appears that changes in baseline levels of markers in epithelial cells, or at least of MNi, are only evident among individuals with a wide age difference (Hopf *et al.*, 2020).

In concordance with our study, no evidence on the influence of weight status on biomarkers in BMC was observed in children by Torres-Bugarin *et al.* (2009); however, the frequencies of the measured parameters were notably different from those obtained in our study. A similar result was reported by Espinosa Arreola *et al.* (2019), who included the nuclear anomalies - except MNi - in a single variable. Additionally, Villarini *et al.* (2018) showed no connection between body mass index (BMI) and this biomarker in oral epithelial cells of minors. On the contrary, Idolo *et al.* (2018) found that obesity had a positive influence on the occurrence of MNi in BMC from a group of children. These results indicate that although obesity is related to an increase in reactive oxygen species (ROS) which induce DNA damage (Usman and Volpi, 2018), it is not clear whether this can be evidenced by cytome assay markers, since most studies about this topic do not take into account the influence of potential confounding factors from diet (Setayesh *et al.*, 2018).

Regarding the exposure to SHS in children and the outcomes of the method, two studies found a positive association between the value of MNi and having a mother who smokes. However, these reports did not fit a multivariate analysis for the rest of the markers (Idolo *et al.*, 2018; Villarini *et al.*, 2018). Following this, Bonassi *et al.* (2011a) reported higher levels of MNi and PYK in heavy smokers (40 cigarettes per day), which could

Table 6. Data obtained from a multivariate model between the genetic damage markers in peripheral blood lymphocytes (PBL) and the consumption level of different food groups.

	MNi			NBUDs			NPBs		
	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*
Tropical fruits									
Tertile 1	N/A	N/A	N/A	N/A	N/A	N/A	0	—	—
Tertile 2							-72.00	0.04/1.23	87
Tertile 3							-42.00	0.05/1.15	75
Fruit vegetables									
Tertile 1	N/A	N/A	N/A	0	—	—	N/A	N/A	N/A
Tertile 2				7.50	0.31/3.35	986			
Tertile 3				112.50	0.60/5.19	297			
Leaf vegetables									
Tertile 1	N/A	N/A	N/A	0	—	—	N/A	N/A	N/A
Tertile 2				-78.72	0.09/2.14	306			
Tertile 3				-59.57	0.21/2.04	466			
Roots and Tubers									
Tertile 1	0	—	—	N/A	N/A	N/A	0	—	—
Tertile 2	-32.39	0.39/0.96	0.032*				-69.57	0.03/1.31	94
Tertile 3	-3.78	0.44/1.17	180				-28.26	0.06/1.34	114
Legumes									
Tertile 1	0	—	—	N/A	N/A	N/A	0	—	—
Tertile 2	-29.61	0.45/1.20	215				233.33		381
Tertile 3	-7.77	0.56/1.47	684				291.67		76
Dairy products and eggs									
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	19.00	0.65/1.77	786						
Tertile 3	46.00	0.67/1.93	625						
Red meata									
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	-48.59	0.38/1.31	273						
Tertile 3	12.60	0.82/1.96	279						
White meat									
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	-26.25	0.42/1.68	616						
Tertile 3	30.97	0.82/1.85	323						

Table 6 (continues). Data obtained from a multivariate model between the genetic damage markers in peripheral blood lymphocytes (PBL) and the consumption level of different food groups.

	MNi			NBUDs			NPBs		
	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*
Lunch meat									
Tertile 1	0	—	—	0	—	—	N/A	N/A	N/A
Tertile 2	0	0.50/1.31	378	-90.09	0.02/1.31	88			
Tertile 3	-27.01	0.48/1.22	265	-88.29	0.05/1.14	72			
Cereals									
Tertile 1	N/A	N/A	N/A	N/A	N/A	N/A	0	—	—
Tertile 2							53.33	0.12/6.77	933
Tertile 3							233.33	0.40/19.56	297
Fat and oil									
Tertile 1	0	—	—	0	—	—	N/A	N/A	N/A
Tertile 2	-61.26	0.22/1.16	105	-74.36	0.05/3.49	415			
Tertile 3	-8.23	0.65/1.59	934	-62.82	0.15/2.05	383			
Candies									
Tertile 1	0	—	—	0	—	—	N/A	N/A	N/A
Tertile 2	-5.86	0.70/2.32	435	-78.00	0.07/7.15	765			
Tertile 3	49.19	0.99/2.55	56	58.00	0.59/4.65	339			

Percentage variation and *p*-values refer to comparison with the first tertile of consumption.

N/A: Not applicable. The variable did not meet the necessary requirements to be included in the multivariate analysis

* Significant difference

MNi: micronuclei; NBUDs: nuclear buds; NPBs: nucleoplasmic bridges; CI confidence interval.

have an analogy with our results; however, it should be taken into consideration that Argentine adolescents were affected by passive smoking, therefore the type of exposure was different. On the other hand, the value of significance for the relationship with PYK in our case was close to the threshold of the set significance level. Interestingly, Nersesyan *et al.* (2011) found controversial results regarding the relationship between nicotine exposure in smokers and BMCyt assay biomarkers, indicating that, in order to determine the association between these variables, controlled intervention trials on smokers are necessary.

It is known that diet provides the necessary nutrients for DNA synthesis and repair, and some foods are even involved in the prevention of genomic damage. Consequently, the evaluation of dietary habits and food-related factors should be considered in biomonitoring studies that use biomarkers of genotoxicity (Fenech, 2020). However, some researchers do not take these parameters into account or do not include them in the statistical analysis when the focus of their research differs from a dietary topic. Such shortcomings should be addressed since it has already been observed that a high intake of vitamin E, retinol, folate, nicotinic acid and calcium decrease the frequency of MNi in adults, while a high consumption of riboflavin, pantothenic acid and biotin increases this value. Additionally, it is known that both the deficiency and excess of micronutrients can be harmful (Fenech *et al.*, 2005).

It is worth mentioning that in the consulted databases and search engines (PubMed, ScienceDirect and Google Scholar) no reports were found on the influence of the intake of macronutrients or specific food groups on the cytome assay markers in PBL from children. Some studies related to micronutrients were evidenced, for instance, Milne *et al.* (2015) found a positive association

Table 7. Data obtained from the analysis of the association between the consumption level of different food groups and markers of genotoxicity and cytokinesis failure in buccal mucosa cells (BMC)

	MNi			NBUDs			BN		
	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*
Citrus									
Tertile 1	0	—	—	0	—	—	0	—	—
Tertile 2	-48.45	0.33/2.70	915	157.14	0.81/4.02	151	111.67	0.88/4.11	100
Tertile 3	-55.88	0.37/2.82	960	122.14	1.14/5.70	0.022*	-16.67	0.23/1.52	276
Tropical fruits									
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	-13.03	0.33/2.48	845						
Tertile 3	-83.35	0.08/0.67	0.007*						
Fruit vegetables									
Tertile 1	N/A	N/A	N/A	0	—	—	N/A	N/A	N/A
Tertile 2				131.82	0.80/4.79	139			
Tertile 3				169.70	0.65/4.84	267			
Leaf vegetables									
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	-69.84	0.19/1.66	295						
Tertile 3	-30.59	0.31/2.33	761						
Legumes									
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	132.94	1.44/12.61	0.009*						
Tertile 3	72.51	0.66/4.78	255						
Red meat									
Tertile 1	0	—	—	0	—	—	0	—	—
Tertile 2	-9.38	0.13/1.33	139	-7.89	0.51/3.65	530	-77.78	0.06/1.30	105
Tertile 3	-68.98	0.10/0.68	0.006*	175.00	1.42/7.96	0.006*	-4.44	0.45/1.78	752
White meat									
Tertile 1	N/A	N/A	N/A	N/A	N/A	N/A	0	—	—
Tertile 2							80.43	0.78/7.31	126
Tertile 3							126.09	1.46/6.98	0.004*
Cereals									
Tertile 1	N/A	N/A	N/A	0	—	—	0	—	—
Tertile 2				-38.61	0.13/0.72	0.007*	-70.00	0.20/1.56	265
Tertile 3				-48.89	0.08/0.47	0.0002*	-35.24	0.31/2.01	622
Fat and oil									
Tertile 1	N/A	N/A	N/A	0	—	—	0	—	—
Tertile 2				-94.94	0.01/0.54	0.011*	-30.21	0.45/4.67	538
Tertile 3				-32.74	0.52/2.08	913	-51.04	0.20/1.32	168
Candies									
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	-60.71	0.12/1.55	200						
Tertile 3	24.71	0.45/2.87	782						

Percentage variation and *p*-values refer to comparison with the first tertile of consumption.
N/A: Not applicable. The variable did not meet the necessary requirements to be included in the multivariate analysis.
* Significant difference

MNi: micronuclei; NBUDs: nuclear buds; BN: binucleated cells; CI: confidence interval

Table 8. Data obtained from the analysis of the association between the consumption level of different food groups and the baseline frequency of cell death biomarkers in buccal mucosa cells (BMC)

	CC			KHC			PYK			KYL		
	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*	Percentage variation	95% CI	*p*
Tropical fruits												
Tertile 1	N/A	N/A	N/A	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2				35.09	0.91/2.09	130						
Tertile 3				6.25	0.75/1.67	576						
Fruit vegetables												
Tertile 1	N/A	N/A	N/A	0	—	—	0	—	—	N/A	N/A	N/A
Tertile 2				40.72	0.74/1.76	539	212.50	0.55/6.49	314			
Tertile 3				-6.58	0.46/1.20	226	90.62	0.48/5.79	427			
Leaf vegetables												
Tertile 1	N/A	N/A	N/A	N/A	N/A	N/A	0	—	—	0	—	—
Tertile 2							-42.00	0.18/2.19	460	15.33	0.76/2.85	249
Tertile 3							100.00	0.74/5.37	173	-35.56	0.43/1.56	548
Roots and Tubers												
Tertile 1	N/A	N/A	N/A	N/A	N/A	N/A	0	—	—	N/A	N/A	N/A
Tertile 2							156.41	0.55/4.81	384			
Tertile 3							28.21	0.22/2.34	585			
Legumes												
Tertile 1	0	—	—	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Tertile 2	-45.82	0.40/1.05	80									
Tertile 3	06.02	0.77/1.87	413									
Red meat												
Tertile 1	N/A	N/A	N/A	0	—	—	0	—	—	0	—	—
Tertile 2				18.46	0.65/1.67	852	129.17	0.73/5.50	178	70.75	0.88/3.67	111
Tertile 3				26.69	0.95/2.28	84	14.58	0.30/2.37	750	39.43	0.79/2.52	244
White meat												
Tertile 1	N/A	N/A	N/A	N/A	N/A	N/A	0	—	—	N/A	N/A	N/A
Tertile 2							-28.26	0.07/2.41	331			
Tertile 3							89.13	0.59/3.49	430			
Cereals												
Tertile 1	0	—	—	N/A	N/A	N/A	0	—	—	0	—	—
Tertile 2	-9.05	0.58/1.51	778				300.00	0.44/6.27	448	-45.93	0.25/1.01	53
Tertile 3	-26.86	0.45/1.13	151				196.00	0.49/7.08	356	-30.36	0.42/1.58	544
Fat and oil												
Tertile 1	0	—	—	N/A	N/A	N/A	0	—	—	N/A	N/A	N/A
Tertile 2	-42.92	0.38/1.20	182				-66.00	0.03/2.49	256			
Tertile 3	-20.36	0.62/1.35	639				90.00	0.51/2.92	651			

Percentage variation and *p*-values refer to comparison with the first tertile of consumption.

N/A: Not applicable. The variable did not meet the necessary requirements to be included in the multivariate analysis.

CC: cells with condensed chromatin; KHC: karyorrhectic cells; PYK: pyknotic cells; KYL: karyolytic cells; CI: confidence interval.

between plasma calcium and MNi, while NPBs showed a positive relationship with lutein and a negative relationship with α-tocopherol. Additionally, Prá *et al.* (2011) observed a negative correlation between iron intake and MNi and NPB values. In our approach, the negative association between a moderate consumption of tubers and roots and the frequency of MNi could be due to the phytoconstituents of these foods, including phenolic compounds, which have known antioxidant activity (Chandrasekara and Josheph Kumar, 2016).

The influence of diet on the BMCyt biomarkers was also evaluated in the pooled analysis by Bonassi *et al.* (2011a), who reported that people who consumed fruits daily had a lower frequency of MNi compared to those who did not. This finding is consistent with our results, i.e., the high intake of tropical fruits had a negative effect on this parameter, possibly due to their antioxidant properties (Lim *et al.*, 2007). On the other hand, research on children from Italy showed that the intake of red or processed meat more than four times a week was positively associated with the occurrence of MNi (Idolo *et al.*, 2018), while adherence to the Mediterranean diet would have a negative effect on this outcome (Panico *et al.*, 2020). In contrast, in the group of Argentine adolescents, a high consumption of red meat had a negative influence on MNi; however, this dietary habit increased the number of NBUDs by up to 175%, which would imply another pathway of expression of DNA damage and corroborate a possible genetic risk caused by excessive intake of this food (Wolk, 2017).

On the other hand, our findings showed a negative association between the intake of cereals and the basal level of NBUDs, which could be associated with the phenolic acids and flavonoids contained in these products (Žilić *et al.*, 2011). It was also observed that the consumption of legumes and citrus fruits had a positive influence on genotoxicity markers. This might seem confusing since both types of foods also contain molecules that act as scavengers for free radicals, preventing their binding with the DNA (Amarowicz and Pegg, 2008; Mahmoud *et al.*, 2019). However, the following three remarks should be considered here. 1) Food contains different types of compounds, not just those that act as antioxidants. Citrus pectin, for instance, is related to an increase in ROS that leads to apoptosis in cancer cells (Salehi *et al.*, 2018). 2) The metabolism of each individual is different due to genetic polymorphisms, which intervenes in the use of nutrients from the diet and its relationship with various diseases (Loktionov, 2003). 3) The amount of macro and micronutrients that enter the body depends on the quality and type of food consumed, for example, lactic fermentation of fruits and vegetables changes the functional properties of these foods (Septembre-Malaterre *et al.*, 2018).

In addition to the genotoxicity biomarkers, BMCyt evaluates other parameters, including the frequency of BN cells. This type of nuclear abnormality could indicate failures in cytokinesis as a result of chromosomal nondisjunction (Shi and King, 2005); however, its biological significance is still unknown, so the association of this marker with the ingestion of white meat is disconcerting. Similarly, there is a lack of information regarding the cell death process associated with CC, KHC, KYL and PYK. These types of cells are found in the superficial layers of the oral epithelium and migrate towards the oral cavity in a process of turnover that culminates in the exfoliation of old cells (Thomas *et al.*, 2009). This tissue composition could be affected by the development of oral dysplasia, a phenomenon associated in part with diet (Morse *et al.*, 2000). Nevertheless, Sánchez-Siles *et al.* (2014) found that while some cell death markers were not altered in patients with oral leukoplakia and dysplasia, the parameters of genotoxicity and cytokinesis failure were significantly increased. This would indicate that biomarkers related to cell death are less prone to the influence of exogenous factors, possibly due to the dynamics of the tissue, which is constantly replaced to protect itself against injuries and maintain its structure (Squier and Kremer, 2001).

CONCLUSIONS

The results of this study provide information on the baseline frequencies of the cytome assay biomarkers in two different cell types from a group of Argentine adolescents. This information is relevant not only to establish the cytogenetic status of the study group, but also because it serves as a starting point for the creation of a database about the health status of children in Argentina, with a focus in populations for which currently little information is available, such as those located in rural areas.

Additionally, the multivariate model helped to assess the association between the biomarkers of both versions of the cytome assay and the main demographic characteristics and dietary habits of the study participants. It was determined that exogenous variables (exposure to SHS and diet) had an influence on markers, particularly those of genotoxicity, both in PBL and BMC. In contrast, the consumption of the evaluated food groups had no effect on parameters related to cell death in epithelial cells. These findings help to better understand the influence of potential confounding factors on the results of the cytome assay and confirm that these variables should be included not only in a descriptive analysis, but also in a multivariate one, especially in child biomonitoring studies.

Finally, it should be noted that the findings of this study are subject to one limitation, that is, the sample

size, which could be seen as small in order to generalize the results to a larger population. However, it is worth mentioning that many of the associations found between the biomarkers and the potential confounding variables were considerably below the set significance value. In other words, it is possible that the size of these effects is large enough to be seen in a small sample. Further research focused on the influence of potential confounders on the cytome assay biomarkers in larger samples of children from different areas is recommended.

BIBLIOGRAPHY

Amarowicz R., Pegg R.B. (2008) Legumes as a source of natural antioxidants. Eur. J. Lipid Sci. Technol. 110(10): 865-78.

Andreassi M.G., Barale R., Iozzo P., Picano E. (2011) The association of micronucleus frequency with obesity, diabetes and cardiovascular disease. Mutagenesis. 26(1): 77-83.

Baier G., Stopper H., Kopp C., Winkler U., Zwirner-Baier I. (2002) Erkrankungen der oberen Atemwege und Gentoxizität bei tabakrauchexponierten Kindern [Respiratory diseases and genotoxicity in tobacco smoke exposed children]. Laryngorhinootologie. 81(3): 217-25.

Beedanagari S., Vulimiri S.V., Bhatia S., Mahadevan B. (2014). Genotoxicity biomarkers: Molecular basis of genetic variability and susceptibility. In: Gupta R.C. (Ed.) Biomarkers in toxicology. Academic Press, Massachusetts, pp. 729-742.

Bonassi S., Coskun E., Ceppi M., Lando C., Bolognesi C., Burgaz S., Holland N., Kirsh-Volders M., Knasmueller S., Zeiger E., Carnesoltas D., Cavallo D., da Silva J., de Andrade V.M., Demircigil G.C., Domínguez Odio A., Donmez-Altuntas H., Gattas G., Giri A., Giri S., Gómez-Meda B., Gómez-Arroyo S., Hadjidekova V., Haveric A., Kamboj M., Kurteshi K., Martino-Roth M.G., Montero Montoya R., Nersesyan A., Pastor-Benito S., Favero Salvadori D.M., Shaposhnikova A., Stopper H., Thomas P., Torres-Bugarín O., Yadav A.S., Zúñiga González G., Fenech M. (2011a) The HUman MicroNucleus project on eXfoLiated buccal cells (HUMN(XL)): the role of life-style, host factors, occupational exposures, health status, and assay protocol. Mutat. Res. 728(3): 88-97.

Bonassi S., El-Zein R., Bolognesi C., Fenech M. (2011b) Micronuclei frequency in peripheral blood lymphocytes and cancer risk: evidence from human studies. Mutagenesis. 26(1): 93-100.

Bonassi S., Neri M., Lando C., Ceppi M., Lin Y.P., Chang W.P., Holland N., Kirsch-Volders M., Zeiger E., Fenech M.; HUMN collaborative group (2003) Effect of smoking habit on the frequency of micronuclei in human lymphocytes: results from the Human MicroNucleus project. Mutat. Res. 543(2): 155-66.

Castañeda-Yslas I.J., Arellano-García M.E., García-Zarate M.A., Ruíz-Ruíz B., Zavala-Cerna M.G., Torres-Bugarín O. (2016) Biomonitoring with Micronuclei Test in Buccal Cells of Female Farmers and Children Exposed to Pesticides of Maneadero Agricultural Valley, Baja California, Mexico. J. Toxicol. 2016: 7934257.

Ceppi M., Biasotti B., Fenech M., Bonassi S. (2010) Human population studies with the exfoliated buccal micronucleus assay: statistical and epidemiological issues. Mutat. Res. 705(1): 11-9.

Chandirasekar R., Suresh K., Jayakumar R., Venkatesan R., Lakshman Kumar B., Sasikala K. (2011) XRCC1 gene variants and possible links with chromosome aberrations and micronucleus in active and passive smokers. Environ. Toxicol. Pharmacol. 32(2): 185-92.

Chandrasekara A., Josheph Kumar T. (2016) Roots and Tuber Crops as Functional Foods: A Review on Phytochemical Constituents and Their Potential Health Benefits. Int. J. Food Sci. 2016: 3631647.

DeMarini D.M. (2004) Genotoxicity of tobacco smoke and tobacco smoke condensate: a review. Mutat. Res. 567(2-3): 447-74.

Donmez-Altuntas H., Sahin F., Bayram F., Bitgen N., Mert M., Guclu K., Hamurcu Z., Arıbas S., Gundogan K., Diri H. (2014) Evaluation of chromosomal damage, cytostasis, cytotoxicity, oxidative DNA damage and their association with body-mass index in obese subjects. Mutat. Res. Genet. Toxicol. Environ. Mutagen. 771: 30-6.

Espinosa Arreola M., Ortega Martínez L.D., Pérez Armendáriz B., Marqués Maldonado A.D.P., Baños Lara M.D.R. (2019) Evaluación del daño genético y de los hábitos alimentarios en niños con normopeso y obesidad en edad escolar [Evaluation of genetic damage and eating habits in children with normal weight and obesity in school age]. Nutr. Hosp. 36(2): 309-314.

Fenech M. (2006) Cytokinesis-block micronucleus assay evolves into a "cytome" assay of chromosomal instability, mitotic dysfunction and cell death. Mutat. Res. 600(1-2): 58-66.

Fenech M. (2007) Cytokinesis-block micronucleus cytome assay. Nat. Protoc. 2(5): 1084-104.

Fenech M. (2020) The Role of Nutrition in DNA Replication, DNA Damage Prevention and DNA Repair. In: Caterina R.D.E., Martinez J.A., Kohlmeier M. (Eds.) Principles of Nutrigenetics and Nutrigenomics. Academic Press, Massachusetts, pp. 27-32.

Fenech M., Baghurst P., Luderer W., Turner J., Record S., Ceppi M., Bonassi S. (2005) Low intake of calcium, folate, nicotinic acid, vitamin E, retinol, beta-carotene and high intake of pantothenic acid, biotin and riboflavin are significantly associated with increased genome instability--results from a dietary intake and micronucleus index survey in South Australia. Carcinogenesis. 26(5): 991-9.

Fenech M., Bonassi S. (2011) The effect of age, gender, diet and lifestyle on DNA damage measured using micronucleus frequency in human peripheral blood lymphocytes. Mutagenesis. 26(1): 43-9.

Fenech M., Knasmueller S., Bolognesi C., Bonassi S., Holland N., Migliore L., Palitti F., Natarajan A.T., Kirsch-Volders M. (2016) Molecular mechanisms by which in vivo exposure to exogenous chemical genotoxic agents can lead to micronucleus formation in lymphocytes in vivo and ex vivo in humans. Mutat. Res. Rev. Mutat. Res. 770(Pt A): 12-25.

Fox J., Weisberg S. (2019) https://socialsciences.mcmaster.ca/jfox/Books/Companion/ (accessed December 2020).

Gajski G., Gerić M., Oreščanin V., Garaj-Vrhovac V. (2013) Cytogenetic status of healthy children assessed with the alkaline comet assay and the cytokinesis-block micronucleus cytome assay. Mutat. Res. 750(1-2): 55-62.

Gómez-Arroyo S., Martínez-Valenzuela C., Calvo-González S., Villalobos-Pietrini R., Waliszewski S.M., Calderón-Segura M.E., Martínez-Arroyo A., Félix-Gastélum R., Lagardeascarrega A. (2013). Assessing the genotoxic risk for mexican children who are in residential proximity to agricultural areas with intense aerial pesticide applications. Rev. Int. Contam. Ambient. 29(3): 217-225.

Holland N., Bolognesi C., Kirsch-Volders M., Bonassi S., Zeiger E., Knasmueller S., Fenech M. (2008) The micronucleus assay in human buccal cells as a tool for biomonitoring DNA damage: the HUMN project perspective on current status and knowledge gaps. Mutat. Res. 659(1-2): 93-108.

Holland N., Fucic A., Merlo D.F., Sram R., Kirsch-Volders M. (2011) Micronuclei in neonates and children: effects of environmental,

genetic, demographic and disease variables. Mutagenesis. 26(1): 51-6.

Hopf N.B., Danuser B., Bolognesi C., Wild P. (2020) Age related micronuclei frequency ranges in buccal and nasal cells in a healthy population. Environ. Res. 180: 108824.

IBM Corp (2015) https://www.ibm.com/products/spss-statistics (accessed December 2020).

Idolo A., Grassi T., Bagordo F., Panico A., De Giorgi M., Serio F., Guido M., Piscitelli P., De Filippis G., Raho A., De Donno A. (2018) Micronuclei in Exfoliated Buccal Cells of Children Living in a Cluster Area of Salento (Southern Italy) with a High Incidence of Lung Cancer: The IMP.AIR Study. Int. J. Environ. Res. Public Health. 15(8): 1659.

Kapka L., Baumgartner A., Siwińska E., Knudsen L.E., Anderson D., Mielzyńska D. (2007) Environmental lead exposure increases micronuclei in children. Mutagenesis. 22(3): 201-7.

Karaman A., Aydın H., Geçkinli B., Çetinkaya A., Karaman S. (2015) DNA damage is increased in lymphocytes of patients with metabolic syndrome. Mutat. Res. Genet. Toxicol. Environ. Mutagen. 782: 30-5.

Kirsch-Volders M., Fenech M., Bolognesi C. (2018) Validity of the Lymphocyte Cytokinesis-Block Micronucleus Assay (L-CBMN) as biomarker for human exposure to chemicals with different modes of action: A synthesis of systematic reviews. Mutat. Res. Genet. Toxicol. Environ. Mutagen. 836(Pt A): 47-52.

Landrigan P.J., Kimmel C.A., Correa A., Eskenazi B. (2004) Children's health and the environment: public health issues and challenges for risk assessment. Environ. Health Perspect. 112(2): 257-65.

Lemos A.T., Lemos C.T., Coronas M.V., Rocha J.R.D., Vargas V.M.F. (2020) Integrated study of genotoxicity biomarkers in schoolchildren and inhalable particles in areas under petrochemical influence. Environ. Res. 188: 109443.

Lim Y. Y., Lim T. T., Tee J. J. (2007) Antioxidant properties of several tropical fruits: A comparative study. Food Chem. 103(3): 1003-1008.

Loktionov A. (2003) Common gene polymorphisms and nutrition: emerging links with pathogenesis of multifactorial chronic diseases (review). J. Nutr. Biochem. 14(8): 426-51.

Long J.A. (2020) https://cran.r-project.org/package=jtools (accessed December 2020).

Louro H., Heinälä M., Bessems J., Buekers J., Vermeire T., Woutersen M., van Engelen J., Borges T., Rousselle C., Ougier E., Alvito P., Martins C., Assunção R., Silva M.J., Pronk A., Schaddelee-Scholten B., Del Carmen Gonzalez M., de Alba M., Castaño A., Viegas S., Humar-Juric T., Kononenko L., Lampen A., Vinggaard A.M., Schoeters G., Kolossa-Gehring M., Santonen T. (2019) Human biomonitoring in health risk assessment in Europe: Current practices and recommendations for the future. Int. J. Hyg. Environ. Health. 222(5): 727-737.

Mahmoud A.M., Hernández Bautista R.J., Sandhu M.A., Hussein O.E. (2019) Beneficial Effects of Citrus Flavonoids on Cardiovascular and Metabolic Health. Oxid. Med. Cell. Longev. 2019: 5484138.

Martínez-Perafán F., Santoli M.F., López-Nigro M., Carballo M.A. (2018) Assessment of the health status and risk of genotoxic and cytotoxic damage in Argentinian adolescents living near horticultural crops. Environ. Sci. Pollut. Res. Int. 25(6): 5950-5959.

Milne E., Greenop K.R., Ramankutty P., Miller M., de Klerk N.H., Armstrong B.K., Almond T., O'Callaghan N.J., Fenech M. (2015) Blood micronutrients and DNA damage in children. Mol. Nutr. Food Res. 59(10): 2057-65.

Mørck T.A., Loock K.V., Poulsen M.B., Siersma V.D., Nielsen J.K., Hertel O., Kirsch-Volders M., Knudsen L.E. (2016) Micronucleus frequency in Danish schoolchildren and their mothers from the DEMOCOPHES population. Mutagenesis. 31(1): 1-8.

Morse D.E., Pendrys D.G., Katz R.V., Holford T.R., Krutchkoff D.J., Eisenberg E., Kosis D.L., Kerpel S., Freedman P., Mayne S.T. (2000) Food group intake and the risk of oral epithelial dysplasia in a United States population. Cancer Causes Control. 11(8): 713-20.

Neri M., Ceppi M., Knudsen L.E., Merlo D.F., Barale R., Puntoni R., Bonassi S. (2005) Baseline micronuclei frequency in children: estimates from meta- and pooled analyses. Environ. Health Perspect. 113(9): 1226-9.

Neri M., Ugolini D., Bonassi S., Fucic A., Holland N., Knudsen L.E., Srám R.J., Ceppi M., Bocchini V., Merlo D.F. (2006) Children's exposure to environmental pollutants and biomarkers of genetic damage. II. Results of a comprehensive literature search and meta-analysis. Mutat. Res. 612(1): 14-39.

Nersesyan A., Fenech M., Bolognesi C., Mišík M., Setayesh T., Wultsch G., Bonassi S., Thomas P., Knasmüller S. (2016) Use of the lymphocyte cytokinesis-block micronucleus assay in occupational biomonitoring of genome damage caused by in vivo exposure to chemical genotoxins: Past, present and future. Mutat. Res. Rev. Mutat. Res. 770(Pt A): 1-11.

Nersesyan A., Muradyan R., Kundi M., Knasmueller S. (2011) Impact of smoking on the frequencies of micronuclei and other nuclear abnormalities in exfoliated oral cells: a comparative study with different cigarette types. Mutagenesis. 26(2): 295-301.

Panico A., Grassi T., Bagordo F., Idolo A., Serio F., Tumolo M.R., De Giorgi M., Guido M., Tutino M., De Donno A. (2020) Micronucleus Frequency in Exfoliated Buccal Cells of Children Living in an Industrialized Area of Apulia (Italy). Int. J. Environ. Res. Public Health. 17(4): 1208.

Perlroth N.H., Castelo Branco C.W. (2017) Current knowledge of environmental exposure in children during the sensitive developmental periods. J. Pediatr (Rio J). 93(1): 17-27

Prá D., Bortoluzzi A., Müller L.L., Hermes L., Horta J.A., Maluf S.W., Henriques J.A., Fenech M., Franke S.I. (2011) Iron intake, red cell indicators of iron status, and DNA damage in young subjects. Nutrition. 27(3): 293-7.

R Core Team (2019) https://www.R-project.org/ (accessed December 2020).

Salehi F., Behboudi H., Kavoosi G., Ardestani S.K. (2018) Oxidative DNA damage induced by ROS-modulating agents with the ability to target DNA: A comparison of the biological characteristics of citrus pectin and apple pectin. Sci. Rep. 8(1): 13902.

Sánchez-Siles M., Camacho-Alonso F., Ros-Llor I., López-Jornet P. (2014) Cytogenetic biomonitoring in oral leukoplakia patients with mild dysplasia. Int. J. Dermatol. 53(12): 1454-9.

Scarpato R., Verola C., Fabiani B., Bianchi V., Saggese G., Federico G. (2011) Nuclear damage in peripheral lymphocytes of obese and overweight Italian children as evaluated by the gamma-H2AX focus assay and micronucleus test. FASEB J. 25(2): 685-93.

Septembre-Malaterre A., Remize F., Poucheret P. (2018) Fruits and vegetables, as a source of nutritional compounds and phytochemicals: Changes in bioactive compounds during lactic fermentation. Food Res. Int. 104: 86-99.

Setayesh T., Nersesyan A., Mišík M., Ferk F., Langie S., Andrade V.M., Haslberger A., Knasmüller S. (2018) Impact of obesity and overweight on DNA stability: Few facts and many hypotheses. Mutat. Res. Rev. Mutat. Res. 777: 64-91.

Shi Q., King R.W. (2005) Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines. Nature. 437(7061): 1038-42.

Silva da Silva C., Rossato J.M., Vaz Rocha J.A., Vargas V.M. (2015) Characterization of an area of reference for inhalable particulate matter (PM2.5) associated with genetic biomonitoring in children. Mutat. Res. Genet. Toxicol. Environ. Mutagen. 778: 44-55.

Sisenando H.A., Batistuzzo de Medeiros S.R., Artaxo P., Saldiva P.H., Hacon S.deS. (2012) Micronucleus frequency in children exposed to biomass burning in the Brazilian Legal Amazon region: a control case study. BMC Oral Health. 12: 6.

Sly P.D., Neira M., Collman G., Carpenter D.O., Landrigan P.J., Van Den Berg M., Barriga F.D., Ruchirawat M., Laborde A., Pascale A., Heacock M., Dalmau M.T., Suk W.A. (2014) Networking to advance progress in children's environmental health. Lancet Glob. Health. 2(3): e129-30.

Sommer S., Buraczewska I., Kruszewski M. (2020) Micronucleus Assay: The State of Art, and Future Directions. Int. J. Mol. Sci. 21(4): 1534.

Sopian N.A., Jalaludin J., Mayusi T.Z.A.T., Latif M.T. (2020) Increased Chromosomal Damage among Children in Proximity to Industrial Zone. Aerosol Air Qual. Res. 20: 944–955.

Squier C.A., Kremer M.J. (2001) Biology of oral mucosa and esophagus. J. Natl. Cancer Inst. Monogr. (29): 7-15.

Thomas P., Holland N., Bolognesi C., Kirsch-Volders M., Bonassi S., Zeiger E., Knasmueller S., Fenech M. (2009) Buccal micronucleus cytome assay. Nat. Protoc. 4(6): 825-37.

Torres-Bugarín O., Fernández-García A., Torres-Mendoza B.M., Zavala-Aguirre J.L., Nava-Zavala A., Zamora-Perez A.L. (2009) Genetic profile of overweight and obese school-age children. Toxicol. Environ. Chem. 91(4): 789-795.

Usman M., Volpi E.V. (2018) DNA damage in obesity: Initiator, promoter and predictor of cancer. Mutat. Res. Rev. Mutat. Res. 778: 23-37.

Venables W.N., Ripley B.D. (2002) http://www.stats.ox.ac.uk/pub/MASS4 (accessed December 2020).

Villarini M., Levorato S., Salvatori T., Ceretti E., Bonetta S., Carducci A., Grassi T., Vannini S., Donato F., Bonetta S., Verani M., de Donno A., Bonizzoni S., Bonetti A., Moretti M., Gelatti U.; MAPEC LIFE Study Group (2018) Buccal micronucleus cytome assay in primary school children: A descriptive analysis of the MAPEC_LIFE multicenter cohort study. Int. J. Hyg. Environ. Health. 221(6): 883-892.

Wolk A. (2017) Potential health hazards of eating red meat. J. Intern. Med. 281(2): 106-122.

WHO (2009) https://www.who.int/growthref/tools/en/ (accessed December 2020).

WHO (2011) Summary of principles for evaluating health risks in children associated with exposure to chemicals. Geneva, Switzerland. 56 p.

WHO (2015) https://www.who.int/europe/publications/i/item/WHO-EURO-2015-3209-42967-60040 (accessed December 2020)

WMA (2013) https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ (accessed February 2020).

Žilić S., Hadži-Tašković Šukalović V., Dodig D., Maksimović V., Maksimović M., Basić Z. (2011) Antioxidant activity of small grain cereals caused by phenolics and lipid soluble antioxidants. J. Cereal Sci. 54(3): 417-24.

ACKNOWLEDGEMENTS

The authors especially thank the volunteers and their parents for their participation in this study. The authors are also grateful to the entire CIGETOX team for their contribution in different stages of this research.

FUNDING STATEMENT

This research was supported by the Universidad de Buenos Aires (UBANEX Grant, 2019 – 2020) and the Fundación Alberto J. Roemmers (Grant for novel researchers, 2018).

AUTHOR CONTRIBUTION STATEMENT

Martínez Perafán F.: conceptualization, formal analysis, funding acquisition, investigation, methodology, visualization, writing - original draft, writing - review and editing.

Casciaro S.: conceptualization, methodology, writing - review and editing.

López Nigro M.M.: conceptualization, funding acquisition, methodology, project administration, resources, supervision, validation, writing - review and editing.

Carballo M.A.: conceptualization, funding acquisition, methodology, project administration, resources, supervision, validation, writing - review and editing.

BAG | Journal of Basic & Applied Genetics

Vol XXXV (1): 39-51; Jan.-July 2024

Journal of the Argentine Society of Genetics

RANDOM FOREST IN PLANT GENETICS AND BREEDING: AN APPLICATION IN TOMATO AS A MODEL CROP

RANDOM FOREST EN GENÉTICA Y MEJORAMIENTO GENÉTICO DE PLANTAS: UNA APLICACIÓN EN TOMATE COMO CULTIVO MODELO

Faviere G.[1], Vitelleschi M.S.[2], Pratta G.R.[3]

[1] Instituto de Investigaciones Teóricas y Aplicadas en Estadística (IITAE), Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario, Rosario, Santa Fe, Argentina.

[2] IITAE, Consejo de Investigaciones de la Universidad Nacional de Rosario (CIUNR), Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario, Rosario, Santa Fe, Argentina..

[3] Instituto de Investigaciones en Ciencias Agrarias de Rosario (IICAR), Facultad de Ciencias Agrarias, Universidad Nacional de Rosario-CONICET, Zavalla, Santa Fe, Argentina

Corresponding author:
Guillermo R. Pratta
gpratta@unr.edu.ar

ORCID 0000-0002-3682-0946

Cite this article as:
Faviere G., Vitelleschi M.S., Pratta G.R. 2024. *RANDOM FOREST* IN PLANT GENETICS AND BREEDING: AN APPLICATION IN TOMATO AS A MODEL CROP. BAG. Journal of Basic and Applied Genetics XXXV (1): 39-51.

Received: 07/13/2023

Revised version received: 03/13/2024

Accepted: 03/13/2024

General Editor: Elsa Camadro

DOI: 10.35407/bag.2024.35.01.03

ISSN online version: 1852-6233

Available online at www.sag.org.ar/jbag

ABSTRACT

Random Forest approaches have been used in phenotyping at both morphological and metabolic levels and in genomics studies, but direct applications in practical situations of plant genetics and breeding are scarce. Random Forest was compared with Discriminant Analysis for its ability in classifying tomato individuals belonging to different breeding populations, exclusively based on phenotypic fruit quality traits. In order to take into account different steps in breeding programs, two populations were assayed. One was composed by a set of RILs derived from an interspecific tomato cross, and the other was composed by two of these RILs and the corresponding F_1, F_2 and backcross generations. Being tomato an autogamous species, the first population was considered a final step in breeding programs because promising genotypes are being evaluated for putative commercial release as new cultivars. Meanwhile, the second one, in which new variation is being generated, was considered as an initial step. Both Random Forest and Discriminant Analysis were able to classify populations with the aim of evaluating general variability and identifying the traits that most contribute to this variability. However, overall errors in classification were lower for Random Forest. When comparing the adequacy of classification between populations, errors of both statistical analyses were greater in the second population than in the first one, though Random Forest was more precise than Discriminant Analysis even in this initial step of plant breeding programs. Random Forest allowed breeders to get a reliable classification of tomato individuals belonging to different breeding populations.

Key words: discriminant analysis, Machine Learning, parametric and non-parametric classification techniques, phenotype identification, traits categorization

RESUMEN

Los enfoques de Random Forest se han utilizado en la fenotipificación, tanto a nivel morfológico como metabólico, y en estudios de genómica, pero las aplicaciones directas en situaciones prácticas de fitomejoramiento y genética son escasas. Random Forest se comparó con el Análisis Discriminante por su capacidad en la clasificación de individuos de tomate pertenecientes a diferentes poblaciones de mejoramiento, exclusivamente en función de los rasgos fenotípicos de calidad de la fruta. Para tener en cuenta los diferentes pasos en los programas de mejoramiento, se ensayaron dos poblaciones. Una estaba compuesta por un conjunto de RILs derivadas de un cruce interespecífico de tomate, y la otra estaba compuesta por dos de estas RILs y las correspondientes generaciones F_1, F_2 y retrocruzas. Siendo el tomate una especie autógama, la primera población se consideró un paso final en los programas de mejoramiento porque se están evaluando genotipos prometedores para su lanzamiento comercial putativo como nuevos cultivares. Mientras tanto, la segunda, en la que se está generando nueva variación, se consideró como un paso inicial. Tanto Random Forest como Análisis Discriminante pudieron clasificar poblaciones con el objetivo de evaluar la variabilidad general e identificar los rasgos que más contribuyen a esta variabilidad. Sin embargo, los errores generales en la clasificación fueron menores para Random Forest. Al comparar la adecuación de la clasificación entre poblaciones, los errores de ambos análisis estadísticos fueron mayores en la segunda población que en la primera, aunque Random Forest fue más preciso que el Análisis Discriminante incluso en este paso inicial de los programas de fitomejoramiento. Random Forest permitió a los criadores obtener una clasificación fiable de individuos de tomate pertenecientes a diferentes poblaciones de cría.

Palabras clave: análisis discriminante, Aprendizaje Automático, técnicas de clasificación paramétricas y no paramétricas, identificación de fenotipos, categorización de rasgos.

INTRODUCTION

In plant breeding, different populations are voluntarily created by crossing selected genotypes to obtain hybrids. Then, the genetic structure of these artificial populations may be predicted, which allows estimating the components of genetic mean values and variances underlying the traits to be improved (Kearsey and Pooni, 1996). Hence, various groups (families, generations) of objects (individuals) are available for evaluation, enabling the application of supervised classification to assess the generated genetic variability (Stephan *et al.*, 2015).

One of the first challenges that biological sciences must deal with is classification (Duda *et al.*, 2000), an inherent process in most human activities that consists in accurately and efficiently assigning a class or a type to a given object under study (Trainor *et al.*, 2017). Objects are considered as factors that are evaluated by a series of variables or attributes with the goal of constructing groups according to their similarities (Hastie *et al.*, 2008). Two principal approaches are distinguished in this common challenge: supervised and unsupervised classification. In the first one, a priori known groups of objects are assessed aiming to establish objective criteria through data analysis for predicting with low uncertainty the belonging of new objects to any of those groups (Alhusain and Hafez, 2017). In unsupervised classification, instead, the belonging of studied objects to a given group is unknown and the goal is to find the underlying structure of data according to similarities found during the assessment (Larose and Larose, 2015).

Tomato (*Solanum lycopersicum* L.) is one of the most important horticultural crops worldwide (FAOSTAT, 2017). Also, it is a model species for plant genetics and breeding by means of both conventional and advanced strategies (Gerszberg *et al.*, 2015). Phenotypic evaluation is essential at different steps of a breeding program, especially when variability for quantitative agronomic traits is increased by crosses to wild germplasm (Dempewolf *et al.*, 2017).

Discriminant Analysis is a parametric method widely used for classifying in biological and agronomic applications (Alhusain and Hafez, 2017) while the non-parametric classification techniques, such as Random Forest, becomes necessary in many studies (Singh *et al.*, 2016).

The objective of this research was to assess the use of Random Forest to classify populations with different genetic structure according to phenotypic variability for fruit quality traits in two different usual situations of plant breeding. The accuracy and robustness of Random Forest in identifying the desired genotypes and the proportional contribution of measured traits in defining such genotypes were compared with the results obtained by Discriminant Analysis.

MATERIALS AND METHODS

Plant populations and traits under study

Two populations were evaluated with the aim of considering two different plant breeding activities. Both of them represent genomic recombination among the same parental genotypes (cv. Caimanta of *S. lycopersicum* and LA0722 of *S. pimpinellifolium*) in extreme situation of linkage disequilibrium, genotypic composition and inbreeding level. The first population comprised eight of the 18 RILs obtained by Rodriguez *et al.* (2006), hereafter named as L1, L5, L6, L8, L9, L15, L17, and L18 (total N=396 plants, because some individuals were lost during the transplant. The final number of individuals per RIL is given in Table 1). These 8 RILs were selected for adequately representing the total variability. In this population, linkage disequilibrium is low (<0.01), inbreeding level is high (>0.99) and genotypes are homozygous, representing potential new tomato commercial cultivars derived after several cycles of artificial selection over both early and advanced generations of selfing from a cross between cultivated and exotic germplasms. Data analyzed in this research are the mean values over six years of agronomic evaluation because its genetic structure is stabilized and this population was considered as a final step in a breeding program.

The second population comprised two selected RILs (L1 and L18), their F_1 (L18 x L1) and its segregating generations F_2 (L18 x L1), obtained by selfing, and both backcrosses F_1 (L18 x L1) x L18 and F_1 (L18 x L1) x L1, hereafter named as F_1, F_2, BC_1 and BC_2, respectively (N=218 plants). In this population of basic breeding generations (Kearsey and Pooni, 1996), linkage disequilibrium is high (>0.20), inbreeding level is relatively low (coefficient F=0.5). Genotypes are both homozygous and heterozygous, representing early generations from a cross among elite genotypes. Gene segregation occurring from the meiosis in F_1 gives new opportunity of recombining and selecting over the genotypes resulting from the previous breeding actions that allowed deriving the parental RILs. In fact, L1 and L18 are registered in the Argentinean National Registry of Cultivars). Data analyzed in this research were measured just in one year of agronomic evaluation because the genetic structure of the segregating generations (F_2 and both BCs) varies with each cycle of selfing, and this population was considered as an initial step in a breeding program.

Both sets of population were assayed under greenhouse conditions at the experimental field station "José F. Villarino", Universidad Nacional de Rosario, Argentina (Latitude: 33.02° S, Longitude: 60.88° O, Altitude: 50 masl) according to a completely randomized design with six replications. Following Mahuad *et al.*

(2013), 11 quantitative traits were evaluated, five of them in fruits harvested at breaker stage (when carotenoids accumulation becomes visible) and the other six in fruits harvested at red ripe stage (with approximately 90% of red surface). In 10 fruits per plant at breaker stage, weight (W, in g), diameter (D, in cm), height (H, in cm), shape index (SI, ratio between H and D), and shelf life (SL, number of days from harvest until the fruit stored at 25 ± 3 ºC loses commercial value due to, for instance, excessive softening), were measured. In fruits at red ripe stage, the following traits were evaluated: soluble solids content (SS, in Brix degrees, as the percentage of sucrose in the fruit juice), pH and titratable acidity (TA, in g of citric acid per 100 g of homogenate) of the fruit juice, firmness (F, measured on two opposite equatorial sides with a digital firmness type Shore A tester Durofel, DFT 100, with a 0.10 cm^2 cap), ratio a/b or chroma index (parameter related to color tone, being "a" the absorbance at 540 nm wavelengths and "b" the absorbance at 675 nm wavelengths), and L value or reflectance percentage (L, parameter related to color intensity, presenting values that range from +100 for white to 0 for black). Values "a", "b" and L were determined with a Chroma Meter CR 400. The color parameters and F were determined in five intact fruits per plant, whereas the SS and the pH were measured in the juice obtained by homogenizing a variable number of three to eight fruits per plant, which depended on the fruit size. In the first set of populations, the mean locule number per fruit (LN) was also measured in five fruits per plant.

Statistical Analyses

Random Forest is a non-parametric classification technique of Machine Learning proposed by Breiman (2001). It is a classifier that generates a big number of decision trees, and each tree is grown from a bootstrap sample of the response variable. The best split is selected from a random subset of variables at each node of the tree, and then the tree grows to the maximum extent without pruning. Each individual is classified by each tree and the most common outcome is used as the final classification. For this classification, contribution of each variable to form the groups is assessed (Breiman, 2001). In plant genetics and breeding, the quantification of this contribution could be considered as a description of phenotypes according to the importance of traits and it could objectively assist in the identification of phenotypes in their belonging to a given breeding population. Random Forest applies a built-in cross-validation, which in this research consisted of a first training step with 2/3 of the data and a validation step with the remaining 1/3, according to Breiman (2001), to estimate set error via the use of Out-Of-Bag (OOB) samples. Firstly, based on the training algorithm, data that did not take part at a given iteration in the bootstrap sample (the so called OOB data) are predicted using the tree grown with the bootstrap sample to be assigned at a given group. This process is known as validation step, and finally, once each individual has been assigned to a given group according to the OOB predictions, the error rate, known as the OOB estimate of error rate, is calculated. In other words, low values of the OOB estimate of the error rate indicate a high precision in the classification (Hastie *et al.*, 2008). Furthermore, two measures on the importance of variables are given by Random Forest (Hastie *et al.*, 2008): the Mean Decrease Accuracy (MDA), which is obtained from the OOB error estimation, and the Mean Decrease Gini (MDG), based in the Gini Index. Hyperparameter tuning of RF greatly influences its accuracy but there is a large scientific discussion on how to accomplish it (Bernard *et al.*, 2009). In this research, the number of tree (NT), the percentage of errors in global classification (GC) associated to NT and the OOB were taken as metrics for tuning hyperparametrization (Probst *et al.*, 2019).

Machine Learning techniques have not been yet widely used in plant genetics and breeding with a similar goal to that of this research. Consequently, Discriminant Analysis, the most common multivariate technique (Lapins and Nash, 1957, Lynch *et al.*, 1987, Sivakumar *et al.*, 2017, Abu-Ellail *et al.*, 2020), was used as a control for group assignment and for measuring the importance of variable contribution detected by Random Forest. Every trait was included in all cases, i.e., no selection of variables was accomplished neither by Random Forest nor by Discriminant Analysis. Random Forest was performed by the statistical package RStudio, version 1.0 ("randomForest" library, randomForest function), while SAS software, version 9.0, was used for Discriminant Analysis, through the procedure "proc discrim". Significance of all statistical tests was assigned at a limit p-value of 0.05.

RESULTS

Population of RILs

Mean values and standard deviations for evaluated traits across six years of agronomic evaluation are presented in Figure 1. A wide phenotypic variability among RILs was found for all traits, though, as expected, the absolute range of variation depended on the scale of each variable. In respect to Random Forest application, three variables were randomly selected for determining each node in the iterative construction of each tree, this sub-conjunct of traits was used for choosing the best partition. The percentages of errors in global classification and in

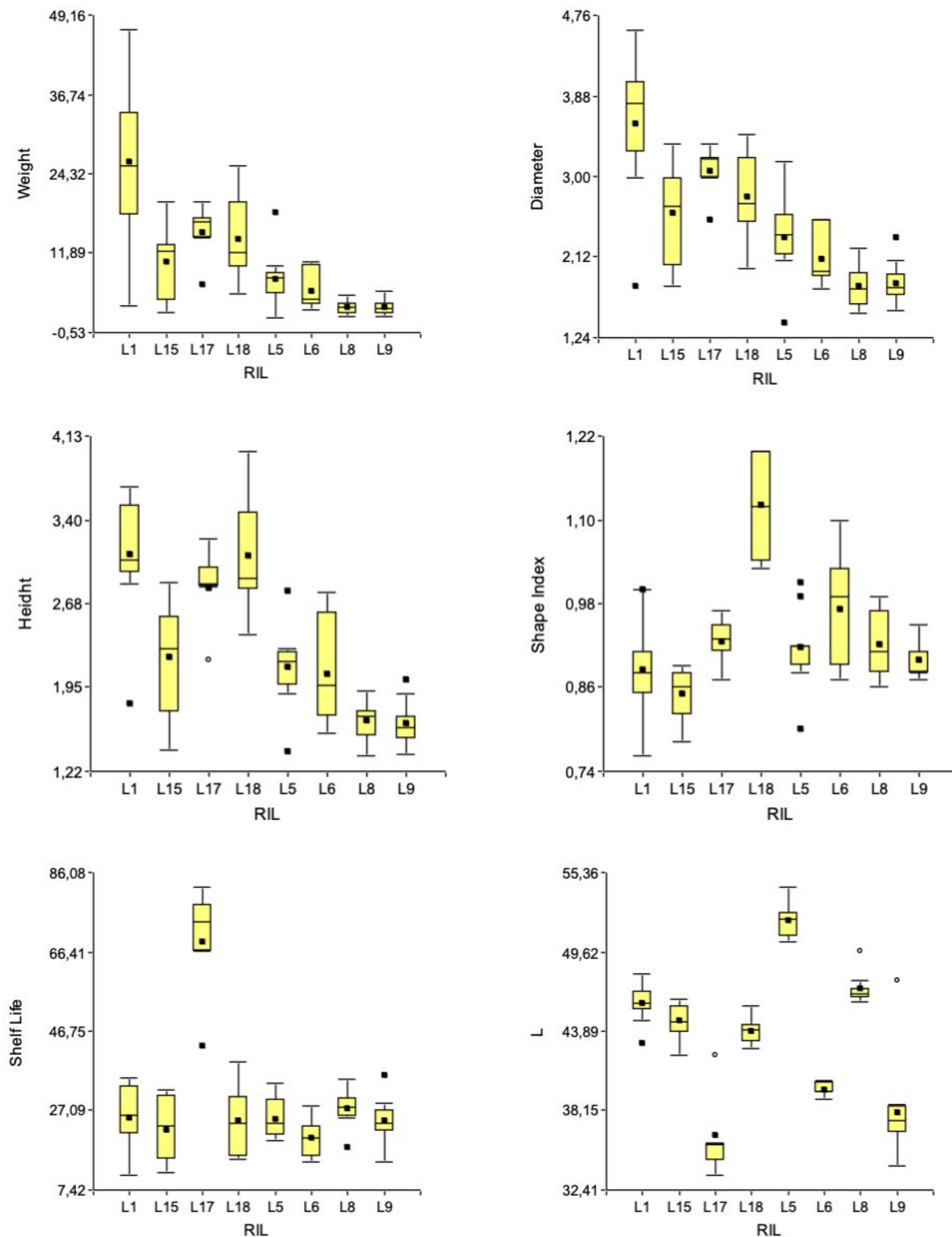


Figure 1 (continues). Values of the fruit traits weight (W, in g), diameter (D, in cm), height (H, in cm), shape index (SI, ratio H/D), shelf life (SL, in days), reflectance percentage (L, in %), chroma index (ratio a/b, being "a" the absorbance at 540 nm wavelengths and "b" the absorbance at 675 nm wavelengths), locule number (LN), soluble solids content (SS, in °Brix), pH, titratable acidity (TA, in g of citric acid per 100 g of homogenate juice), and firmness (F, in %) in eight tomato RILs obtained by Rodriguez *et al.* (2006).

the classification per group (RIL) are shown in Table 1. From 200 constructed trees onwards, both classification errors were stabilized in a null value, hence there are none misclassifications after this number. As the most frequent number of trees constructions in the literature is at least 500, the 100% of plants were perfectly classified into the expected group, the OOB error being 0%. Therefore, as 0% OOB error has been obtained with just 200 constructed trees (Table 1), this RIL population was accurately classified by Random Forest. The most important traits to obtain this excellent classification were L, TA, pH, SS and F, according to their MDA and MDG values (Table 2).

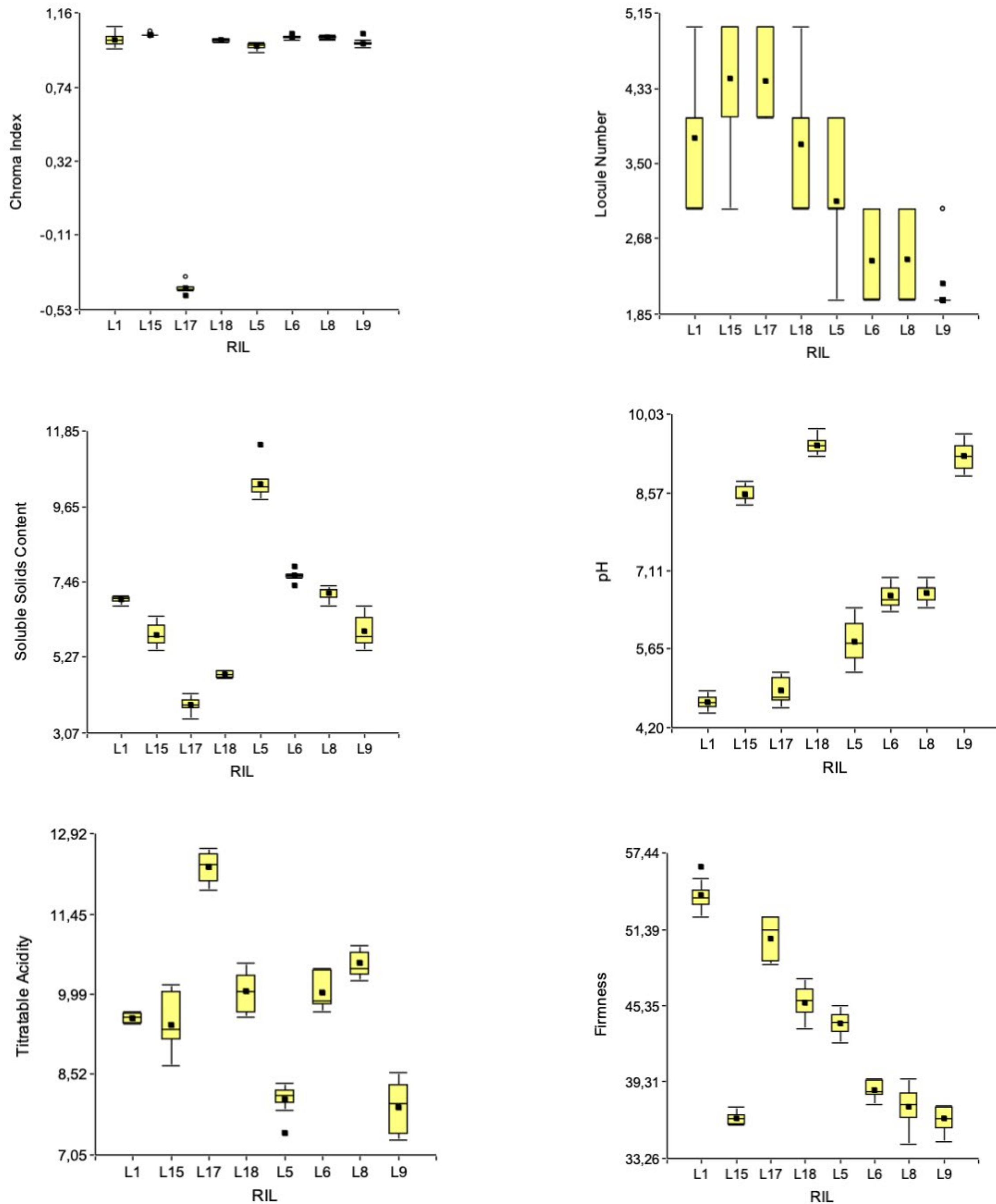


Figure 1 (continuation). Values of the fruit traits weight (W, in g), diameter (D, in cm), height (H, in cm), shape index (SI, ratio H/D), shelf life (SL, in days), reflectance percentage (L, in %), chroma index (ratio a/b, being "a" the absorbance at 540 nm wavelengths and "b" the absorbance at 675 nm wavelengths), locule number (LN), soluble solids content (SS, in °Brix), pH, titratable acidity (TA, in g of citric acid per 100 g of homogenate juice), and firmness (F, in %) in eight tomato RILs obtained by Rodriguez *et al.* (2006).

With the control technique, Discriminant Analysis, six linear discriminant functions (LDF) were obtained that allowed classifying RILs and measuring the contribution of each trait to the total variability (Table 3, LDF 1 and LDF 2 are shown as footnotes). However, when it was contrasted with Random Forest, errors in misclassifying were greater in Discriminant Analysis (Table 3). This misclassification made by Discriminant Analysis could be explained by two reasons. Firstly, Rao's F test ($p<0.0001$, Table 3) indicated that there were at least two groups of RILs which have different average vectors or LDF, i.e., the eight RILs are not univocally different among them but they could be clustered in either two, three, four, five or six groups, all these groups being

Table 1. Percentage of errors in global classification (GC) and per RIL (indicated with L and the respective number, these genotypes are eight tomato RILs obtained by Rodriguez *et al.*, 2006) considering different number of trees (NT) in each applied Random Forest

NT	GC	L1 (54)*	L15 (42)	L17 (30)	L18 (42)	L5 (66)	L6 (42)	L8 (54)	L9 (66)
50	3.03	0	14.29	0	0	0	14.29	0	0
100	1.52	0	14.29	0	0	0	0	0	0
150	1.52	0	14.29	0	0	0	0	0	0
200	0	0	0	0	0	0	0	0	0
300	0	0	0	0	0	0	0	0	0
400	0	0	0	0	0	0	0	0	0
500	0	0	0	0	0	0	0	0	0

* Numbers in parenthesis indicate the final number of plants per RIL evaluated in the research

Table 2. Contribution of each fruit trait to the classification by Random Forest of eight tomato RILs obtained by Rodriguez *et al.* (2006), according to Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG). Fruit traits: weight (W, in g), diameter (D, in cm), height (H, in cm), shape index (SI, ratio H/D), shelf life (SL, in days), reflectance percentage (L, in %), chroma index (ratio a/b, being "a" the absorbance at 540 nm wavelengths and "b" the absorbance at 675 nm wavelengths), locule number (LN), soluble solids content (SS, in °Brix), pH, titratable acidity (TA, in g of citric acid per 100 g of homogenate juice), and firmness (F, in %).

Fruit Trait	MDA	MDG
D	5.02	1.74
H	6.27	2.37
SI	6.37	3.01
W	6.69	2.53
SL	5.31	2.01
L	15.11	8.34
a/b	7.12	2.92
LN	4.39	1.14
SS	13.53	8.64
pH	13.86	8.31
F	13.25	7.20
TA	14.92	8.18

statistically significant. In fact, in Table 3 it is shown sequentially that all the obtained LDF were significant, i.e. classification RILs according to these six LDF is not robust. Secondly, traits which were identified as the most important in their contribution to general variability, varied in respect to those identified by Random Forest. For instance, a/b was identified by Discriminant Analysis as having the most important contribution to LDF 1 (footnote to Table3) while this trait had low contribution to general variability in the analysis with Random Forest (Table 1). Hence a lower robustness of classification is accomplished with Discriminant Analysis compared to Random Forest. In fact, differences in identifying traits contribution to RILs classification between Random Forest and Discriminant Analysis resulted in a wrong predicted assignation to RIL 15 of six plants actually belonging to RIL 9. This misclassification caused a global apparent error of 1.51% in Discriminant Analysis, while the global apparent error was null in Random Forest.

Population of Basic Generations

Mean values and standard deviations for the six basic generations evaluated are presented in Figure 2. Though some difference due to environmental influences were detected on the mean values of parental lines between both databases, general tendencies for morphological traits were observed since L1 had flattened fruits with higher weight and size than L18, whose fruits were elongated. Also, the F_1 phenotype agreed to gene actions reported by Pereira da Costa *et al.* (2014). For instance, the lower weight of F_1 fruits was explained by negative dominance of exotic alleles early contributed by LA0722 that are segregating in dispersion among L1 and L18 (Cabodevila *et al.*, 2021). In agreement, individuals from backcross to L1 (BC_2) had slightly heavier fruits than those of backcross to L18 (BC_1). Regarding variances, and as expected, the F_2 generation had a larger dispersion than both backcrosses for all traits; the F_1, genetically uniform, and the parents were the least variable generations.

The different genetic structure among generations in this set of populations, in contrast to the previous set in which all populations were homozygous, provoked some not unexpected effects on applying classification

Table 3. Significance of the Sequential Test (ST) for the Linear Discriminant Functions (LDF) in the Discriminant Analysis applied to eight tomato RILs obtained by Rodriguez *et al.* (2006)

Steps of ST	Null Hypothesis	Canonical Correlation (CC)	Square CC	F-value	*p*-value
1	None LDF is significant	0.99	0.98	103.01	<0.0001
2	Only LDF 1 is significant	0.99	0.98	60.71	<0.0001
3	Only LDF 1 and 2 are significant	0.98	0,96	41.61	<0.0001
4	LDF 1, 2, and 3 are significant	0.96	0,92	27.33	<0.0001
5	LDF 1, 2, 3, and 4 are significant	0.94	0.88	20.34	<0.0001
6	LDF 1, 2, 3, 4, and 5 are significant	0.83	0.66	12.00	<0.0001
7	LDF 1, 2, 3, 4, 5, and 6 are significant	0.72	0.52	9.46	<0.0001

The first two LDF, with standardized coefficients and traits, were:
LDF 1 = - 0.41 D + 0,89 H – 0.32 SI – 0.57 W + 0,17 SL – 0.03 L – 1.01 a/b + 0.09 LN – 0.34 SS – 0.30 pH + 0.02 F + 0.05 TA
LDF 2 = - 1.04 D + 0.84 H – 0.37 SI – 0.03 W + 0,003 SL – 0.04 L + 0.15 a/b + 0.13 LN + 0.37 SS – 0.80 pH + 0.40 F + 0.11 TA

techniques. Firstly, when including the F_2 generation either in Random Forest or in Discriminant Analysis, plants from all other generations were misclassified as belonging to the F_2 (data not shown). Even though differences in population size, common to plant breeding process, could explain this observation, the actual cause of this misclassification is the segregation of genes observed in the F_2, BC_1, and BC_2 generations while both parents and the F_1 are uniformly homozygote and heterozygote, respectively. Then the higher level of segregation and recombination among the F_2 in comparison to both BC_1 and BC_2 explains that even for not genetically uniform generation, all plant were classified as belonging to the F_2. Hence, data from F_2 generation were not taken into account for applying both classification techniques.

Similar to the previous study in RIL population, randomly selecting based on three variables was chosen for dividing the nodes in the iteratively construction of the trees during training the algorithm. However, Table 4 shows that, even considering 500 trees, errors in global classification and in classification per groups are not eliminated. Given that in 200 trees errors are stabilized in minimum values, this number is retained for continuing the analysis. However, the robustness of classification in this population is lower than in the RILs population previously analyzed according to this data respecting the number of trees construction. In fact, of the 120 evaluated plants, only 78 (65%), all belonging to the backcrosses generations (Table 5), could be adequately assigned to their respective group. Of these two backcrosses, BC_1 plants were better assigned (91% classified as BC_1 and 9% as BC_2) than BC_2 plants (76.6 % classified as BC_2, 17% as BC_1, and 6.4% as L1). All parental and F_1 plants were misclassified as either BC_1 or BC_2. The importance of the variables to classification algorithm explain these observations, given that SI (particularly higher in L18 and its BC_1) had the greatest MDA and MDG, i.e., the greatest contribution to construct the decision trees (Table 6). Other important traits were SS, D, and H, though their MDA and MDG values were low compared to SI.

In respect to the control technique Discriminant Analysis, the existence of at least two generations with different mean vectors or LDF was contrasted with Rao's F test ($p<0.0001$). Though four LDF were obtained in the analysis, the latter 2 were no significant in this population of basic generations. In fact, just the two first LDF, whose composition is shown as a footnote in Table 7, had strictly statistical significance. Traits mostly contributing to LDF 1 were D (1.97), H (-1.14), and a/b (-0.72), while again D (-1.04) and H (0.84) together with pH (-0.80) were the most important in LDF 2. Interestingly, SI was not detected as a highly important trait by Discriminant Analysis, possibly due to heteroscedasticity. Accordingly, Discriminant Analysis had lower ability than Random Forest for classifying plants, and only 67 over 120 plants (56%) were correctly assigned to their respective group (Table 8). However, though both techniques got a best classification for both backcrosses (72.7% in BC_1 and 53.9% in BC_2), Discriminant Analysis, in opposition to Random Forest could adequately assign some F_1 and L1 plants. Despite this observation, its estimated apparent error is very high (43.48%). Though many tomato plants from these basic breeding generations were misclassified by both

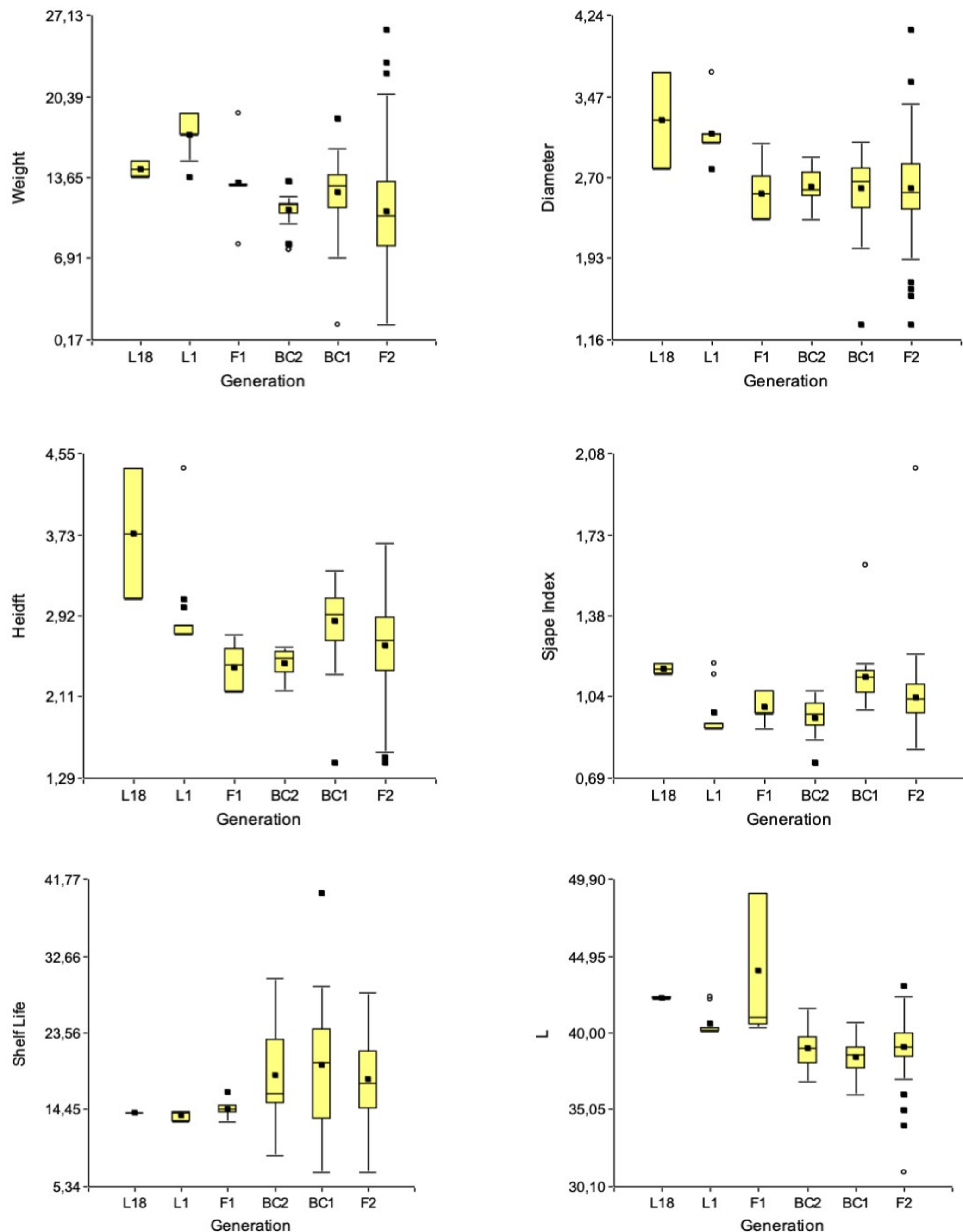


Figure 2 (continues). Values of the fruit traits weight (W, in g), diameter (D, in cm), height (H, in cm), shape index (SI, ratio H/D), shelf life (SL, in days), reflectance percentage (L, in %), chroma index (ratio a/b, being "a" the absorbance at 540 nm wavelengths and "b" the absorbance at 675 nm wavelengths), soluble solids content (SS, in °Brix), pH, titratable acidity (TA, in g of citric acid per 100 g of homogenate juice), and firmness (F, in %) in the population composed by two parental tomato RILs (L18 and L1; obtained by Rodriguez *et al.* 2006), their F_1 (second cycle hybrid L18 x L1), and the segregating generations F_2 (obtained by selfing the F_1), BC_1 (F1 x L18) and BC_2 (F1 x L1).

techniques, it is noteworthy that just plant 7 from RIL 18 was assigned to a different group, concretely to BC_2 by Random Forest and to BC_1 by Discriminant Analysis. In all other cases, misclassification on individuals was to the same erroneous group (data not shown).

DISCUSSION

Random Forest and other Learning Machine approaches have been used in phenotyping at both morphological and metabolic level (Amit and Geman, 1997; Singh *et*

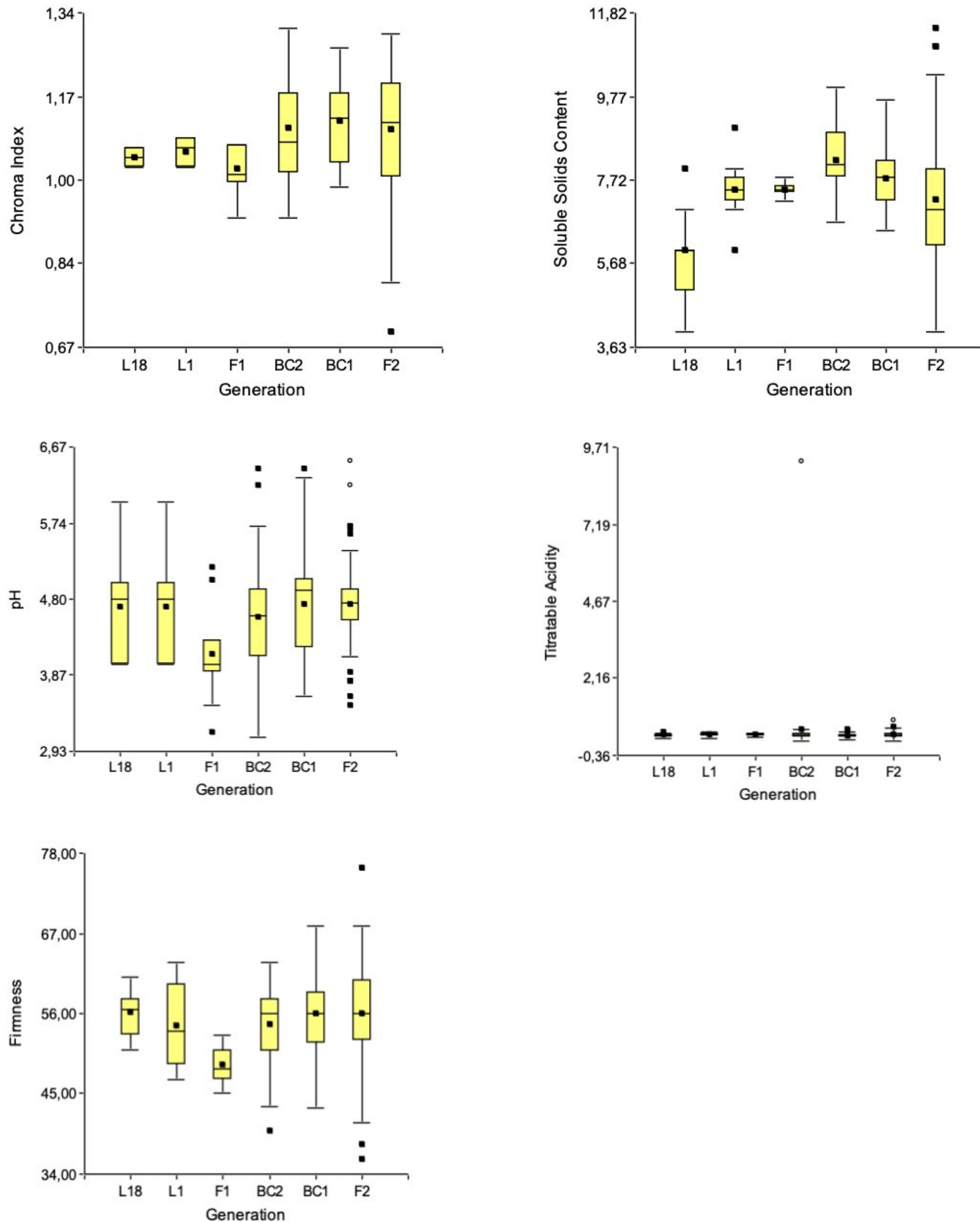


Figure 2 (continuation). Values of the fruit traits weight (W, in g), diameter (D, in cm), height (H, in cm), shape index (SI, ratio H/D), shelf life (SL, in days), reflectance percentage (L, in %), chroma index (ratio a/b, being "a" the absorbance at 540 nm wavelengths and "b" the absorbance at 675 nm wavelengths), soluble solids content (SS, in °Brix), pH, titratable acidity (TA, in g of citric acid per 100 g of homogenate juice), and firmness (F, in %) in the population composed by two parental tomato RILs (L18 and L1; obtained by Rodriguez *et al.* 2006), their F_1 (second cycle hybrid L18 x L1), and the segregating generations F_2 (obtained by selfing the F_1), BC_1 (F1 x L18) and BC_2 (F1 x L1).

al., 2016; Zhao *et al.*, 2016, Trainor *et al.*, 2017) and in genomic studies (Chen and Ishwaran, 2012). However, the direct applications in practical situations of plant genetics and breeding, as were reported in this paper, have been infrequent. Though Random Forest was used in studies on wide genomic associations, detection of correlation among phenotypic traits and molecular markers and identification of different fruits, classification of breeding populations exclusively based in phenotypes is a vacant application (Biau, 2012; Chen

Table 4. Percentage of errors in global classification (GC) and per generation (L1: parental RIL 1, L18: parental RIL 18, F_1: second cycle hybrid L18 x L1, BC_1: backcross F_1 x L18, BC_2: backcross F_1 x L1, these genotypes are the six basic generations derived from RILs 1 and 18 obtained by Rodriguez *et al*, 2006, to initiate a breeding program) considering different number of trees (NT) in each applied Random Forest

NT	GC	F_1 (10)*	L1 (9)	L18 (8)	BC1 (46)	BC_2 (47)
50	41.30	100	100	100	18.18	47.06
100	32.61	100	100	100	9.09	35.29
150	30.43	100	100	100	9.09	29.41
200	28.26	100	100	100	9.09	23.53
300	28.26	100	100	100	9.09	23.53
400	30.43	100	100	100	9.09	29.41
500	30.43	100	100	100	9.09	29.41

* Numbers in parenthesis indicate the final number of plants per generation evaluated in the research

Table 5. Predicted classification of plants into groups (basic breeding generations L1: parental RIL 1, L18: parental RIL 18, F_1: second cycle hybrid L18 x L1, BC1: backcross F_1 x L18, BC_2: backcross F_1 x L1) in the training and validation of Random Forest (Method Out of Bag, OOB). L1 and L18 are tomato RILs obtained by Rodríguez *et al.* (2006)

Actual Group	Predicted Group					Total (%)
	F_1	L1	L18	BC_1	BC_2	
F_1	0 (0)	0 (0)	0 (0)	5 (50)	5 (50)	**10 (100)**
L_1	0 (0)	0 (0)	0 (0)	0 (0)	9 (100)	**9 (100)**
L18	0 (0)	0 (0)	0 (0)	4 (50)	4 (50)	**8 (100)**
BC_1	0 (0)	0 (0)	0 (0)	42 (91)	4 (9)	**46 (100)**
BC_2	0 (0)	3 (6.4)	0 (0)	8 (17.0)	36 (76.6)	**47 (100)**

and Ishwarab, 2012). Regarding this vacancy, a common but not desirable situation in plant breeding programs is the loss of identification of plant material in a given plot or genotype's mix when manipulating seeds, especially when exotic germplasm was introgressed (Dempewolf *et al.*, 2017). Hence the availability of reliable classification methods based on fast and easy to evaluate phenotypic traits and generated by supervised tools would be greatly advantageous. However, due to the different genetic structure of the various breeding populations coexisting in the same program, it is necessary to evaluate the adequacy of developing classification tools exclusively based on phenotypic variability using Random Forest for each types of population. Classification of individuals for improving their management in hybridization, recombination and selection, taking into account not only general variability but also traits mostly contributing to its conformation is a key step in breeding programs, as well as in defining the best phenotype for each situation (Niazian and Niedbała, 2020). According to our results, Random Forest was a better classifying technique than the most widely applied Discriminant Analysis. However, classification was better by either technique in final stages than in early stages of breeding programs. In final stages, when variability is dispersed and fixed among pure lines or other uniform population by effect of both artificial selection and inbreeding, best phenotypes appear to be more precisely classified than in early stages, when variability is created by crosses and recombination. Both the greater level of gene segregation and particularly the higher linkage disequilibrium of the population of basic generations, account for the less reliable classification obtained in it by both statistical methods. In consequence, the different genetic structure between both populations hinders the establishment of a robust algorithm for

Table 6. Contribution of each fruit trait to the classification with Random Forest of five breeding basic generations according to Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG). The genotypes are the basic generations derived from RILs 1 and 18 obtained by Rodriguez *et al*, 2006, to initiate a breeding program. Fruit traits: weight (W, in g), diameter (D, in cm), height (H, in cm), shape index (SI, ratio H/D), shelf life (SL, in days), reflectance percentage (L, in %), chroma index (ratio a/b, being "a" the absorbance at 540 nm wavelengths and "b" the absorbance at 675 nm wavelengths), soluble solids content (SS, in °Brix), pH, titratable acidity (TA, in g of citric acid per 100 g of homogenate juice), and firmness (F, in %).

Traits	MDA	MDG
D	2.99	3.62
H	2.49	2.23
SI	5.26	6.08
W	0.18	2.69
SL	-0.79	1.65
L	0.84	2.28
a/b	0.59	1.78
SS	3.82	2.82
pH	-1.27	1.84
F	-1.52	1.19
TA	1.41	2.01

Table 7. Significance of the Sequential Test (ST) for the Linear Discriminant Functions (LDF) in the Discriminant Analysis applied to the population of basic generations derived from RILs 1 and 18 obtained by Rodriguez *et al*, 2006, to initiate a breeding program

Steps of ST	Null Hypothesis	Canonical Correlation (CC)	Square CC	F-value	*p*-value
1	None LDF is significant	0.86	0.73	3.06	<0.0001
2	Only LDF 1 is significant	0.73	0.53	2.19	0.0022
3	Only LDF 1 and 2 are significant	0.67	0.45	1.80	0.0632
4	LDF 1, 2, and 3 are significant	0.43	0.18	0.94	0.4968

The first two LDF, with standardized coefficients and traits, were:
LDF 1 = 1.97 D – 1.14 H + 0.14 SI – 0.50 W – 0.07 SL – 0.14 L – 0.72 a/b – 0.57 SS – 0.22 pH - 0.72 F + 0.15 TA
LDF 2 = 1.67 D + 0.14 H + 0.59 SI – 1.16 W – 0.26 SL + 0.83 L + 0.41 a/b - 0.44 SS – 0.06 pH + 0.26 F + 0.18 TA

categorizing earlier breeding step generations. However it allows a better classification in the final step (the population of RILs) since they represent different gene associations from the same single cross. In fact, RILs also had a high level of segregation (Pratta *et al.*, 2011) but their linkage disequilibrium was low (Cambiaso *et al.*, 2019). Though imbalance of data could also partially explain a less robust classification, in this research the number of individuals composing the unities of classification was certainly more noticeable in the population of basic generations than in the population of RILs but this is a common fact in plant breeding assays due to the different genotypic constitution of parents, F_1 and segregating F_2 and BCs.

The identification of variables most important for classification was more accurate and robust with Random Forest than with Discriminant Analysis, hence description of phenotypes was more precise using the first technique. Additionally, the genetic structure of a population may be assessed by multivariate methods, as Principal Component Analysis, whose main application is related to the characterization of general variability. Though this method is often used for classification, it is not adequate enough for categorizing. However, it can be used in a preliminary approach to reduce the data dimensionality and then apply classification methods such as Random Forest or Discriminant Analysis to obtain an appropriate classification (Hastie *et al.*, 2008). Finally, it is interesting to point out that Discriminant Analysis is a parametric statistical technique while Random Forest is a non parametric one. One of the assumptions in Discriminant Analysis is that the variables come from a multivariate normal distribution. However, this assumption is not a requirement for Random Forest application, which becomes an additional advantage.

Random Forest was more accurate and robust than

Table 8. Predicted classification of plants into groups (breeding basic generations: L1: parental RIL 1, L18: parental RIL 18, F_1; second cycle hybrid L18 x L1, BC_1: backcross F_1 x L18, BC_2: backcross F1 x L1) in the Discriminant Analysis, RILs 1 and 18 were obtained by Rodriguez et al, (2006)

Actual Group	Predicted Group					Total (%)
	F_1	L1	L18	BC_1	BC_2	
F_1	0 (0)	0 (0)	0 (0)	5 (50)	5 (50)	**10 (100)**
L_1	0 (0)	0 (0)	0 (0)	0 (0)	9 (100)	**9 (100)**
L18	0 (0)	0 (0)	0 (0)	4 (50)	4 (50)	**8 (100)**
BC_1	0 (0)	0 (0)	0 (0)	42 (91)	4 (9)	**46 (100)**
BC_2	0 (0)	3 (6.4)	0 (0)	8 (17.0)	36 (76.6)	**47 (100)**

Discriminant Analysis for classifying tomato genotypes by phenotypic fruit quality traits at two different usual situations of plant breeding. Though a specific application such as identification of eventual unknown group of plants, was approached in the present research, a wide use of this technique in plant genetics and breeding can be proposed from these results. For instance, the evaluation of general variability, the identification of traits that most contribute to this variability, and even the definition of the best phenotype at different steps in breeding programs, are potential areas of application for Random Forest.

BIBLIOGRAPHY

Abu-Ellail F.F.B., Hussein E.M.A., El-Bakry A. (2020) Integrated selection criteria in sugarcane breeding programs using discriminant function analysis. Bull. Natl. Res. Cent. 44: 21-35. https://doi.org/10.1186/s42269-020-00417-6

Alhusain L., Hafez A.M. (2017) Cluster ensemble based on Random Forests for genetic data. BioData Min. 10: 101-125. https://doi.org/10.1186/s13040-017-0156-2

Amit Y., Geman D. (1997) Shape quantization and recognition with randomized trees. Neural Comput. 9: 1545-1588. https://doi.org/10.1162/neco.1997.9.7.1545

Bernard S., Heutte L., Adam S. (2009) Influence of hyperparameters on Random Forest accuracy. In: Benediktsson J.A., Kittler J. and Roli F. (Eds.) Multiple Classifier Systems. MCS 2009. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 171-180. https://doi.org/10.1007/978-3-642-02326-2_18

Biau G. (2012) Analysis of a Random Forests model. J. Mach. Learn. Res. 13: 1063-1095. https://dl.acm.org/doi/10.5555/2188385.2343682

Breiman L. (2001) Random Forests. Mach. Learn. 45: 5-32. https://doi.org/10.1023/A:1010933404324

Cabodevila V.G., Cambiaso V., Rodríguez G.R., Picardi L.A., Pratta G.R., Capel C., Lozano R., Capel J. (2021) A segregating F_2 population from a tomato second cycle hybrid allows the identification of novel QTL for fruit quality traits. Euphytica. 217: 453-461. https://doi.org/10.1007/s10681-020-02731-6

Cambiaso V., Giménez M.D., Pereira da Costa J.H., Vazquez D.V., Picardi L.A., Pratta G.R., Rodríguez G.R. (2019) Selected genome regions for fruit weight and shelf life in tomato RILs discernible by markers based on genomic sequence information. Breed. Sci. 69: 447–454. https://doi.org/10.1270/jsbbs.19015.

Chen X., Ishwaran H. (2012) Random Forests for genomic data analysis. Genomics. 99: 323-329. https://doi.org/10.1016/j.ygeno.2012.04.003

Dempewolf H., Baute G., Anderson J., Kilian B., Smith C., Guarino L. (2017) Past and future use of wild relatives in Crop Breeding. Crop Sci. 57: 1070-1082. https://doi.org/10.2135/cropsci2016.10.0885

Duda R., Hart P., Stork D. (2000) Pattern Classification. Wiley, Hoboken NJ, USA.

FAOSTAT (2017) https://www.fao.org/faostat/en/#data/QCL (accessed July 2021).

Gerszberg A., Hnatuszko-Konka K., Kowalczyk T., Kononowicz A.K. (2015) Tomato (*Solanum lycopersicum* L.) in the service of biotechnology. Plant Cell Tissue Organ Cult. 120: 881–902. https://doi.org/10.1007/s11240-014-0664-4

Hastie T., Tibshirani R., Friedman J. (2008) The elements of Statistical Learning. Data Mining, Inference, and Predictions. Springer, New York, NY, USA.

Kearsey M.J., Pooni H.S. (1996) The Genetical Analysis of Quantitative Traits. Chapman and Hall, London, UK.

Lapins K., Nash S.W. (1957) Discriminant function analysis in identification of peach varieties in nursery trees. Can. J. Plant Sci. 37: 12-25. https://doi.org/10.1007/BF02853700

Larose D., Larose C. (2015) Data Mining and Predictive Analytics. Wiley, Hoboken NJ, USA.

Lynch D.R., Schaalje G.B., Tai G.C.C., Young, D.A. (1987) Use of canonical discriminant analysis in assessing the merit of crosses in terms of breeding goals. Am. Potato J. 64: 385–395. https://doi.org/10.1007/BF02853700

Mahuad S.L., Pratta G.R., Rodriguez G.R., Zorzoli R., Picardi L.A. (2013) Preservation of *Solanum pimpinellifolium* genomic fragments in recombinant genotypes increased tomato fruit quality. J. Genet. 92: 195-203. https://doi.org/10.1007/s12041-013-0245-z

Niazian M., Niedbała G. (2020) Machine Learning for Plant Breeding and Biotechnology. Agriculture. 10: 615-640. https://doi.org/10.3390/agriculture10100436

Pereira da Costa J.H., Rodríguez G.R., Pratta G.R., Picardi L.A., Zorzoli R. (2014) Pericarp polypeptides and SRAP markers associated with fruit quality traits in an interespecific tomato backcross. Genet. Mol. Res. 13: 2539-2547. https://doi.org/10.4238/2014.January.24.10

Pratta G.R., Rodriguez G.R., Zorzoli R., Valle E.M., Picardi L.A. (2011) Phenotypic and molecular characterization of selected tomato recombinant inbred lines derived from a cross *Solanum lycopersicum* x *S. pimpinellifolium*. J. Genet. 90: 229-237. https://doi.org/10.1007/s12041-011-0063-0

Probst P., Wright M.N., Boulesteix A.L. (2019) Hyperparameters and tuning strategies for random forest. Data Min. Knowl. Discov. 9: 1-15. https://doi.org/10.48550/arXiv.1804.03515

Rodriguez G.R., Pratta G.R., Zorzoli R., Picardi L.A. (2006) Recombinant lines obtained from an interspecific cross among *Lycopersicon* species selected by fruit weight and fruit shelf life. J. Am. Soc. Hortic. Sci. 131: 651-656. https://doi.org/10.21273/JASHS.131.5

Singh A., Ganapathysubramanian B., Singh A.K., Sarkar S. (2016) Machine Learning for high-throughput stress phenotyping in plants. Trends Plant Sci. 21: 110-124. https://doi.org/10.1016/j.tplants.2015.10.015

Sivakumar V., Celine V.A., Venkata Ramana C. (2017) Discriminant function method of selection in vegetable cowpea genotypes. Int J. Curr. Microbiol. Appl. Sci. 10: 4954-4958. https://doi.org/10.20546/ijcmas.2017.610.469

Stephan J., Stegle O., Beyer A. (2015) A random forest approach to capture genetic effects in the presence of population structure. Nat. Commun. 6: 7432-7442. https://doi.org/10.1038/ncomms8432

Trainor P.J., De Filippis A.P., Rai S.N. (2017) Evaluation of classifier performance for multiclass phenotype discrimination in untargeted metabolomics. Metabolites. 21: 742-762. https://doi.org/10.3390/metabo7020030.

Zhao J., Bodner G., Rewald B. (2016) Phenotyping: using Machine Learning for improved pairwise genotype classification based on root traits. Front. Plant Sci. 7: 1083-1100. https://doi.org/10.3389/fpls.2016.01864

ACKNOWLEDGEMENTS

Authors are grateful to Drs. Sabina Mahuad and Victoria Cabodevila for generating phenotypic data during their respective Doctoral Thesis, and to Agencia Nacional de Promoción Científica y Tecnológica, Ministerio de Ciencia y Tecnología, Argentina, for financial supporting.

AUTHOR CONTRIBUTION STATEMENT

Conceptualization: Pratta G.R. Data curation: Vitelleschi M.S., Pratta G.R. Formal analysis: Vitelleschi M.S., Faviere G. Funding acquisition: Pratta G.R. Investigation: Vitelleschi M.S., Faviere G., Pratta G.R. Methodology: Vitelleschi M.S. Project administration: Vitelleschi M.S., Pratta G.R. Resources: Vitelleschi M.S., Faviere G. Supervision: Vitelleschi M.S., Pratta G.R. Writing-original draft: Pratta G.R. Writing-review & editing: Vitelleschi M.S., Pratta, G.R.

BAG | Journal of Basic & Applied Genetics

Vol XXXV (1): 53-64; Jan.-July 2024

Journal of the Argentine Society of Genetics

DEVELOPMENT OF GENETIC STOCKS OF SUNFLOWER WITH RESISTANCE TO SUNFLOWER CHLOROTIC MOTTLE VIRUS

CONFORMACIÓN DE RECURSOS GENÉTICOS DE GIRASOL CON RESISTENCIA AL VIRUS DEL MOTEADO CLORÓTICO DEL GIRASOL

Fernández Moroni I.[1], Lenardon S.[2,3,4], Giolitti F.[2,3], Álvarez D.[5], Poverene M.[1,6], Presotto A.[1,6], Cantamutto M.[7]

[1]Dpto. Agronomía, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina.

[2]Instituto Nacional de Tecnología Agropecuaria (INTA), Córdoba, Argentina.

[3]Instituto de Patología Vegetal (IPAVE), UFYMA, Córdoba, Argentina.

[4]Universidad Nacional de Río Cuarto (UNRC), Facultad de Agronomía y Veterinaria (FAV), Río Cuarto, Argentina.

[5]Estación Experimental Agropecuaria Manfredi (E.E.A. Manfredi), INTA, Manfredi, Córdoba, Argentina.

[6]Centro de Recursos Naturales Renovables de la Zona Semiárida-Consejo Nacional de Investigaciones Científicas y Técnicas (CERZOS-CONICET), Bahía Blanca, Argentina.

[7]E.E.A. Hilario Ascasubi, INTA, Hilario Ascasubi, Argentina.

Corresponding author:
Ivana Fernández Moroni
ivana.fernandez@uns.edu.ar

ORCID 0000-0001-9779-1184

Cite this article as:
Fernández Moroni I., Lenardon S., Giolitti F., Álvarez D., Poverene M., Presotto A., Cantamutto M. 2024. DEVELOPMENT OF GENETIC STOCKS OF SUNFLOWER WITH RESISTANCE TO SUNFLOWER CHLOROTIC MOTTLE VIRUS. BAG. Journal of Basic and Applied Genetics XXXV (1): 53-64.

Received: 04/12/2023
Revised version received: 05/06/2024
Accepted: 05/05/2024

General Editor: Elsa Camadro

DOI: 10.35407/bag.2024.35.01.04

ISSN online version: 1852-6233

Available online at
www.sag.org.ar/jbag

ABSTRACT

The common race of sunflower chlorotic mottle virus (SCMoV-C) can cause severe yield losses in susceptible genotypes of sunflowers if infection occurs at early plant stages. In Argentina, SCMoV-C is widespread in sunflower production fields and even if its incidence is generally low, in some cases it can reach up to 95%. To date, no complete resistance to SCMoV-C has been detected in commercial cultivars. In the search for resistant germplasm, wild sunflower (*Helianthus annuus* L.) populations from Argentina were tested, as they were exposed to natural selective pressure during their naturalization. After artificial inoculation with SCMoV-C, symptom-free plants were selected and grown for controlled self-pollination, sibling crosses and crosses with inbred lines. Recurrent selection for non-symptomatic plants and self-fertility significantly increased the frequency of asymptomatic individuals after SCMoV-C inoculation in the development germplasm. After eight generations of recurrent selection and controlled crosses, four genetic stocks with complete SCMoV-C resistance were developed. These genetic stocks could be used for breeding programs and genetic studies. The genetic stocks were registered in the Active Sunflower Germplasm Bank of the National Institute of Agricultural Technology (INTA, EEA-Manfredi), for maintenance and public distribution.

Key words: germplasm, *Helianthus annuus*, potyvirus, pre-breeding.

RESUMEN

En los genotipos susceptibles de girasol, la cepa común del *Virus del moteado clorótico del girasol* (SCMoV-C) puede causar graves pérdidas de rendimiento si la infección ocurre en las primeras etapas del desarrollo de la planta. En Argentina, el SCMoV-C está muy extendido en los campos de producción de girasol y aunque su incidencia es generalmente baja, en algunos casos puede llegar hasta el 95%. Hasta ahora, no se ha detectado resistencia completa a SCMoV-C en cultivares comerciales. En la búsqueda de germoplasma resistente, poblaciones de girasol silvestre (*Helianthus annuus* L.) de Argentina fueron testeadas, ya que durante su naturalización estuvieron expuestas a presión selectiva natural. Después de la inoculación artificial con SCMoV-C, se seleccionaron y cultivaron plantas libres de síntomas y se realizaron cruzamientos controlados, entre hermanos, con líneas endogámicas y autofecundaciones. La selección recurrente de plantas asintomáticas y autofértiles aumentaron considerablemente la frecuencia de individuos asintomáticos después de la inoculación con SCMoV-C en el germoplasma en desarrollo. Después de ocho generaciones de selección recurrente y cruces controlados, se desarrollaron cuatro stocks genéticos con resistencia completa a SCMoV-C. Este germoplasma podría utilizarse para programas de mejoramiento y estudios genéticos. Los stocks genéticos fueron registrados en el Banco de Germoplasma Activo de Girasol del Instituto Nacional de Tecnología Agropecuaria (INTA, EEA-Manfredi), para su mantenimiento y distribución pública.

Palabras clave: germoplasma, *Helianthus annuus*, potyvirus, premejoramiento.

INTRODUCTION

Sunflower (*Helianthus annuus* L.) yield is threatened by several diseases that cause economic losses. Generally, the disease problems are caused by specific or generalist fungal pathogens, but viruses are a potential hazard due to their degree of symptomatology (Vasquez and de Romano, 2006; Gontcharov, 2014; Gulya *et al.*, 2019). In Argentina, sunflower chlorotic mottle virus (SCMoV) (Dujovny *et al.*, 1998; 2000) is the most widespread virus infecting cultivated and wild sunflower (Cabrera Mederos *et al.*, 2020). Although its incidence is generally low, i.e. less than 3% in some cases, it can reach up to 95% (Lenardon, 1994). In susceptible genotypes, SCMoV infection at the early plant stages can cause generalized chlorosis, reduced growth, and yield losses exceeding 50% (Lenardon *et al.*, 2001).

SCMoV (*Potyvirus helichloromaculae*) belongs to family *Potyviridae* (ICTV, 2024). Viruses of this genus are transmitted from plant to plant mainly during the feeding action of infected aphid vectors. Five natural hosts of SCMoV were identified: *H. annuus*, *H. petiolaris* L., *Eryngium* sp., *Dipsacus fullonum* L. and *Ibicella lutea* L. In the epidemiology of the virus, the most important are *D. fullonum* and *Eryngium* sp., which are biennial and perennial weeds, respectively, that allow the virus to pass from one growing season to the next (Cabrera Mederos *et al.*, 2020). Preventive measures are needed to manage plant virus diseases since there are no curative treatments when crops are established. Resistant or tolerant genotypes are a simple, economical, and sustainable way to manage viral diseases. In combination with cultural practices (including chemical applications against vectors) and biological control, integrated disease management maximizes the likelihood of reducing yield losses (Jones, 2004; Tatineni and Hein, 2023).

Currently, two strains of SCMoV virus affect sunflower, the chlorotic ringspot strain (Giolitti *et al.*, 2010) and the common (C) strain (Dujovny *et al.*, 1998; 2000). The latter is the most widely distributed in Argentina (Cabrera Mederos *et al.*, 2020).

Lenardon *et al.* (2005) explored the susceptibility to SCMoV-C of more than 200 public and private sunflower inbred lines. Of these, only three lines showed partial resistance to this pathogen. The best response was observed in line L33 (Advanta Semillas S.A.I.C), which was linked to limited systemic infection due to scarce and isolated symptoms of chlorotic mottling and moderate reduction of yield components.

In plants, disease resistance can be genetically controlled by one, a few, or many genes, and it can be partial or total (Agrios, 2005). Total resistance implies that the virus cannot colonize its host, while in partial resistance there is colonization by the pathogen, although it is suppressed. Generally, the severity of symptoms reflects the level of virus replication and accumulation in the host (Revers *et al.*, 1999; de Ronde *et al.*, 2014).

Many crops have undeniably benefited from the useful traits of their wild relatives (Hajjar and Hodgkin, 2007). In sunflower, cytoplasmic male sterility, herbicide tolerance, modified fatty acid profile, disease resistance, among other traits have been successfully introgressed into the cultivated gene pool with very important economic consequences (Seiler *et al.*, 2017).

Naturalized wild sunflower, *H. annuus var. annuus* L., is distributed across the central region of Argentina (Poverene *et al.*, 2002). The invasive process has been associated with high phenotypic (Cantamutto *et al.*, 2010a; Presotto *et al.*, 2009) and genetic diversity (Garayalde *et al.*, 2011; Hernández *et al.*, 2019), probably exacerbated by an intense gene flow from cultivated sunflower (Ureta *et al.*, 2008).

Wild Argentine sunflowers could have developed resistance genes under natural selective pressure, due to the wide diffusion of the virus. To date, no specific research has been carried out on the reaction of wild sunflower germplasm to SCMoV-C. If wild Argentine sunflowers are virus resistant, there is a challenge of introgressing this trait into domestic sunflower. This paper reports the results of 8-years testing and selection for resistance to SCMoV-C in wild Argentine sunflowers, aimed at forming a useful source of germplasm for sunflower breeding.

MATERIALS AND METHODS

Screening for resistance to SCMoV-C in wild sunflower naturalized in Argentina

Nine wild *H. annuus* populations collected from representative geographical habitats (Cantamutto *et al.*, 2008), expressing different phenotypes in a common garden experiment (Cantamutto *et al.*, 2010b), were selected to initiate the evaluation and selection procedure (Table 1). The evaluations were conducted in November 2004, August 2005, and January 2006 on a minimum of 43 plants of each population. Every wild population was evaluated at least twice completing 575 screened plants. The commercial hybrid Contiflor 7 was used as a susceptible control.

Virus maintenance and artificial inoculation

The artificial inoculation and selection of plants without disease symptoms were carried out in greenhouses at the Instituto de Patología Vegetal, Instituto Nacional de Tecnología Agropecuaria (IPAVE-INTA), Córdoba, Argentina. SCMoV-C was maintained on susceptible sunflower cultivars in the greenhouse. Infected leaves

Table 1. Wild *Helianthus annuus* populations from Argentina

Population code	Location	Province	Eco-region
AAL	Adolfo Alsina	Buenos Aires	Pampa
BAR	Colonia Barón	La Pampa	Spinal
DIA	Diamante	Entre Ríos	Spinal
LMA	Las Malvinas	Mendoza	Plains and tablelands forest
CAR	Carhué	Buenos Aires	Pampa
RAN	Rancul	La Pampa	Spinal
MAG	Media Agua	San Juan	Plains and tablelands forest
JUM	La Carlota	Córdoba	Pampa
RCU	Río Cuarto	Córdoba	Spinal

were collected and stored at -80 °C until used as a source of inoculum. Infected leaves were ground in a buffer solution, pH 7, containing silicon carbide added as abrasive. The inoculum was applied to expanding sunflower leaves at V2-6 (Schneiter and Miller, 1981), using a high-pressure airbrush apparatus. Further details about virus maintenance and inoculation protocol were described by Lenardon *et al.* (2005).

Reproduction and crosses of select plants

Selected plants were transplanted in the experimental field of Universidad Nacional del Sur (UNS) Agronomy Department, Bahía Blanca, Argentina, with a spacing of 30 cm between plants and 100 cm between plots. Drip irrigation was provided to supply the water demand of the plants.

The fertile heads used for the controlled crosses were covered with paper or polyamide bags at the R4 stage, for insect and pollen exclusion. The male fertile plants used as females were emasculated manually in the morning and pollinated in the late afternoon. Pollination was carried out with fresh pollen collected from covered heads. Crosses involving the wild resource as the maternal parent were not emasculated because of its high degree of self-incompatibility (Gutierrez *et al.*, 2014).

Generation of the SCMoV-C resistant genetic stocks

Two cycles of selection were performed to conform the SCMoV-C resistant genetics stocks.

First cycle of selection: introgression of the SCMoV-C resistant trait into a domestic strain background

The first cycle of selection comprised five round of SCMoV-C inoculation, selection of asymptomatic plants and its controlled reproduction. Started with the selection of non-symptomatic plants of wild sunflower accessions, BAR and CAR. SCMoV-C free symptoms plants of wild sunflower, as well as new plant selected on next generations, were self-pollinated, sibling mated or interbreeding with the male sterile inbred lines (IL): A10, HA89 and A09, susceptible to SCMoV-C.

Second cycle of selection: fixing the SCMoV-C resistant trait

In 2009, the S10 family, a segregant germplasm developed in the first cycle of selection (Figure S1), was chosen to be the donor of the resistant trait on the second cycle of selection. Four asymptomatic plants after SCMoV-C infection of S10 (S10aRR, S10bRR, S10cRR, and S10dRR) were crossed with A09 and B09 inbred lines. Both lines are susceptible to SCMoV-C. A09 is a male-sterile inbred line with PET1 cytoplasm; while B09 is the male-fertile maintainer line of A09, with normal *H. annuus* cytoplasm (Garayalde *et al.*, 2015; González *et al.*, 2015). Due to the branching condition of S10RR plants, several heads were used in reciprocal controlled crosses between B09 on manual emasculated flowers, and as a pollen donor in crosses with A09. The progenies of those crosses were cultivated in the experimental field during the following season. At R4, heads were bagged until maturity to produce self-pollinated seeds.

In G5 (fifth generation), the progeny of self-fertile plants of the crosses (A09xS10RR; S10RRxB09 and B09xS10RR), were selected to generate the next generation. In G6 (sixth generation), SCMoV-C asymptomatic and self-fertile plants of F_2 families were chosen to ongoing selection. In G7 (seventh generation), F_3 families were submitted to another round of SCMoV-C inoculation. Selected F_3 plants, belonging to different families, were crossed between them or self-pollinated. The offspring of these crosses constituted four genetic stocks named GS-1, GS-2, GS-3, and GS-4. A detailed description of the the successive crosses performed on the second cycle of selection was described in Figure S2, in the supplementary material.

Phenotypic reaction to virus infection and serological virus analysis of resistant plants

The evaluated plants were grown in three-liter plastic pots; in each pot, 1-2 seeds were sown. To break dormancy, seeds had previously undergone a stratification treatment in plastic trays on moistened paper for one week at 4-7 °C (ISTA, 2004). The

susceptible cultivars Contiflor 17 or Contiflor 7 were used as a susceptible control.

SCMoV-C inoculation and selection of plants without the disease symptoms were performed artificially and individually on each plant. The reaction of the plants to SCMoV-C infection was assessed 15 days after inoculation and, on selected asymptomatic plants, after transplanting to the experimental field in the UNS at the reproductive stage. Inoculated plants were classified visually according to the leaf symptoms expression as: 1) SS, when chlorotic mottling was confluent over the entire lamina; 2) MR, when chlorotic mottling was mild; and 3) RR, when they had no visible disease symptoms (Figure 1). Some plants that had very few isolated, chlorotic pinpoints compatible with SCMoV-C after transplant, at the reproductive stage, were recorded as RR*.

On G8 (eighth generation), the visual diagnosis of inoculated plants was confirmed by the DAS-ELISA test. DAS-ELISA was performed as described by Lenardon *et al.* (2005) using an antiserum obtained by Dujovny *et al.* (1998) at least 15 days after SCMoV-C inoculation.

Field SCMoV-C inoculation of the genetic stocks

The selected genetic stocks were cultivated and artificially inoculated with SCMoV-C at V2-V4 in the experimental field of INTA in Manfredi, Córdoba, Argentina, during the 2013-14 growing season. Plant reaction was observed 15 days after inoculation and classified as previously described.

Phenotypic characterization of the SCMoV-C genetic stocks

Seeds of the genetic stocks, preconditioned as previously indicated, were arranged in multi-cell trays (70cc) filled with organic substrate (Terrafertil MULTIPRO®) in October 2014. They were grown in greenhouses until being transplanted in the experimental field of the UNS Agronomy Department, approximately one month later, at V2-4. Between 8 and 31 plants were spaced every 30 cm and grouped in plots according to the germplasm origin. Water demand was supplied by drip irrigation.

Morphological and physiological traits were used to describe the genetic stocks. The following characteristics were determined on mature plants: height (PH), number of leaves on the main stem (LP), leaf area (LA) of a leaf located halfway up the main stem (Aguirrezábal *et al.*, 1996), presence of a main head (MH), anthocyanins (ANT), branches (BRA) on the main stem, and head number per plant (HP). Branching was classified into apical, basal, or fully branching. The number of self-fertilized seeds per head (SAF) was counted in one or two heads per plant. Cypsela length (CL) and width (CW) were determined for 20 to 40 cypselas, and the fresh biomass of 1000 cypselas (B1000) was estimated, based on the fresh weight of at least four samples of 30 to 150 cypselas of each genetic stock, according to Equation 1.

B1000 (g) =(1000*PFcip)/Ncip [Equation 1]

where B1000: fresh biomass of 1000 cypselas, PFcip (g): fresh weight of cypselas, Ncip: number of cypselas.

The duration of the ontogenetic cycle was characterized as the duration in days elapsed from germination until 50% of the plants reached the R5 stage (G-R5) and until 50% of the plants were at the R9 stage (G-R9).

RESULTS AND DISCUSSION

Screening for resistance to SCMoV-C in wild sunflower naturalized in Argentina

No local or systemic symptoms of SCMoV-C disease were observed in at least 13% of the plants of all wild Argentine *H. annuus* populations after artificial inoculation (Table 2). This discovery was of interest because many public and private sunflower varieties from Argentina appear to be susceptible to SCMoV-C (Lenardon *et al.*, 2005), and once infected the virus can cause severe yield losses (Lenardon *et al.*, 2001). Until recently, Argentina was the only country with the presence of this virus, but Bello *et al.* (2023) recorded the first case in Brazil in 2021. This virus will continue to expand to other countries because it is transmitted by aphids, some weeds could act as natural reservoirs (Dujovny *et al.* 1998; Cabrera Mederos *et al.*, 2020) and many sunflower cultivars seem to be susceptible (Lenardon *et al.*, 2001).

Among the wild Argentine accessions, BAR showed the highest frequency, with almost 60% of the plants expressing no disease symptoms after inoculation. The BAR accession might have introgressed with *H. petiolaris* because it was collected in an area where both wild species coexist and it shows morphological evidence of gene flow (Gutierrez *et al.*, 2009). This agroecological situation could have made the emergence of novel variability possible through interspecific crosses.

Generation of the SCMoV-C resistant genetic stocks

The first cycle of virus resistance introgression into the domestic strain was started from a few plants selected from the BAR and CAR accessions, which reached the reproductive stage without symptoms after being inoculated with SCMoV-C and transplanted. Despite the high level of self-incompatibility of wild sunflowers (Fick and Miller, 1997; Gutierrez *et al.*, 2014), some seeds were obtained from self-fertilized and sibling-mated BAR plants. However, all its scarce progeny were susceptible to SCMoV-C. Instead, the crosses BAR-RRxCAR-RR and A10xBAR-RR produced abundant

Table 2. Mean percentage of asymptomatic plants of the wild *H. annuus* populations from Argentina after artificial inoculation with SCMoV-C between 2004 and 2006. Contiflor 7 (CF7) was used as susceptible control.

Population code	Artificial inoculations rounds (No.)	Inoculated plants (No.)	Asymptomatic plants (%)
RCU	3	126	56
RAN	2	43	35
DIA	3	130	49
LMA	2	48	30
MAG	3	105	19
JUM	3	109	35
BAR	3	113	61
AAL	3	115	13
CAR	2	52	35
CF7	3	33	0

seeds and SCMoV-C symptoms-free plants to continue the selection process (Table 3). Plants without disease symptoms after infection were self-pollinated, sibling-mated, or acted as a pollen donor to pollinate inbred sterile plants.

Three new rounds of SCMoV-C inoculation and controlled crosses between asymptomatic plants were performed in 2007-08, 2008-09, and 2009-10. However, after four generations of recurrent selection, the resultant germplasm lacked complete fixation of the SCMoV-C resistance trait, which denoted the absence of homozygosity of the character.

Reduced selection efficacy probably was because some plants without disease symptoms may have been erroneously diagnosed for further selection. Specifically, this situation may have happened in 2008-09, since 40% of plants of the susceptible control had no visual symptoms of SCMoV-C disease after inoculation (Table 3). Additionally, the high level of self-incompatibility of the genetics resources difficult the increment of homozygosity of the resistant trait, which was probably due to the self-incompatibility inherited from wild populations (Gutierrez *et al.*, 2014). In both wild and old sunflower cultivars, self-fertility is prevented mainly by a sporophytic self-incompatibility mechanism (Fick and Miller, 1997). However, sunflower breeding has broken down this reproductive mechanism and encouraging self-pollination (Gandhi *et al.*, 2005; Sun *et al.*, 2012) and now, cultivars can produce up to 100% of seed set under self-pollination (Astiz *et al.*, 2011). Segregation or co-segregation of the male-sterile trait, a product of crossing with male sterile inbred lines, contributes as a barrier to prevent self-fertilization of the resistant individuals, too.

To enhance the chance of fixing the resistant trait, in 2009-10 began the second cycle of crosses and selection. Disease symptom free plants of G4 families derived from crosses that produce abundant seeds, more than 50, by sibling crosses or self-pollination were selected. But, at maturity, only one of the selected crosses produced male fertile plants with a profuse production of pollen. The selected resource was named S10 family (Family identifier code: 25, Table 3). S10 was the result of the first cycle of selection, a segregating cross with wild cytoplasm and about half a percent of resistance.

S10 plants were crossed with the maintainer B09 and the male sterile A09 line. The objective of these crosses was to increase the level of self-fertilization to facilitate the subsequent introgression process of the SCMoV-C resistant trait into the elite germplasm of sunflower (Seiler *et al.*, 2017; Warburton *et al.*, 2017).

Only F_1 self-pollinated seed families were chosen to continue the second cycle of SCMoV-C selection. From that moment on, the chosen individuals were SCMoV-C symptom-free plants selected from progenies obtained by self-fertilization of families with abundant seed production. A higher degree of self-fertility can facilitate breeding since it allows a rapid increase in the homozygosity of the trait under selection (Cubero, 2003).

Once the offspring of highly self-fertilized seed production plants were used to continue the selection process, the proportion of progenies without SCMoV-C symptoms increased quickly as selection progressed. After another two generations submitted to selection, the proportion of SCMoV-C symptoms-free progenies increased up to 100% (Table 4). Similar results were obtained by Jan and Gulya (2006a; 2006b) during the development of genetic stocks with resistance to sunflower mosaic virus (SuMV) using wild-resistant material of American wild *H. annuus*.

Lenardon *et al.* (2005), in reference to SCMoV, as well as Jan and Gulya (2006b) in SuMV, stated that the resistance of their germplasms to the viruses of the *Potyviridae* family was controlled by a dominant gene. Although in many cases plant resistance to virus appears to be under simple dominant or recessive genetic control (Maule *et al.*, 2007), some reports show that resistance may also be under a few genes or polygenic inheritance (Gómez *et al.*, 2009; de Ronde *et al.*, 2014, Rossi *et al.*, 2015).

The segregation pattern of resistance suggests that more than one gene could participate in complete resistance against SCMoV-C. As in this study, research by Melchinger *et al.* (1998) in maize, Gore *et al.* (2002) in soybean, and Lee *et al.* (2017) in pepper, revealed a segregation pattern that did not fit a single-gene model for resistant and susceptible plants. Although it seems

Table 3. Seed production and offspring (O) reaction after SCMoV-C inoculation (RAI), of the controlled crosses performed during the first cycle of selection. Intercrosses with selected wild *Helianthus annuus* conformed the first generation (G1), and the successive generations (G2, G3, and G4) were obtained from new selections (highlighted in bold) and controlled crosses. ORAI: RR, plant without symptoms; MR, plant with mild chlorotic symptoms; SS, susceptible to SCMoV-C. SIB: sibling cross. ▽: self-pollination. Contiflor 17 (CF17) were used as susceptible control. A09, A10, and HA89 are inbred lines. The number at the beginning of an inoculated family indicates the family identifier code (e.g., 3RR SIB= a sibbing cross of resistant plants of BAR-RR x CAR-RR).

Family identifier code	Inoculated family	Produced seeds (No.)	Inoculated plants (No.)	ORAI		
				SS (%)	MR (%)	RR (%)
	G1	First generation	2006-07			
1	BAR-RR ▽	< 5	No plants were obtained			
2	BAR-RR SIB	< 10	3	0	100	0
3	**BAR-RR x CAR-RR**	**> 100**	**33**	**39**	**6**	**55**
4	**A10 x BAR-RR**	**> 100**	**53**	**63**	**24**	**13**
	CF17		18	100	0	0
	G2	Second generation	2007-08			
5	**A10 x 3RR**	**> 50**	**18**	**72**	**0**	**28**
6	A09 x 4RR	> 50	21	76	0	24
7	A10 x 4RR	> 50	22	83	0	17
8	**A09 x 3RR**	**> 50**	**18**	**67**	**17**	**16**
9	**HA89 x 3RR**	**> 50**	**24**	**54**	**42**	**4**
10	HA89 x 3RR	> 50	24	46	54	0
11	**3RR SIB**	**> 25**	9	56	22	22
12	4RR SIB	> 25	24	55	12	33
	CF17		6	100	0	0
	G3	Third generation	2008-09			
13	**5 RR SIB**	**> 40**	**5**	**0**	**0**	**100**
14	8 RR SIB	> 50	28	0	18	82
15	8 RR SIB	> 50	30	10	7	83
16	**8 RR SIB**	**> 50**	**18**	**0**	**0**	**100**
17	8 RR SIB	> 50	22	19	0	81
18	9 RR ▽	> 20	15	7	0	93
19	9 RR x 7 RR	> 50	30	10	3	87
20	**11 RR x 5 RR**	**> 50**	**13**	**0**	**0**	**100**
	CF17		5	60	0	40
	G4	Fourth generation	2009-10			
21	13 RR ▽	< 20	8	13	0	87
22	16 RR SIB	> 50	23	52	0	48
23	16 RR ▽	< 5	1	0	0	100
24	**20 RR SIB**	**> 50**	**42**	**52**	**0**	**48**
25	20 RR ▽	> 50	40	52	0	48
	CF17		19	100	0	0

plausible that resistance to SCMoV-C is governed by more than one gene, genetic analysis was outside the scope of the present study. As resistance evaluations on F_2 and F_3 progeny considered only two to ten inoculated individuals, a segregation model was not statistically determined. Future studies, focused on the inheritance of resistance could elucidate the genetic mechanism involved.

In G8, none of the progenies of the selected crosses had any visual symptoms of SCMoV-C after inoculation (Table 4). Negative DAS-ELISA results individually obtained from each inoculated plant confirmed that virus accumulation was strongly inhibited in these germplasms. This is the first time that germplasms completely resistant to SCMoV-C have been produced without segregation of the resistance trait. These genetic stocks represent an important source of genetic variability of SCMoV-C resistance, since no resistant sunflower cultivars are available, and no cases of complete resistance have been reported up to the present time. These germplasms will provide sunflower breeders with a source of resistance against SCMoV-C, should it become an economic problem.

Phenotypic characterization of the resistant genetic stocks

The use of wild relatives in breeding programs can be a great challenge because their useful traits could be masked by agronomically inferior background characteristics (Dempewolf *et al.*, 2017). Even beyond this, there is strong agreement on the benefits of conserving and expanding genetic breeding resources to address production constraints (Seiler *et al.*, 2017; Warburton *et al.*, 2017; Khoury *et al.*, 2022).

Table 4. Parent (P) and offspring (O) reaction after SCMoV-C inoculation (RAI) of the crosses performed during the second cycle of selection. F_2 progeny of crosses performed with selected S10, segregating resistant SCMoV-C population (family code 24, Table 3), with two inbred lines, A09 and B09, conformed the sixth generation (G6), and the successive generations (G7 and G8) were obtained from new selections (highlighted in bold) and controlled crosses. RAI: RR, plant without symptoms; MR, plant with mild chlorotic symptoms; SS, susceptible to SCMoV-C. B09, B10 and Contiflor 17 (CF17) were used as susceptible controls. SIB: sibling cross. ▽: self-pollination. ND: no data available. GS: Genetic Stock. The number at the beginning of an inoculated family indicates the family identifier code (e.g., 30 ▽= a self-pollination of resistant plants of F2(A09xS10cRR)). The letter RR followed by the superscript characters MR or * indicates that a selected RR plant showed mild chlorotic mottling or very few isolated chlorotic spots compatible with SCMoV-C infection after transplanting, respectively.

Family identifier code	Inoculated family	PRAI	ORAI (No.)		
			SS	MR	RR
	G6	**Sixth Generation**	**2011-12**		
26	F_2(A09xS10cRR)	ND	0	3	2
27	F_2(A09xS10cRR)	ND	0	3	1
28	F_2(A09xS10cRR)	ND	3	0	1
29	**F_2(A09xS10cRR)**	**ND**	**1**	**0**	**4**
30	**F_2(A09xS10cRR)**	**ND**	**0**	**0**	**4**
31	F_2(B09xS10bRR)	ND	0	2	1
32	F_2(B09xS10bRR)	ND	1	1	2
33	F_2(B09xS10bRR)	ND	0	1	0
34	F_2(B09xS10bRR)	ND	0	1	0
35	F_2(B09xS10bRR)	ND	1	0	0
36	**F_2(B09xS10dRR)**	**ND**	**0**	**1**	**3**
37	F_2(B09xS10dRR)	ND	1	0	1
38	**F_2(S10cRRxB09)**	**ND**	**0**	**0**	**5**
39	F_2(S10cRRxB09)	ND	0	2	2
40	F_2(S10cRRxB09)	ND	0	1	1
41	F_2(S10cRRxB09)	ND	3	0	2
42	F_2(S10cRRxB09)	ND	0	1	3
43	F_2(S10cRRxB09)	ND	1	0	3
44	F_2(S10cRRxB09)	ND	0	2	3
45	F_2(S10cRRxB09)	ND	1	0	3
	CF17		19	0	0
	G7	**Seventh Generation**	**2012-13**		
46	**29 ▽**	**RR**	**0**	**0**	**7**
47	29 ▽	RR	7	0	2
48	30 ▽	RR	2	0	5
49	30 ▽	RR	4	0	4
50	30 ▽	RR	0	0	2
51	36 ▽	RR	2	0	8
52	**36 ▽**	**RR**	**0**	**0**	**5**
53	**38 ▽**	**RR**	**0**	**0**	**10**
54	38 ▽	RR	1	0	9
55	38 ▽	RR	0	0	3
56	38 ▽	RR	0	0	6
	CF17		20	0	0
	G8	**Eighth Generation**	**2013-14**		
GS-1	53 ▽	RR^{MR}	0	0	68
GS-2	46 x 52	RR x RR^{MR}	0	0	34
GS-3	46 x 52	RR x RR*	0	0	34
GS-4	52 ▽	RR*	0	0	68
	B09		34	0	0
	B10		34	0	0

Table 5. Morphological and physiological traits (mean ± standard error) of SCMoV-C resistant genetic stocks (GS). PH: plant height; LP: leaves per plant; LA: foliar surface; HP: heads per plant; SAF: self-fertilized cypselas per head; B1000: thousand cypselas biomass; CL: cypsela length; CW: cypsela width; G-R5: days between germination and flowering; G-R9: days between germination and plant senescence.

Genetic stock	PH (cm)	LP (No.)	LA (cm²)	HP (No.)	SAF (No.)	B1000 (g)	CL (mm)	CW (mm)	G-R5 (days)	G-R9 (days)
GS-1	144,2 ± 5,5	27 ± 1	235,9 ± 21,9	17 ± 3	101 ± 41	33,3 ± 0,2	8,6 ± 0,2	4,7 ± 0,2	104	154
GS-2	176,4 ± 6,8	26 ± 1	295,6 ± 21,4	22 ± 3	83 ± 15	28,8 ± 2,3	7,3 ± 0,2	4,2 ± 0,1	104	147
GS-3	126,5 ± 10,3	27 ± 1	171,1 ± 28,5	18 ± 6	62 ± 14	23,8 ± 2,5	7,2 ± 0,2	4,2 ± 0,1	107	140
GS-4	95,5 ± 5,2	25 ± 1	96,1 ± 13,6	6 ± 1	54 ± 15	12,4 ± 0,7	6,2 ± 0,1	3,4 ± 0,1	107	133

As SCMoV-C resistant genetic stocks were the result of crosses between cultivated and wild biotype parents, they showed intermediate morphological, reproductive, and phenological characteristics between them, a fact commonly observed in this type of crossing (Gandhi *et al.*, 2005; Baack *et al.*, 2008; Presotto *et al.*, 2011).

The height of SCMoV-C resistant germplasms ranged from 95 to 177 cm. In GS-3 and GS-4 the height was less than 130 cm (Table 5) resembling the height of the parental inbred lines. The height of the lines involved in the development of a hybrid should be 40 cm shorter than the desired height of the hybrid, generally between 150-180 cm (Škorić, 2012).

Due to their wild origin, plants of the four resistant genetic stocks showed branches with several heads. In addition, they presented anthocyanins on the stem or leaf petioles, except for GS-4 (Table 6). These traits resembled the wild Argentine sunflower biotypes (Presotto *et al.*, 2011). However, all the SCMoV-C genetic stocks exhibited a main larger inflorescence (Table 6), 22 to 27 leaves on the main stem, with a leaf area exceeding 90 cm² (Table 5). Apical dominance is a characteristic of domesticated sunflowers, which are generally monoecious with large heads that produce large seeds (Škorić, 2012; Radanović *et al.*, 2018). The intermediate condition, larger main heads, and branching are the product of hybridization between wild and cultivated sunflowers (Presotto *et al.*, 2011).

The level of seed set under self-pollination of the resistant genetic stocks was higher than that observed in the wild populations from which they originated. The developed genetic stocks produced more than 50 seeds per head and the GS-1 genetic stock produced the highest number of seeds by self-fertilization, about 100 per head.

Full flowering (R5) took place between 99 to 107 days after seed germination. The flowering period was long, more than two weeks, because of the pluri-head plant morphology. The phenological stage R9 was reached between 26 and 58 days after flowering (Table 5). The length of the ontogenetic cycle was more than 130 days. The SCMoV-C-resistant germplasms could be considered to have a long growth cycle (Fick, 1978; Škorić, 2012).

The cypsela biomass of resistant genetic stocks was greater than 20 mg, except in GS-4, and the fruit length and width were equal to or greater than 6 and 3 mm, respectively (Table 5). Seed size and biomass were intermediate between the sunflower inbred lines (Pekcan *et al.*, 2015) and wild sunflower (Presotto *et al.*, 2009; 2011).

Male sterile plants were observed in GS-2 and GS-3. These germplasms have PET1-type male-sterile cytoplasm donated by A09 male-sterile line (Figure S2); presumably the nuclear fertility restoration genes are in a heterozygous state (Fick and Miller, 1997; Acquaah, 2012). The GS-2, GS-2, and GS-3 showed *Verticillium* wilt symptoms in the common garden. *V. dahliae* produces early leaf drying (Gulya *et al.*, 1997) and a reduction in crop production in susceptible genotypes (Wang *et al.*, 2021). Therefore, it is possible that evaluated quantitative traits had been undervalued in those genetic stocks. GS-1 did not exhibit any *Verticillium* wilt symptoms

Self-fertilized seeds of the resistant genetic stocks, GS-1, GS-2, GS-3 and GS-4, were deposited in the Active Germplasm Bank of the INTA Manfredi, registered with codes CGGI1351, CGGI1352, CGGI1353 and CGGI1354, respectively.

CONCLUSIONS

The wild Argentine *H. annuus* showed abundant plants without any SCMoV-C symptoms after inoculation. BAR showed the highest proportion among the nine accessions tested.

Crossing of the symptoms-free SCMoV-C BAR plants with inbred lines, followed by a recurrent selection for SCMoV-C resistance and self-fertility, led to the generation of four SCMoV-C resistant genetic stocks, GS-1, GS-2, GS-3 and GS-4, without phenotypic segregation for the resistance trait.

Among the resistant germplasms, GS-1 presented the most suitable phenotype for use as a SCMoV-C resistance donor to the cultivated sunflower. The traits in which it stood out were the absence of segregation for male fertility, high production of self-fertilized seeds,

Table 6. Morphological traits of SCMoV-C resistant genetic stocks (GS). BRA: plants with branches; A: absent; AB: apical branching; BB: basal branching; or FB: full branching. MH: plants with main head. ANT: plants with stem anthocyanin. (*) Plants with signs of Verticillium.

Genetic stock	BRA (%)				MH (%)	ANT (%)	Other traits
	A	AB	BB	FB			
GS-1	0	0	0	100	100	100	Presence of plants with only tubular flowers, larger on the margins than at the center of the head.
GS-2	0	16	0	84	100	68	Presence of male-sterile plants. Presence of plants with only tubular flowers, larger on the margins than at the center of the head. (*)
GS-3	5	28	0	67	100	67	Presence of male-sterile plants. Presence of plants with only tubular flowers, larger on the margins than at the center of the head. (*)
GS-4	10	20	13	57	100	0	(*)

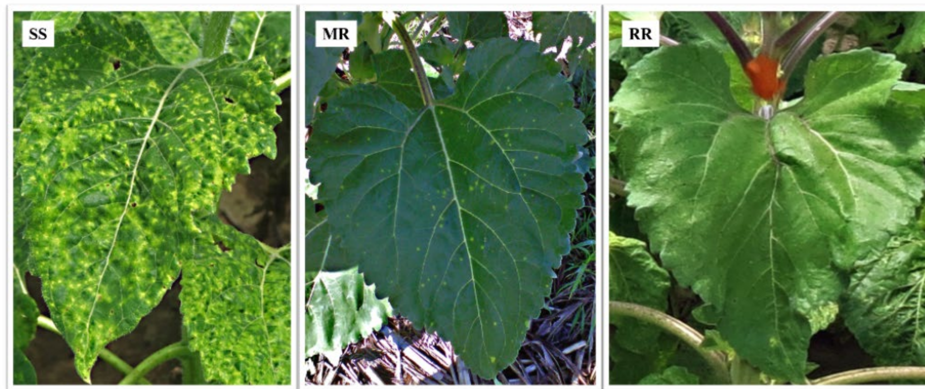


Figure 1. Classification of sunflower plant reaction after artificial SCMoV-C inoculation by lamina symptoms: SS: confluent chlorotic mottling over the entire lamina; MR: mild chlorotic mottling; RR: no visible disease symptoms of SCMoV-C.

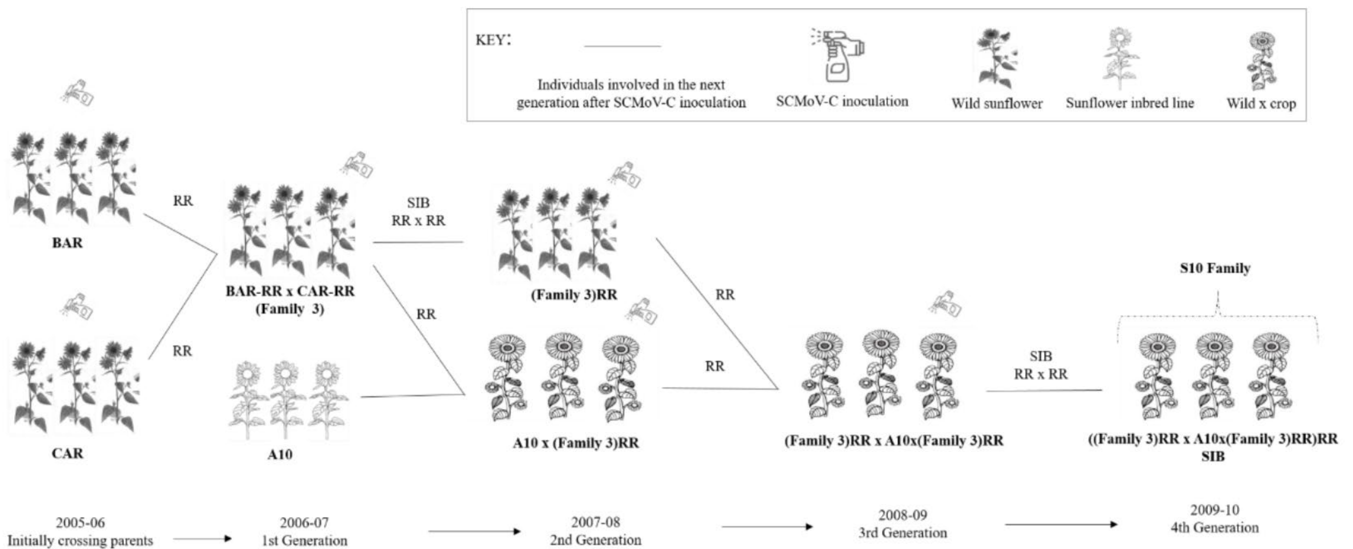


Figure S1. Overview of the crosses performed and inoculation rounds on the first cycle of selection against SCMoV-C infection to conform the S10 family. Plants without symptoms (RR) of SCMoV-C disease after artificial inoculation were chose to continuous selection. A10: male sterile inbred line, BAR: wild sunflower from Colonia Barón. CAR: wild sunflower from Carhué. SIB: sibling cross.

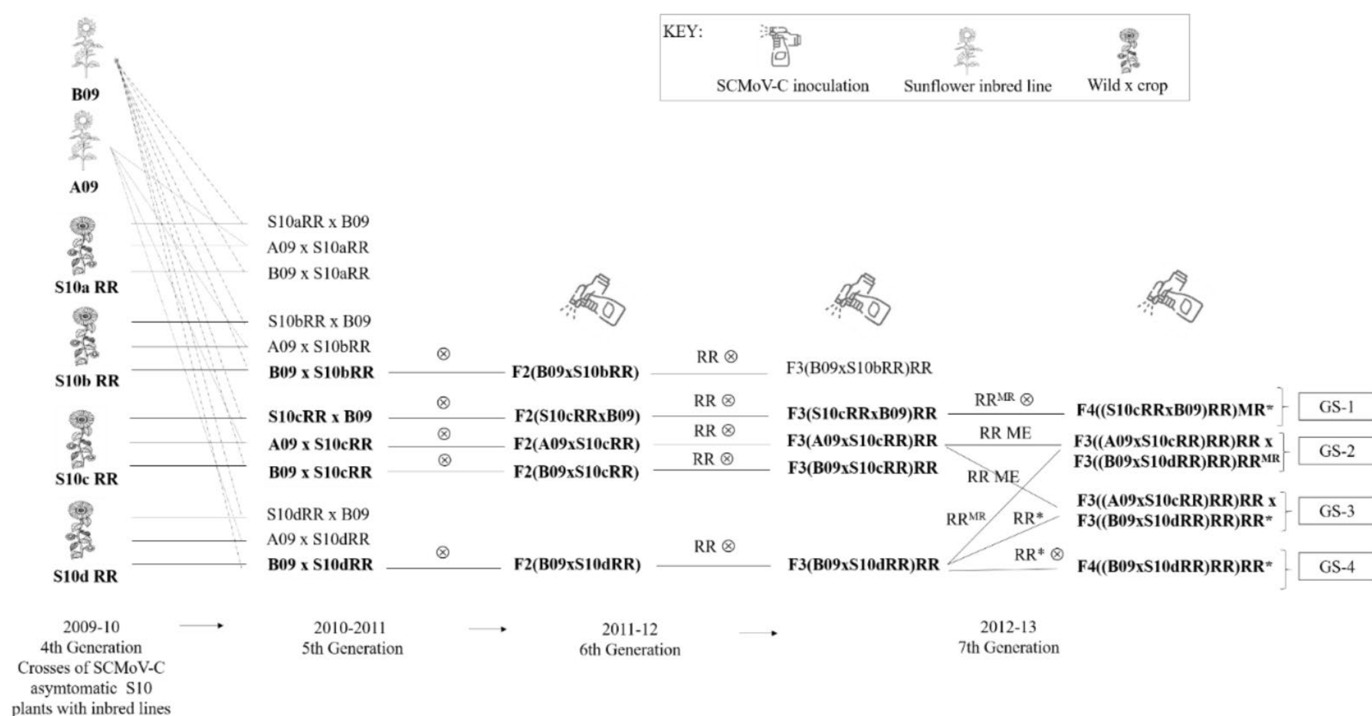


Figure S2. Overview of the successive intercrossings and inoculation rounds carried out in the second cycle of selection to conform the SCMoV-C resistant genetic stocks (GS). RR: plants without symptoms of SCMoV-C disease after artificial inoculation. The letter RR followed by the superscript characters MR or * indicates that a selected RR plant showed mild chlorotic mottling or very few isolated chlorotic spots compatible with SCMoV-C infection after transplanting, respectively. A09: male sterile inbred line. B09: A09 maintainer line. S10: segregant germplasm developed in the first cycle of selection. Lower-case letters added to the S10 to indicate different plants. ⊗: self-pollination.

presence of a main head larger than the secondary ones and seminal biomass greater than 30 mg.

These germplasms are open-access resources for breeders, their contribution to research or the development of new breeding lines or cultivars implies a commitment to appropriate recognition.

ACKNOWLEDGMENTS

We thank the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Agencia Nacional de Promoción Científica y Tecnológica (ANPCYT) and Universidad Nacional del Sur (UNS) for the fellowship to Ivana Fernández Moroni. This work was supported by grants from ANPCYT (PAE 37100 PICT 020; PICT-2012-2854), INTA and UNS (PGI 24A204) of Argentina.

REFERENCES

Acquaah G. (2012) Principles of plant genetics and breeding. Wiley-Blackwell Editorial, Oxford, UK.

Agrios G. (2005) Plant Pathology. Elsevier Academic Press Editorial, Amsterdam, Netherlands.

Aguirrezábal L., Orioli G., Hernández L., Pereyra V., Miravé J. (1996) Girasol. Aspectos fisiológicos que determinan el rendimiento. Unidad Integrada Balcarce Editorial. Buenos Aires. Argentina.

Astiz V., Iriarte L.A., Flemmer A., Hernández L.F. (2011) Self-compatibility in modern hybrids of sunflower (*Helianthus annuus* L.). Fruit set in open and self-pollinated (bag isolated) plants grown in two different locations. Helia 34 (54):129-138.

Baack E., Sapir Y., Chapman M., Burke J., Rieseberg L. (2008) Selection on domestication traits and quantitative trait loci in crop–wild sunflower hybrids. Mol. Ecol. 17: 666–677.

Bello V.H., Madoglio Favara G., Verruma Bernardi G., Marques Rezende J.A., Barbosa Salaroli R., Watanabe Kitajima E. (2023) First report of sunflower chlorotic mottle virus infecting sunflower plants in Brazil. Sci. agric. 80. https://doi.org/10.1590/1678-992X-2022-0035

Cabrera Mederos D., Torres C., Bejerman N., Trucco V., Lenardon S., Leiva Mora M., Giolitti F. (2020) Phylodynamics of sunflower chlorotic mottle virus, an emerging pathosystem. Virolog. 545: 33-39.

Cantamutto M., Poverene M., Peinemann N. (2008) Multi-scale analysis of two annual Helianthus species naturalization in Argentina. Agric. Ecosyst. & Environ. 123 (1–3): 69-74.

Cantamutto M., Poverene M., Presotto A., Alvarez D., Lenardon S., Rodríguez R., Martín Sánchez J., Fernández Moroni I., Giolitti F., Garayalde A., Haucke A., Bellido A., Fraysse M. (2010a) The argentine wild *Helianthus annuus* L. genetic resource. Helia 33 (52): 47-62.

Cantamutto M., Presotto A., Fernández Moroni I., Alvarez D., Poverene M., Seiler G. (2010b) High infraspecific diversity of wild sunflowers (*Helianthus annuus* L.) naturally developed in central Argentina. Flora 205 (5): 306–312.

Cubero J. (2003) Introducción a la mejora genética vegetal. Wiley-Blackwell Editorial, Oxford, UK.

de Ronde D., Butterbach P., Kormelink R. (2014) Dominant resistance against plant viruses. Front. Plant Sci. 5. https://doi.org/10.3389/fpls.2014.00307

Dempewolf H., Baute G., Anderson J., Kilian B., Smith C., Guarino L. (2017) Past and future use of wild relatives in crop breeding. Crop Sci. 57: 1070-1082.

Dujovny G., Sasaya T., Koganesawa H., Usugi T., Shoara K, Lenardon S. (2000) Molecular characterization of a new potyvirus infecting sunflower. Arch. Virol. 145: 2249-2258.

Dujovny G., Usugi T., Shoara K, Lenardon S. (1998) Characterization of a potyvirus infecting sunflower in Argentina. Plant Disease 82 (5): 470-474.

Fick G. (1978) Breeding and genetics. In: American Society of Agronomy (Ed.) Sunflower science and technology. Madison, US, pp. 279-338.

Fick G., Miller J. (1997) Sunflower breeding. In: American Society of Agronomy (Ed.) Sunflower technology and production. Madison, US, pp. 395-440.

Gandhi S., Heesacker A., Freeman C., Argyris J., Bradford K., Knapp S. (2005) The self-incompatibility locus (S) and quantitative trait loci for self-pollination and seed dormancy in sunflower. Theor. Appl. Genet. 111: 619-629.

Garayalde A.F., Poverene M., Cantamutto M., Carrera A.D. (2011) Wild sunflower diversity in Argentina revealed by ISSR and SSR markers: an approach for conservation and breeding programs. Ann. Appl. Biol. 158 (3): 305-317.

Garayalde A.F., Presotto A., Carrera A., Poverene M., Cantamutto M. (2015) Characterization of a new male sterility source identified in an invasive biotype of *Helianthus annuus* (L.). Euphytica 206 (3): 579-595.

Giolitti F., Bejerman N., de Breuil S., Lenardon S. (2010) Identification and characterization of a new strain of Sunflower chlorotic mottle virus, a Potyvirus infecting Asteraceae in Argentina. J. Phytopathol. 158 (7-8): 536-541.

Gómez P., Rodríguez Hernández A., Moury B., Aranda M. (2009) Genetic resistance for the sustainable control of plant virus diseases: breeding, mechanisms and durability. Eur. J. Plant. Pathol. 125:1-22.

Gontcharov S.V. (2014) Dynamics of hybrid sunflower disease. Helia 37 (60): 99-104.

González J., Mancuso N., Alvarez D., Cordes D., Vázquez A. 2015. Contribution of the Argentine germplasm to the improvement of sunflower. Helia 38 (62): 121-140.

Gore M.A., Hayes A.J., Jeong S.C., Yue Y.G., Buss G.R., Saghai Maroof M.A. (2002) Mapping tightly linked genes controlling potyvirus infection at the Rsv1 and Rpv1 region in soybean. Genome 45: 592–599.

Gulya T., Harveson R., Mathew F., Block C., Thompson S., Kandel H., Berglund D., Sandbakken J., Kleingartner L., Markell S. (2019) Comprehensive Disease Survey of U.S. Sunflower: Disease Trends, Research Priorities and Unanticipated Impacts. Plant Disease 103: 601-618.

Gulya T., Rashid K., Masirevic S. (1997) Sunflower diseases. In: American Society of Agronomy (Ed.) Sunflower technology and production. Madison, US, pp. 263-380.

Gutierrez A., Carrera A., Basualdo J., Rodriguez R., Cantamutto M., Poverene M. (2009) Gene flow between cultivated sunflower and *Helianthus petiolaris* (Asteraceae). Euphytica 172 (1): 67-76.

Gutierrez A., Rueda F., Cantamutto M., Poverene M. (2014) Self-pollination and its implication in invasiveness of *Helianthus annuus* ssp. *annuus* and *H. petiolaris*. BAG 25: 5-15.

Hajjar R., Hodgkin T. (2007) The use of wild relatives in crop improvement: a survey of development over the last 20 years. Euphytica 156: 1-13.

Hernández F., Presotto A., Poverene M., Mandel J. (2019) Genetic diversity and population structure of wild sunflower (*Helianthus annuus* L.) in Argentina: reconstructing its invasion history. J. Hered. 110 (6): 746–759.

ICTV-International Committee on Taxonomy of Viruses (2024) https://ictv.global/report/chapter/potyviridae/potyviridae/potyvirus (accessed June 2024).

ISTA- International Seed Testing Association (2004) International rules for seed testing. International Seed Testing Association Editorial, Zürich, Switzerland.

Jan C.C., Gulya T.J. (2006a) Registration of a sunflower germplasm resistant to rust, downy mildew, and virus. Crop Sci. 46 (4): 1829.

Jan C.C., Gulya T.J. (2006b) Registration of three virus resistant sunflower genetic stocks. Crop Sci. 46 (4):1834-1835.

Jones R.A.C. (2004) Using epidemiological information to develop effective integrated virus disease management strategies. Virus Res. 100: 5–30.

Khoury C., Brush S., Costich D., Curry H., Haan S., Engels J.M.M., Guarino L., Hoban S., Mercer C.L., Miller A.J., Nabhan G.P., Perales H.R., Richards C., Riggins C., Thormann I. (2022) Crop genetic erosion: understanding and responding to loss of crop diversity. New Phytol. 233: 84-118.

Lee J.H., An J.T., Siddique M.I., Han M., Choi S., Know J.K., Kang B.C (2017) Identification and molecular genetic mapping of Chili venial mottle virus (ChiVMV) resistance genes in pepper (*Capsicum annuum*). Mol. Breeding 37 (121). https://doi.org/10.1007/s11032-017-0717-6.

Lenardon S. (1994) Síntomas de etiología viral en cultivos de girasol. In: INTA (Ed.) Enfermedades del girasol en la Argentina. Balcarce, Argentina, pp. 99-103.

Lenardon S., Bazzalo M., Abratti G., Cimmino C., Galella M., Grondona M., Giolitti F., León A. (2005) Screening sunflower for resistance to Sunflower chlorotic mottle virus and mapping the Rcmo-1 resistance gene. Crop Sci. 45: 735-739.

Lenardon S., Giolitti F., León A., Bazzalo M., Grondona M. (2001) Effect of Sunflower chlorotic mottle virus infection on sunflower yield parameters. Helia 24 (35): 55-66.

Maule A., Caranta C., Boulton M. (2007) Sources of natural resistance to plant viruses: status and prospects. Mol. Plant Pathol. 8 (2): 223-231.

Melchinger A., Kuntze L., Gumber R., Lübberstedt T., Fuchs E. (1998) Genetics basis of resistance to sugarcane mosaic virus in European maize germplasm. Theor. Appl. Genet. 96: 1151-1161.

Pekcan V., Evci G., Yilmaz M.I., Nalcaiyi A.S.B., Erdal S.Ç., Cicek N., Ekmekci Y., Kaya Y. (2015) Drought effects on yield traits of some sunflower inbred lines. J. Agric. For. 61 (4): 101-107.

Poverene M., Cantamutto M., Carrera A., Ureta M., Salaberry M., Echeverria M., Rodriguez R. (2002) El girasol silvestre (*Helianthus* spp.) en la Argentina: caracterización para la liberación de cultivares transgénicos. RIA 31 (2): 97-116.

Presotto A., Cantamutto M., Poverene M., Seiler G. (2009) Phenotypic diversity in wild *Helianthus annuus* from Argentina. Helia 32 (50): 37-50.

Presotto A., Fernández Moroni I., Poverene M., Cantamutto M. (2011) Sunflower crop-wild hybrids: Identification and risks. Crop Prot. 30: 611-616.

Radanović A., Miladinović D., Cvejić S., Jocković M., Siniša J. (2018) Sunflower genetics from ancestors to modern hybrids- A review. Genes 9 (11): 528. https://doi.org/10.3390/genes9110528.

Revers F., Le Gall O., Candresse T., Maule A. (1999) New advances in understanding the molecular biology of plant/potyvirus interactions. MPMI 12 (5): 367–376.

Rossi E., Borghi M., Di Renzo M., Bonamico M. (2015) Quantitative trait loci (QTL) identification for resistance to Mal de Río Cuarto virus (MRCV) in maize based in segregate population. Open Agric. J. 9: 48-55.

Schneiter A., Miller J. (1981) Description of sunflower growth stages. Crop Sci. 21: 901-903.

Seiler G., Qi L.L., Marek F. (2017) Utilization of sunflower crop wild relatives for cultivated sunflower improvement. Crop Sci. 57 (3): 1083-1101.

Škorić D. (2012) Sunflower genetics and breeding. Int. Monography. Editorial Serbian Acad. Sci. Arts, Novi Sad, Serbia.

Sun Y., Godwin I.D., Arief V.N., Delacy I.H., Jackway P.T., Lambrides C. J. (2012) Factors controlling self-fertility in sunflower: the role of GCA/SCA effects, alleles, and floret characteristics. Crop Sci. 52:128-135.

Tatineni S., Hein G.L. (2023) Plant viruses of agricultural importance: current and future perspective of virus disease management strategies. Phytopathology® 113 (2):117-141.

Ureta S., Carrera A., Cantamutto M., Poverene M. (2008) Gene flow among wild and cultivated sunflower, *Helianthus annuus* in Argentina. Agric. Ecosyst. Environ. 123: 343–349.

Vasquez A., de Romano A. (2006) Sunflower crop in Argentina to date. Helia 29:159-164.

Wang D., Su Z., Ning D., Zhao Y., Meng H., Dong B., Zhao J., Zhou H. (2021) Different appearance period of *Verticillium* wilt symptoms affect sunflower growth and production. J. Plant Pathol. 103: 513-517.

Warburton M.L., Rauf S., Marek L., Hussain M., Ogunola O., Sanchez Gonzalez J.J. (2017) The use of crop wild relatives in maize and sunflower breeding. Crop Sci. 57: 1227-1240.

BAG | Journal of Basic & Applied Genetics

Vol XXXV (1): 65-72; Jan.-July 2024

Journal of the Argentine Society of Genetics

REPRODUCTIVE BIOLOGY AND POLLEN-PISTIL COMPATIBILITY RELATIONSHIPS IN AN ARGENTINIAN COLLECTION OF *Stevia rebaudiana* BERTONI

BIOLOGÍA REPRODUCTIVA Y RELACIONES DE COMPATIBILIDAD POLEN-PISTILO EN UNA COLECCIÓN DE *Stevia rebaudiana* BERTONI EN ARGENTINA

Budeguer C.J.[1], Camadro E.L.[2], Erazzú L.E.[3]

[1]Facultad de Agronomía, Zootecnia y Veterinaria, Universidad Nacional de Tucumán (UNT), Argentina.

[2]Unidad Integrada Estación Experimental Agropecuaria (EEA) "Domingo Pasquale", Instituto Nacional de Tecnología Agropecuaria (INTA)- Facultad de Ciencias Agrarias (FCA), Universidad Nacional de Mar del Plata (UNMdP), Argentina. Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

[3]EEA Famaillá, INTA, Tucumán, Argentina.

Corresponding author:
Carlos Budeguer
carlos.budeguer@faz.unt.edu.ar

ORCID 0000-0002-4649-956X

Cite this article as:
Budeguer C.J., Camadro E.L., Erazzú L.E. 2024. REPRODUCTIVE BIOLOGY AND POLLEN-PISTIL COMPATIBILITY RELATIONSHIPS IN AN ARGENTINIAN COLLECTION OF *Stevia rebaudiana* BERTONI. BAG. Journal of Basic and Applied Genetics XXXV (1): 65-72.

Received: 02/04/2023
Revised version received: 06/17/2024
Accepted: 06/24/2024

General Editor: Elsa Camadro

DOI: 10.35407/bag.2024.35.01.05

ISSN online version: 1852-6233

Available online at www.sag.org.ar/jbag

ABSTRACT

Stevia rebaudiana Bertoni (Asteraceae) is a diploid species (2n=2x=22) with sexual and asexual reproduction. The sexual propagules are seeds produced by cross-pollination (allogamy) whereas the asexual propagules are either vegetative shoots or apomictic seeds. Various authors have reported that allogamy in this species is promoted by the expression of a sporophytic self-incompatibility (SSI) system. To introduce the cultivation of *S. rebaudiana* as a production alternative in Tucumán, Argentina, a germplasm collection of this species was established with accessions from four Argentinian provinces in the Famaillá Agropecuarian Experimental Station (EEA Famaillá), National Institute of Agropecuarian Technology (INTA). The reproductive biology of the collection was studied between 2014 and 2021 to develop strategies for breeding and conservation of these genetic resources. Fifty-six genotypes were analyzed, all of them were 2n=2x=22. Pollen viability was high (69.4 to 99.6%) in all the genotypes except in four of them, which exhibited low viability (36.0 to 51.5%) in 2015 and 2017. Forty-eight genotypic combinations were obtained by manual controlled crosses. In 12 of these combinations, one pollen tube was observed in the style zone and, in four of them, one pollen tube was observed in the embryo sac; these observations indicate, respectively, incompatible and compatible pollen-pistil relationships. Normal plump seeds were obtained in all compatible genotypic combinations. The observed incompatibility might be due to the functioning of the sporophytic homomorphic system and/or a cross-incompatibility system. The observed compatibility will allow the planification of controlled crosses within and between accessions of different geographical origins to generate genetically variable progenies for breeding purposes.

Key words: genetic resources, incompatibility, internal hybridization barriers, pollen viability

RESUMEN

Stevia rebaudiana Bertoni (Asteraceae) es una especie diploide (2n=2x=22) con reproducción sexual y asexual. Los propágulos sexuales son semillas producidas por polinización cruzada (alogamia), mientras que los propágulos asexuales son brotes vegetativos y semillas apomícticas. Varios autores han señalado que la alogamia en esta especie se ve favorecida por la expresión de un sistema de autoincompatibilidad esporofítica (SSI). Para introducir el cultivo de *S. rebaudiana* como alternativa productiva en Tucumán, Argentina, se estableció una colección de germoplasma de esta especie proveniente de cuatro provincias de la Argentina en la Estación Experimental Agropecuaria (EEA) Famaillá, Instituto Nacional de Tecnología Agropecuaria (INTA). Se estudió la biología reproductiva de la colección entre 2014 y 2021 para desarrollar estrategias de mejoramiento y conservación de estos recursos genéticos. Se analizaron 56 genotipos, que fueron 2n=2x=22. La viabilidad del polen fue alta (69,4 a 99,6%) excepto en cuatro de ellos que exhibieron baja viabilidad (36,0 a 51,5%) en 2015 y 2017. Se obtuvieron 48 combinaciones genotípicas mediante cruzamientos controlados manuales. En 12 de estas combinaciones, se observó un tubo polínico en la zona estilar y, en cuatro de ellas, un tubo polínico en el saco embrionario; estas observaciones indican, respectivamente, relaciones polen-pistilo incompatibles y compatibles. Se obtuvieron semillas rellenas normales en todas las combinaciones genotípicas compatibles. La incompatibilidad observada podría deberse al funcionamiento del sistema de autoincompatibilidad homomórfica esporofítica, un sistema de incompatibilidad cruzada, o ambos. La compatibilidad observada permitirá la planificación de cruzamientos controlados dentro y entre introducciones de diferentes orígenes geográficos para generar progenies genéticamente variables con fines de mejoramiento genético.

Palabras clave: recursos genéticos, barreras internas a la hibridación, viabilidad de polen, incompatibilidad.

INTRODUCTION

Stevia rebaudiana Bertoni (Asteraceae) is a diploid species (2n=2x=22) (Galiano, 1987; Frederico *et al.*, 1996; Wulff *et al.*, 1996; Oliveira *et al.*, 2004) of commercial interest due to the presence of glycosylated diterpenes steviol glycosides (SG). *S. rebaudiana* is the only species in genus *Stevia* that exhibits an intense and persistent sweet taste (Soejarto *et al.*, 1982) due to a high SG content in the leaves. These molecules are 250-300 times sweeter than sucrose and non-caloric sweeteners (Ceunen and Geuns, 2013), and are considered as a natural alternative to controversial synthetic sweeteners (Hastoy *et al.*, 2019). In nature, this species has two forms of reproduction: sexual and asexual. Sexual propagules (seeds) are produced by cross-pollination (allogamy), whereas asexual propagules develop from the base of the main stem (vegetative shoots) or by apomixis (apomictic seeds). It has been reported that sexual seeds have low viability, that viable seeds (those with embryos) are dark in color and that the non-viable ones are light and embryoless (Monteiro, 1980; Goettemoeller and Ching, 1999; Yadav *et al.*, 2011).

Commercial field stands are established with asexual propagules obtained by either stem cutting or micropropagation (Kryvenki *et al.*, 2008; Autade *et al.*, 2014; Modi and Kumar, 2018). However, in Argentina, the use of sexual seeds by small-scale farmers results in stands of very variable plants, both morphologically and phenologically.

A major difficulty in seed propagation is the low seed germination percentage, which can vary from 9.0 to 83.0%, depending on seed quality and germination conditions (Macchia *et al.*, 2007; Angelini *et al.*, 2018).

The flowers of *S. rebaudina* are very small (about 3 mm in length), with white petals, hermaphrodites, grouped in capitula which are arranged irregular cymes; they have protandry and mainly entomophilous pollination (Monteiro, 1980; Yadav *et al.*, 2011; Gantait *et al.*, 2018). Studies carried out in Italy determined that the most frequent pollinators were Hymenoptera, Diptera, and Lepidoptera (Martini *et al.*, 2016; Benelli *et al.*, 2017).

Various authors have reported that the species has a sporophytic self-incompatibility (SSI) system (Monteiro, 1980; Gantait *et al.*, 2018), whereas others have not specified whether the observed incompatibility is either sporophytic or gametophytic (Galiano, 1987; Frederico *et al.*, 1996; Yadav *et al.*, 2014; Caponio *et al.*, 2016). In the Asteraceae family, SSI has been described as the characteristic system of self-incompatibility (Frankel and Galun, 1977; Hiscock and McInnis, 2003).

Furthermore, there are contrasting proposals regarding the type of reproduction. Monteiro (1980), in a study of the reproductive biology of three groups of 60 plants each, all from Brazil, concluded that *S. rebaudiana* is an obligate apomictic species. However, Caponio *et al.* (2016) ruled out the existence of apomixis by analyzing the reproductive biology over three years of plants from the Argentinian provinces of Misiones (four) and Entre Ríos (two). The genotypes (and the number of them) characterized in each work were different and the contrasting conclusions do not allow for a generalization; on the contrary, they point out to the necessity of studying the reproductive biology of the accessions conserved at each genebank.

The genotypes of *S. rebaudiana* cultivated in Argentina derive from populations of Paraguay introduced into cultivation in the 1980´s without previous breeding processes, and there are no records on how they were collected. The variety *Criolla* derives from natural populations from the Amambay region in the Paraguay highlands. It is an open-pollinated variety which served as a source of raw material for the development of other current varieties that were incorporated into the international market with the world-wide expansion of the crop, more than two decades ago (Liaudat, 2021).

In order to introduce the cultivation of *S. rebaudiana* as a productive alternative in Tucumán, Argentina, we initiated an active genebank with 75 plants of this species collected from farmers' fields in four provinces (Tucumán, Jujuy, Misiones, and Formosa). These plants were established in the experimental field of the Famaillá Agropecuarian Experimental Station (EEA), National Institute of Agropecuarian Technology (INTA) (27º 01' 05" S, 65º 22' 42" W) in 2013. The aim of the present work was to study the reproductive biology of the active genebank collection of *S. rebaudiana* to develop strategies for breeding and conservation of these genetic resources.

MATERIALS AND METHODS

Plant material

Fifty-six plants of each of four accessions from Tucumán (13), Jujuy (14), Misiones (14) and Formosa (15), were randomly provided by the *S. rebaudiana* INTA Famaillá genebank (SRG). To study pollen-stigma/style compatibility relationships, individual flowers were emasculated and hand-pollinated following an incomplete diallel crossing design. The flowering stages were asynchronous, the flowering behavior of the same plants was variable in each year, and the size of the flowers was very small; thus, 56 plants were selected and each of them was identified with one letter according to origin -T (Tucumán), J (Jujuy), M (Misiones) and F (Formosa)- and numbers for each plant (genotype). Thirty of these plants were further selected for carrying out the crossing work in 2017, 2020, and 2021 because they had high pollen viability (69.4 to 99.6%) (Table 1). The plants were maintained in 2 L individual pots

Table 1. Procedence of the 30 plants of *Stevia rebaudiana* from the INTA Famaillá Genebank used in the study of pollen-stigma/style relationships.

Province	No. plants	Genotype Identification
Tucumán	7	T1.2, T2.3, T2.4, T2.6, T1.8, T2.16, T2.18
Jujuy	8	J3.10, J4.3, J4.5, J5.5, J6.1, J6.5, J8.5, J8.12
Misiones	7	M9.8, M9.9, M9.12, M9.18, M10.4, M11.10, M11.11
Formosa	8	F12.1, F12.2, F12.5, F12.6, F12.8, F12.10, F4, F6

outdoors, with a 1:1 mixture (v:v, peat moss:perlite) and a drip irrigation system.

Methods

Chromosome number and ploidy level

Chromosome number and ploidy level were determined by either root tip chromosome counting (20 plants) or chloroplast counting in the occlusive cells of leaf stomata (36 plants). For the former, root tips were pre-treated with 8-hydroxyquinoline 0.002 M for 4 h, fixed in ethanol 96º: glacial acetic acid (3:1, v/v) for 24 h, and transferred to 70% ethanol until use. For microscopic observations, root tips were hydrolysed in 1N HCl for 10 min at 60º C, rinsed with distilled water and stained with 2% acetic hematoxylin on a glass slide, squashed with a glass bar, covered with a cover slip, and observed under an optical microscope at x1000 magnification.

In plants in which direct chromosome counting could not be performed, the number of chloroplasts in the occlusive cells of the stomata was counted to estimate the ploidy level. In this regard, epidermal tissue was removed with tweezers from the abaxial leaf side for microscopic observations, placed in a glass slide on a drop of lugol (1g KI and 1g I in 100 ml of 70% ethanol), and covered with a cover slip. The number of chloroplasts was determined in either one of the two occlusive cells, in at least 10 stomata/genotype. An average of five to eight chloroplasts per occlusive cell were considered to be indicative of diploidy (Ordoñez *et al.*, 2016).

Pollen viability

For estimating pollen viability, pollen samples were taken at bloom from the dehiscent anthers of four flowers per plant and stained with an 0.5% acetocarmine solution (0.5 g carmine, 45 ml of glacial acetic acid, and 55 ml of distilled water) (Marks, 1954). To this end, a small amount of pollen was placed on a drop of the staining on a glass slide, and covered with a cover slip; then, observations were performed under an optical microscope at x125 magnification. Approximately 300 grains per plant were recorded. Fully stained pollen grains with well-defined contours were considered viable, whereas those that were colorless or poorly stained were considered non-viable.

Pollen-stigma/style compatibility relationships

Crosses between individual genotypes to determine pollen-stigma/style compatibility relationships were carried out following an incomplete diallel mating design, using the plants with high pollen viability as male parents. Flowering stems with receptive stigmata were removed from the plant and pollinated under a stereomicroscope to ensure that enough pollen was deposited on the stigmata. Forty-eight hours after pollination, and following Martin´s (1958) technique, pollinated pistils without ovaries were fixed in FAA (9:0.5:0.5, v/v/v, ethanol: glacial acetic acid: 37% formaldehyde) for at least 1 day. Fixed pistils were rinsed with distilled water, softened in 8 N NaOH solution for 2 h, rinsed again with distilled water, stained with a 0.1% aniline blue solution in tribasic potassium phosphate (0.1N PO_4K_3), mounted in a drop of glycerin on a glass slide, squashed with a cover slip and observed under a fluorescent microscope.

In the compatible genotypic combinations, three to four flowers were manually pollinated for seed formation, in order to confirm that fertilization would indeed occurred. On the other hand, to rule out the formation of viable seeds as a result of autogamy, branches with flower buds were isolated in the plants with voile cloth to avoid the arrival of foreign pollen.

RESULTS

Chromosome number and ploidy level

The number of chromosomes determined by root-tip chromosome counting in 20 plants was 2n=2x=22 (Fig. 1A). The number of chloroplast/occlusive cells of the

Table 2. Percentage of pollen viability -in four years- of 56 genotypes of *Stevia rebaudiana* from various geographic procedences.

Year	Genotype Identification (pollen viability %)[a]
2015	T1.1 (98), T1.2 (89,8), T1.5 (94,0), T1.6 (98,6), T1.8 (87,3), T2.2 (97,0), T2.4 (95,6), T2.5 (99,5), T2.10 (92,4), T2.12 (99,2), T2.15 (95,5), T2.16 (97,3), T2.18 (79,4)
	J3.12 (98,2), J3.14 (97,4), J3.15 (83,8), J3.16 (97,7), J3.2 (99,4), J5.5 (98,8), J6.16 (99), J6.20 (96,8), J6.5 (98,7), J7.5 (94,3), J8.12 (69,4), J8.7 (98,4), J8.9 (98,5)
	M9.1 (99,1), M9.4 (88,5), M9.8 (95,6), M9.12 (97,8), M9.18 (98,3), M10.2 (78,5), M10.6 (89,7), M10.8 (36,5), M10.10 (95,5), M11.2 (98,5), M11.6 (98,0), M11.7 (38,4), M11.11 (38,1)
	F12.1 (51,5), F12.2 (97,1), F12.3 (98,5), F12.4 (97,1), F12.5 (98,7), F12.6 (87,8), F12.7 (99,3), F12.8 (99,3), F12.10 (98,4), F12.11 (99,6), F12.12 (98,3), F12.13 (95,5), F12.14 (79,6), F12.15 (92,4)
2017	T1.1 (98), T1.2 (89,8), T1.6 (98,6), T2.12 (99), T2.18 (79,4)
	J3.14 (97,4), J4.3 (98,8), J5.5 (98,8), J6.20 (96,8), J8.12 (59,9)
	M9.18 (98,3), M10.4 (98,9), M11.6 (98), M11.7 (36), M11.11 (48,5)
	F12.1 (56,1), F12.3 (99,7), F12.4 (90,1), F12.7 (99,3), F12.8 (99,3)
2020	T2.4 (99,5), T2.3 (98,2), T2.16 (98,8), T2.18 (89,7), T9 (88,6)
	J4.5 (98), J 5.16 (99,6), J5.5 (99,7), J6.5 (88,2), J8.12 (100)
	M11.11 (98), M11.10 (87,5), M9.17 (99,4), M9.8 (99,6), M11.7 (83,5)
	F12.1 (89,5), F12.5 (98,4), F12.6 (100), F12.7 (99,8), F12.10 (100)
2021	T1.12 (99,5), T2.4 (97,7), T2.6 (98,4), T1.8 (99,2), T2.18 (99,6)
	J3.10 (99,8), J6.5 (99,5), J6.1 (99,8), J8.5 (97,8), J4.3 (99,2)
	M10.4 (98,3), M11.11 (98,7), M9.18 (99,6), M9.9 (99,2), M9.12 (99,6)
	F12.5 (100), F12.2 (99,7), F12.8 (100), F4 (99,8), F6 (99,7)

[a]High (> 80.0%), medium (80.0 to 60.0%), low (< 60.0%).

stoma in 36 of the 56 analyzed plants varied between six and seven and was, therefore, indicative of the diploid level (Fig. 1B). According to Huamán (1995), an average of six to eight chloroplasts/occlusive cell is indicative of the diploid level in potatoes.

Pollen viability

Pollen viability in the genotypes that reached the flowering stage was estimated in 2015, 2017, 2020 and 2021 (Table 2). High pollen viability was observed in all plants studied in 2017, 2020 and 2021, including those that had low viability in 2015 and 2017.

Pollen-stigma/style relationships

A total of 168 genotypic combinations were carried out in three years: 51 in 2017, 69 in 2020, and 48 in 2021. Between two and three stigmata per floret, and three to four florets per genotype were manually pollinated (Table 3).

In 2017 and 2020, only germinated pollen grains -with pollen tubes approximately of equal length than their diameters- were observed on the stigmata surfaces without stigma penetration (Figure 2A, B). Such genotypic combinations were considered incompatible. In stigmata that were self-pollinated, there was no

Table 3. Number of compatible (C) and incompatible (I) genotypic combinations obtained over three years in controlled crosses of *S. rebaudiana.*

Pollen-stigma/style relationship	Year		
	2017	2020	2021
No. of analyzed genotypic combinations	51	69	48
Type			
C	0	0	12
I	51	69	36

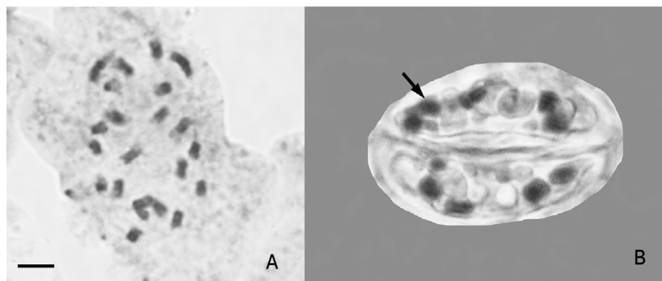


Figure 1. *Stevia rebaudiana.* Determination of the ploidy level by chromosome and chloroplast countings. (A) Mitotic metaphase in plant J4.5 with 2n=2x=22; (B) chloroplasts (arrow) in occlusive cells of a stomata in plant M9.9. Bar = 3 µm.

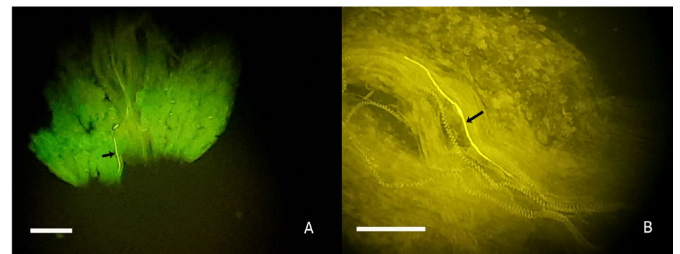


Figure 4. *Stevia rebaudiana.* Pollen tubes (arrow) growing towards the base of the style in compatible genotypic combinations. (A) genotypic combination M10.4 x J3.10; (B) genotypic combination 12.8 x T1.8. Bar = 20 µm.

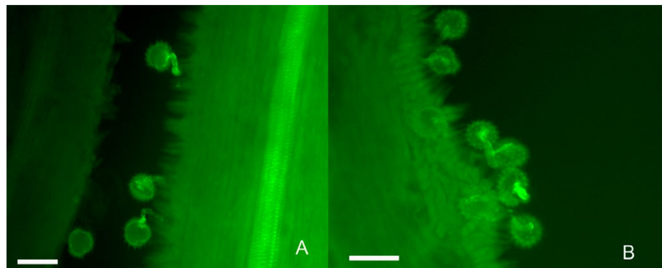


Figure 2. *Stevia rebaudiana.* Incompatibility with pollen tubes arrested in the stigma. (A) genotypic combination M9.18 x T2.18; (B) genotypic combination F12.8 x T2.6. Bar = 20 µm.

Figure 5. Seeds of *S. rebaudiana.* (A) clear and empty, obtained from self-fertilization of plant F12.5; (B) dark with embryo, obtained from genotypic combination T2.4 x M9. Bar = 2 mm.

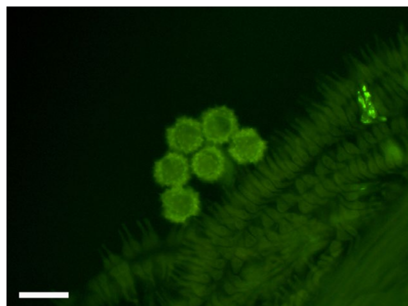


Figure 3. *Stevia rebaudiana.* Self-incompatibility with non-germinated pollen grains on a self-pollinated stigma of plant F12.8. Bar = 20 µm.

pollen grain germination (Figure 3). On the other hand, one pollen tube was observed in the style in each of 12 genotypic combinations out of the 48 performed in 2021; some of these pollen tubes reached the embryo sac (Figure 4A, B).

From all of the pollen-pistil compatible genotypic combinations, well-formed seeds were obtained (Table 4). All the plants whose florets were isolated with voile cloth to favor self-pollination produced empty, light-colored seeds without embryos (Figure 5A, B).

Table 4. Number (No.) of pollinated flowers and No. of seeds obtained in compatible genotypic combinations of *S. rebaudiana.*

Genotypic combination	No. pollinated flowers	No. seeds
F6 x T2.6	3	2
F12.8 x T1.8	3	3
M10.4 x J3.10	4	3
M9.18 x J6.5	4	2
J3.10 x M10.4	3	1
J6.5 x M9.18	3	3
T2.4 x M9.9	3	2
T2.6 x F12.5	3	2
F4 x F12.5	4	3
M9.18 x F4	4	4
J6.1 x M9.12	4	3

DISCUSSION

The chromosome number determined in root tip cells (2n=2x=22) in all the analyzed plants is the same reported by Frederico *et al.* (1996) for *S. rebaudiana* from Brazil. It is also the same reported by Oliveira *et al.* (2004) for 11 genotypes of *S. rebaudiana* from CENARGEN/EMBRAPA, Brazil. The basic number (x= 11) is common in most South American species of the *Stevia* genus (Galiano, 1987; Frederico *et al.*, 1996). Chloroplast counting in the stomata occluding cells was an effective technique to estimate the diploid level of the remaining analyzed plants; even though its application does not give an exact result, it allows for the distinction among ploidy levels (Poulsen Hornum and Camadro, 2021).

The percentage of pollen viability was high over the four years, with the exception of four plants in 2015 and 2017. Abdullateef *et al.* (2012) reported that pollen viability of three introductions of *S. rebaudiana* grown in the field in Malaysia ranged from 88.6 to 93.3%. Similarly, Caponio *et al.* (2016) reported 94.8 to 97.9% of pollen viability in plants from the provinces of Misiones (four) and Entre Rios (two), Argentina. Likewise, Monteiro (1980), working with 30 plants in Campinas, Brazil, determined that pollen viability ranged from 55% to 65%.

The analysis of pollen viability is of fundamental importance to identify fertile male parents to use as pollinators in breeding. Previously, Budeguer *et al.* (2018) reported abnormalities in meiosis and at tetrad stage in plants M10.8, M11.7, M11.11 and F12.1, which would explain their low pollen viability. In the present study, low pollen viability was observed in very few plants, and only in 2015 and 2017; these results could be attributed to the expression of genotype x environment interactions. In any case, pollen viability would not represent a problem in planning the crossing work with the SGR´s collection conserved at EEA Famaillá.

For the study of the pollen-stigma/style compatibility relationships, the manual pollination technique was extremely difficult to perform due to the very small size of the flowers and the poor adherence of the pollen on the stigmata; thus, a fine-tuning of the technique was previously required to perform the crosses on cut stems under a stereomicroscope. Flowering asynchrony in each year was an additional difficulty in the obtainment of the genotypic combinations.

Despite the difficulties mentioned above, 168 genotypic combinations were achieved in three years (2017, 2020, and 2021). Flower fixation 24 h after hand pollination proved adequate to observe pollen tube growth in the ovary. Caponio *et al.* (2016) studied the germination of pollen grains *in situ* in three ecotypes of *S. rebaudiana*; they observed that 88% of the attached pollen grains developed pollen tubes 8 h after

pollination, and that pollen tubes reached the lower third of the ovary 16 h later.

In 2021, pollen tubes were observed in the base of the style and growing into the ovary in 12 of the 48 genotypic combinations; in the remaining combinations, only tubes above the stigmatic papillae were observed, none of which penetrated the stigmata. This could be due to the functioning of the sporophytic homomorphic self-incompatibility system characteristic of the Asteraceae family (Frankel and Galun, 1977; Allen *et al.*, 2011; Gantait *et al.*, 2018), a cross-incompatibility system, or both. Frankel and Galun (1977) pointed out that the SSI is controlled by one multi-allelic *S*-locus (*S*-haplotype), being the reaction of the n pollen determined by the 2n genotype of the sporophytic tissue in which it was formed; thus, upon self-pollination all pollen grains of a plant will exhibit the same incompatibility reaction regardless of their own genotypes. However, the *S*-alleles may exhibit dominance/independence relationships that may differ in the pollen and pistils of the same plant, generating complex incompatibility reactions (Frankel and Galun, 1977; Hiscock and McInnis, 2003). Pollen-pistil cross-incompatibility -as observed in other species of various families of Dicots and Monocots (see examples in Marcellán and Camadro, 1996; Arias *et al.*, 2003; Ibañez and Camadro, 2014; Maune *et al.* 2018)- could also explain the observed incompatible reactions. Nevertheless, the simultaneous action of both systems cannot be discarded.

The results of this work allow us to conclude that the collection conserved in the SRG at EEA Famaillá, Tucumán, are diploid (2n=2x=22) and also allogamous because no seeds were formed in the flowers isolated with voile cloth. This lack of seed formation does not allow speculation on the possible apomictic behavior reported by Monteiro (1980) for the species. Pseudogamous apomixis requires the growth of the pollen tube in the embryo sac and the discharge of the generative cells into the central cell for the formation of the first endospermic cell after fusion of their respective nuclei, a phenomenon that cannot take place if a sporophytic incompatibility system with the same type of *S*-allele expression in the maternal and paternal tissues is acting.

In the compatible genotypic combinations, it was possible to achieve fertilization and viable seed formation, so it would be possible to plan crosses between genotypes from the SRG, taking into consideration different geographical origins, to increase the chances of obtaining genetically variable progenies. It is also advisable to previously determine the viability of the pollen of the plants that would act as paternal progenitors.

REFERENCES

Abdullateef R.A., Zakaria N.H., Hasali N.H., Osman M. (2012) Studies on pollen viability and germinability in accessions of *Stevia rebaudiana* Bertoni. Int. J. Biol. 4(3): 72-79.

Allen A.M., Thorogood C.J., Hegarty M.J., Lexer C., Hiscock S.J. (2011) Pollen–pistil interactions and self-incompatibility in the Asteraceae: New insights from studies of *Senecio squalidus* (Oxford ragwort). Ann. Bot. 108(4): 687–698.

Angelini L.G., Martini A., Passera B., Tavarini S. (2018) Cultivation of *Stevia rebaudiana* Bertoni and Associated Challenges. In Mérillon J.M. & Ramawat K.G. (Eds.) Sweeteners. Springer International Publishing, pp. 35–85. https://doi.org/10.1007/978-3-319-27027-2_8

Arias M., Camadro E.L., Díaz Ricci J., Castagnaro A.P. (2003) Breeding barriers between the cultivated strawberry, *Fragaria* x *ananassa*, and related wild germoplasm Euphytica, 136: 139-149.

Autade R., Fargade S., Borhade P., Udmale S., Choudhary R. (2014) In Vitro Propagation of *Stevia rebaudiana* (Bert.) a Natural, Non-Caloric Sweetener Herb. J. Cell Tissue Research, 14(3): 4659–4664.

Benelli G., Canale A., Romano D., Flamini G., Tavarini S., Martini A., Angelini L. G. (2017) Flower scent bouquet variation and bee pollinator visits in Stevia rebaudiana Bertoni (Asteraceae), a source of natural sweeteners. Arthropod-Plant Inte. 11(3), 381–388.

Budeguer C.J, Camadro E.L, Erazzú L.E. (2018) Pollen viability and meiotic irregularities in a collection of *Stevia rebaudiana* (Bertoni) Bertoni cultivated in Tucumán, Argentina. BAG. Journal of Basic and Applied Genetics, 29(2): 49-57.

Caponio I., Torres S.C., Andersson G., Norrmann G.A. (2016) Aspectos de la biología reproductiva de *Stevia rebaudiana* (Asteraceae). Bonplandia, 25(1): 59. https://doi.org/10.30972/bon.2511279.

Ceunen, S., Geuns, J.M.C. (2013) Steviol Glycosides: Chemical Diversity, Metabolism, and Function. J. Nat. Prod. 76: 1201–1228.

Frankel R., Galun E. (1977) Pollination Mechanisms, Reproduction and Plant Breeding (Vol. 2). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-81059-6.

Frederico A.P., Ruas P.M., Marin-Morales M.A., Ruas C.F., Nakajima, J.N. (1996) Chromosome studies in some *Stevia*. Cav. (Compositae) species from Southern Brazil. Braz. J. Genet. 19: 605–609.

Galiano N. (1987) Estudios cromosómicos en especies argentinas de *Stevia* (Compositae). Darwiniana, 28: 311–315.

Gantait S., Das A., Banerjee J. (2018) Geographical distribution, botanical description and self-incompatibility mechanism of genus *Stevia*. Sugar Tech. 20: 1–10.

Goettemoeller J., Ching A. (1999) Seed germination in *Stevia rebaudiana.* In: Janick J. (Ed.) Perspectives on new crops and new uses. ASHS Press, Alexandria, VA, pp. 510–511.

Hastoy, C., Cosson, P., Cavaignac, S., Boutie, P., Waffo-Teguo, P., Rolin, D., Schurdi-Levraud, V. (2019) Deciphering performances of fifteen genotypes of *Stevia rebaudiana* in southwestern France through dry biomass and steviol glycoside evaluation. Ind. Crops Prod. 128: 607-619.

Hiscock S.J., McInnis S.M. (2003) The Diversity of Self-Incompatibility Systems in Flowering Plants. Plant Biol. 5(1): 23–32. https://doi.org/10.1055/s-2003-37981.

Huamán Z. (1995) Técnicas citológicas para determinar el número cromosómico y la fertilidad de las papas. Guía de Investigación CIP 10. Centro Internacional de la Papa, Lima, Perú.

Ibañez M.S., Camadro E.L. (2014) Reproductive behavior of the wild carrots *Daucus pusillus* and *D. montanus* from Argentina Botany, 93: 279-286. dx.doi.org/10.1139/cjb-2014-0243.

Kryvenki M., Kosky R.G., Guerrero D., Dominguez M., Reyes M. (2008) Obtención de callos con estructuras embriogénicas de *Stevia rebaudiana* Bert. en medios de cultivo semisólidos. Biotecnol. Veg. 8(2): 91–98.

Liaudat S. (2021) *Stevia*. Conocimiento, propiedad intelectual y acumulación de capital . Prometeo Libros. 336 p.

Macchia M., Andolfi L., Ceccarini L., Angelini L.G. (2007) Effects of Temperature, Light and Pre-Chilling on Seed Germination of *Stevia rebaudiana* (Bertoni) Bertoni Accessions. Ital. J. Agron. 2(1): 55. https://doi.org/10.4081/ija.2007.55.

Marcellán O.N., Camadro E.L. (1996) Self- and cross- incompatibility in *Asparagus officinalis* and *A. densiflorus* cv. Sprengeri. Can. J. Bot. 74 (10): 1621-1625.

Marks G.E. (1954) An aceto-carmine glycerol jelly for use in pollen fertility counts. Stain Technol. 29: 277. https://doi.org/10.3109/10520295409115483.

Martin F.W. (1958) Staining and Observing Pollen Tubes in the Style by Means of Fluorescence. Stain Tech. 34(3): 125–128. https://doi.org/10.3109/10520295909114663.

Martini A., Tavarini S., Macchia M., Benelli G., Canale A., Romano D., Angelini L.G. (2016) Influence of insect pollinators and harvesting time on the quality of *Stevia rebaudiana* (Bert.) Bertoni seeds. Plant Biosyst. 151(2): 341–351. https://doi.org/10.1080/11263504.2016.1174173.

Maune J.F., Camadro E.L., Erazzú L.E. (2018) Cross-incompatibility and self-incompatibility: unrelated phenomena in wild and cultivated potatoes? Botany, 96 (1): 33 - 45.

Modi A., Kumar N. (2018) Conventional and Biotechnological Approaches to Enhance Steviol Glycosides (SGs) in *Stevia rebaudiana* Bertoni. In: Kumar N. (Ed.) Biotechnological Approaches for Medicinal and Aromatic Plants. Springer Singapore, pp. 53–62. https://doi.org/10.1007/978-981-13-0535-1_3.

Monteiro R. (1980) Taxonomia e Biologia da Reproduçao de *Stevia reabaudiana* (Bert.) Bert. Tesis Doctoral. Universidad de Campinas, São Paulo, Brasil.

Oliveira V.M., Forni-Martins E.R., Magalhães P.M., Alves M.N. (2004) Chromosomal and morphological studies of diploid and polyploid cytotypes of *Stevia rebaudiana* (Bertoni) Bertoni (Eupatorieae, Asteraceae). Genet. Mol. Biol. 27(2): 215–222.

Ordoñez, B., Orrillo, M., Bonierbale, M. (2016) Manual biología reproductiva y citológica de la papa. Lima (Perú). Centro Internacional de la Papa (CIP). ISBN 978-92-9060-477-8. 2a ed. 65 p.

Poulsen Hornum A., Camadro E.L. (2021) Expression of internal reproductive barriers in a germplasm bank accession of the wild potato *Solanum chacoense* Bitter in three ex situ regeneration cycles. Genet. Resour. Crop Evol. 68(3): 915–938. https://doi.org/10.1007/s10722-020-01034-y.

Soejarto D.D., Kinghorn A.D., Farnsworth N.R. (1982) Organoleptic evaluation of Stevia leaf herbarium samples for sweetness. J. Nat. Prod. 45: 590–599.

Wulff A.F., Hunziker J.H., Escobar A. (1996) Estudios cariológicos en Compositae VII. Darwiniana, 34: 213–231.

Yadav A.K., Singh S., Dhyani D., Ahuja P.S. (2011) A review on the improvement of stevia [*Stevia rebaudiana* (Bertoni)]. Can. J. Plant Sci. 91(1): 1–27.

Yadav A.K., Singh S., Rajeev R. (2014) Self-incompatibility evidenced through scanning electron microscopy and pollination behaviour in *Stevia rebaudiana*. Indian J. Agric. Sci. 84(1), 93–100.

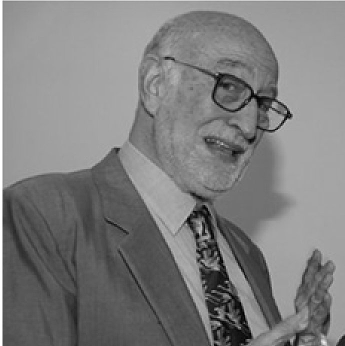
FUNDING

This study was partially funded by Instituto Nacional de Tecnología Agropecuaria (INTA; 2023-PD-L01-I071) and Universidad Nacional de Tucumán (PIUNT A705).

BAG Journal of Basic & Applied Genetics

Vol XXXV (1): 73-74; Jan.-July 2024

Journal of the Argentine Society of Genetics



MEMORIA - MEMOIR

DR. NÉSTOR OSCAR BIANCHI

1931 - 2023

Al cumplirse un nuevo aniversario del natalicio,
la SAG honra su memoria con la siguiente semblanza

El Dr. Néstor Oscar Bianchi fue un destacado investigador del CONICET, considerado uno de los principales y más eminentes genetistas de Argentina, reconocido con numerosos premios por aportes en el país y en el extranjero. Fue miembro fundador de la Sociedad Argentina de Genética (SAG) y de la Asociación Latinoamericana de Genética (ALAG).

Sus discípulos/as e integrantes del Instituto Multidisciplinario de Biología Celular (IMBICE, CONICET-UNLP-CIC) resaltan su labor científica y calidad humana. Lo recuerdan con las siguientes palabras:

Nacido en La Plata el 18 de abril de 1931. El Dr. Bianchi era Doctor en Medicina por la Universidad Nacional de La Plata (UNLP) (1955). Fue profesor de la cátedra de Citología de UNLP (1972-1975), de Genética Molecular de Eucariontes en el curso de Master en Genética del Instituto Nacional de Tecnología Aplicada (INTA) de Pergamino (1979-1986) y Profesor Visitante en las universidades de Louisiana (USA, 1970), San Francisco (USA, 1984-1985) y Helsinki (Finlandia, 1986-1992).

En 1974 fundó el Instituto Multidisciplinario de Biología Celular (IMBICE, CONICET- UNLP-CIC), del cual fue su Director desde ese mismo año hasta el año 2000. Su gestión como Director del IMBICE se caracterizó por un gran interés en el crecimiento del Instituto en todos los aspectos y por brindar apoyo a todo aquel que lo necesitara, ya sea en forma de ideas o de recursos económicos para poder llevar a cabo las investigaciones científicas en beneficio del Instituto. Su impacto en la historia del IMBICE fue de tal magnitud que resulta imposible hacer referencia al Instituto sin mencionar al Dr. Bianchi. También fue el gestor y ejecutor de la actividad de investigación científica en Tierra del Fuego, donde fue Director del Programa de EcoGenética, con lo que inició su actividad en el Centro Austral de Investigaciones Científicas (CADIC-CONICET) (1979).

Su profusa labor científica se vio reflejada en sus casi 300 publicaciones (artículos científicos, capítulos de libros, entre otros) sobre diversos temas relacionados con la genética; y en la formación de numerosos tesistas doctorales, tesistas de grado, becarios/as e investigadores/as, muchos de los cuales continúan en actividad, tanto en el IMBICE como en otras instituciones científicas locales y del exterior. Realizó aportes significativos en las áreas de estructura y función de los cromosomas humanos y animales, daño inducido a nivel cromosómico y molecular por radiaciones y agentes químicos, mecanismos de determinación sexual, evolución cromosómica, carcinogénesis, antropología y genética molecular poblacional humana y aspectos éticos y legales de la genética. Fue sin dudas uno de los investigadores de mayor prestigio a nivel local e internacional en genética y pionero de la citogenética en la Argentina y en Latinoamérica.

Obtuvo numerosas becas nacionales e internacionales y fue miembro de distintas instituciones nacionales e internacionales, incluyendo la Sociedad Médica de la Plata, la

Sociedad Argentina de Diabetes, la Sociedad Argentina de Genética (Miembro Fundador), la American Society for Human Genetics (Miembro Titular) y la New York Academy of Sciences, siendo además Académico Correspondiente de la Academia Nacional de Ciencias Exactas, Físicas y Naturales de Buenos Aires y Miembro de la Academia Latinoamericana de Ciencias.

Su gran jerarquía científica lo llevó a trabajar junto a destacados investigadores en el área de genética, tales como Lima de Faría, Albert de la Chapelle, Gerald Holmquist y James Cleaver, entre otros.

Su destacada labor en el campo de la investigación científica le permitió obtener numerosos premios, entre ellos: el «Premio Rubén Cherny» (1977), «Premio Konex» en el Área de Genética y Citogenética (1993) y el «Premio Ángel Gallardo 1986-1988» (1990) otorgado por la Academia Nacional de Ciencias Exactas, Físicas y Naturales de nuestro país. Asimismo, fue designado Doctor Honoris Causa de la Universidad Central de Ecuador (1997).

Sin dudas, el gran legado científico y humano del Dr. Bianchi permanecerá en la memoria de todos los que tuvimos oportunidad de compartir con él nuestra trayectoria científica. Vale recordar aquí las palabras que el propio Dr. Bianchi escribió en una reseña de su actividad científica: “Espero tener el privilegio de persistir en el recuerdo de otros cuando mi tiempo se haya acabado”. No tenga dudas, Doctor Bianchi. Lo recordaremos por siempre.

Dra. Graciela Bailliet (IMBICE),
Dra. Marta Lizarralde (CADIC)
y Dr. Alejandro Bolzán (IMBICE)

BAG | Journal of Basic & Applied Genetics