

1 The Biologist (Lima), 2026, vol. 24 (1), XX-XX.

2 DOI: <https://doi.org/10.62430/rtb20262412138>

3 Este artículo es publicado por la revista The Biologist (Lima) de la Facultad de Ciencias Naturales y Matemática, Universidad
4 Nacional Federico Villarreal, Lima, Perú. Este es un artículo de acceso abierto, distribuido bajo los términos de la licencia Creative
5 Commons Atribución 4.0 Internacional (CC BY 4.0) [<https://creativecommons.org/licenses/by/4.0/deed.es>] que permite el uso,
6 distribución y reproducción en cualquier medio, siempre que la obra original sea debidamente citada de su fuente original.



8 COMPUTATIONAL ANALYSIS OF DOUBLESEX ISOFORMS IN *ANOPHELES*

9 *GAMBIAE*

10 ANÁLISIS COMPUTACIONAL DE ISOFORMAS DE *DOUBLESEX* EN *ANOPHELES*

11 *GAMBIAE*

12
13 Mabel Serrano¹, Cheryl Marte¹, Nasla Gomez¹, Manuel Colomé-Hidalgo², Edian

14 Franco^{1,3}, Pedro María Alarcón-Elbal⁴ & Alejandro Vallejo Degaudenzi^{1,2}

15
16 ¹ Instituto Tecnológico de Santo Domingo (INTEC), Santo Domingo, República
17 Dominicana.

18 ² Universidad Autónoma de Santo Domingo (UASD) Santo Domingo, República
19 Dominicana.

20 ³ Instituto de Innovación en Biotecnología e Industria (IIBI) Santo Domingo, República
21 Dominicana.

22 ⁴ Grupo de Investigación de Zoonosis Transmitidas por Vectores (ZOOVEC),
23 Departamento de Producción y Sanidad Animal, Salud Pública Veterinaria y Ciencia y
24 Tecnología de los Alimentos (PASAPTA), Facultad de Veterinaria, Universidad CEU
25 Cardenal Herrera, CEU Universities, Alfara del Patriarca, Valencia, España.

26 *Corresponding author: avdegaudenzi@live.com

27 Running Head: Dsx isoforms in *Anopheles gambiae*

28 Serrano *et al.*

29

30 Mabel Serrano:  <https://orcid.org/0009-0002-3908-4195>

31 Cheryl Marte:  <https://orcid.org/0009-0001-2335-443X>

32 Nasla Gomez:  <https://orcid.org/0009-0003-2227-562X>

33 Manuel Colomé-Hidalgo:  <https://orcid.org/0000-0002-4562-6491>

34 Edian Franco:  <https://orcid.org/0000-0001-9715-9437>

35 Pedro María Alarcón-Elbal:  <https://orcid.org/0000-0001-5319-4257>

36 Alejandro Vallejo-Degaudenzi:  <https://orcid.org/0000-0003-2057-2332>

37

38 **ABSTRACT**

39 The doublesex (*dsx*) gene is involved in sex determination in *Anopheles gambiae sensu*
40 *stricto* and is a potential target for vector control. The objective was to computationally
41 characterize its sex-specific isoforms. Materials and methods: An *in silico* analysis of RA
42 and RB isoforms was performed using VectorBase data, evaluating gene architecture,
43 CDS length, and predicted protein products. RA and RB showed structural and protein-
44 length differences while preserving the reading frame. RB-derived theoretical variants
45 allowed the effect of terminal modifications to be explored without introducing frameshift
46 changes. The functional diversity of *dsx* is associated with alternative splicing and
47 supports its relevance as a molecular target for future genetic vector control research.

48 **Keywords:** alternative splicing – biosafety – *in silico* analysis – malaria – sex
49 determination – vector control

50

51

52 **RESUMEN**

53 El gen *doublesex* (*dsx*) participa en la determinación sexual de *Anopheles gambiae*
54 sensu stricto y constituye una posible diana para control vectorial. El objetivo fue
55 caracterizar computacionalmente sus isoformas sexo-específicas. Se realizó un análisis
56 *in silico* de las isoformas RA y RB con datos de VectorBase, evaluando arquitectura
57 génica, longitud del CDS y productos proteicos predichos. RA y RB mostraron diferencias
58 en organización estructural y longitud proteica, conservando el marco de lectura. Las
59 variantes teóricas derivadas de RB permitieron explorar el efecto de modificaciones
60 terminales sin introducir cambios de fase. La diversidad funcional de *dsx* se asocia al
61 splicing alternativo y refuerza su interés como diana molecular para futuras
62 investigaciones en control genético de vectores.

63 **Palabras clave:** análisis *in silico* – bioseguridad – control vectorial – determinación
64 sexual – malaria – splicing alternativo

65

66 **INTRODUCCIÓN**

67 La malaria persiste como una de las enfermedades infecciosas con mayor impacto en la
68 salud pública global, vinculada estrechamente a la biología de sus vectores. *Anopheles*
69 *gambiae* Giles, 1902 (Arthropoda: Diptera: Culicidae) es el principal transmisor de
70 *Plasmodium falciparum* (Welch 1897) (Apicomplexa: Haemosporida: Plasmodiidae) en

71 África subsahariana, región que soporta la mayor carga de esta parasitosis (World Health
72 Organization, 2025). Esta situación ha impulsado el estudio de los mecanismos
73 biológicos que regulan su desarrollo, destacando la determinación sexual por su
74 relevancia operativa: solo las hembras transmiten el patógeno. En *A. gambiae sensu*
75 *stricto*, este proceso depende de mecanismos regulatorios precisos donde el gen
76 *doublesex (dsx)* actúa como regulador terminal, codificando un factor de transcripción
77 que controla la formación de caracteres sexuales específicos (Wexler *et al.*, 2019;
78 Chikami *et al.*, 2022).

79 Desde la salud pública, los blancos moleculares que permiten sesgar la proporción
80 sexual o reducir la fertilidad femenina son clave para complementar las estrategias
81 convencionales de control, especialmente ante la creciente resistencia a insecticidas.
82 (Beier *et al.*, 2008; World Health Organization, 2021; Messenger *et al.*, 2023). La función
83 de *dsx* reside en el *splicing* alternativo, permitiendo que un único locus genere isoformas
84 con funciones distintas según el sexo (Price *et al.*, 2015). Aunque la caracterización de
85 estas isoformas es vital para entender la biología reproductiva del vector, las limitaciones
86 técnicas del análisis experimental dificultan su exploración sistemática.

87 Históricamente, la introducción accidental de *A. gambiae* s.s. en Brasil en la década de
88 1930 demostró que un vector altamente competente puede alterar drásticamente el
89 riesgo de transmisión si no existen sistemas de vigilancia y gobernanza sólidos (Killeen
90 *et al.*, 2002). Aunque actualmente ausente en las Américas, la malaria persiste mediante
91 vectores como *Anopheles darlingi* Root, 1926 y *Anopheles albimanus* Wiedemann, 1820
92 (Sinka *et al.*, 2010). La conservación de la vía de determinación sexual sugiere la
93 existencia de ortólogos de *dsx* en vectores americanos, cuya caracterización podría

94 orientar futuras investigaciones y discusiones regulatorias bajo marcos de evaluación de
95 riesgos (Krzywinska *et al.*, 2021).

96 En este escenario, los enfoques *in silico* permiten analizar la arquitectura génica e
97 identificar la plausibilidad funcional de variantes del *splicing*. El presente estudio realiza
98 un análisis computacional de las isoformas de *dsx* en *A. gambiae* s.s. para evaluar sus
99 diferencias estructurales y discutir sus implicaciones en bioseguridad y control genético
100 orientado a reducir la transmisión de la malaria.

101

102 MATERIAL Y MÉTODOS

103 Se realizó un estudio observacional, descriptivo y comparativo de tipo *in silico* sobre las
104 isoformas del gen *dsx* (AGAP004050) en *A. gambiae* s.s., utilizando anotaciones génicas
105 y secuencias CDS y proteicas de VectorBase. Se analizaron las isoformas canónicas RA
106 (macho) y RB (hembra) mediante scripts en Python 3.12.3 para reconstruir su
107 arquitectura génica y evaluar exones, segmentos codificantes, longitud del CDS y
108 productos proteicos, con anotación de dominios en UniProt. De forma exploratoria, se
109 generaron variantes teóricas de RB por truncaciones en los extremos 5' y 3' del CDS,
110 realizadas en múltiplos de tres nucleótidos para conservar el marco de lectura. Por
111 tratarse de un análisis descriptivo exploratorio, no se aplicaron pruebas estadísticas
112 inferenciales.

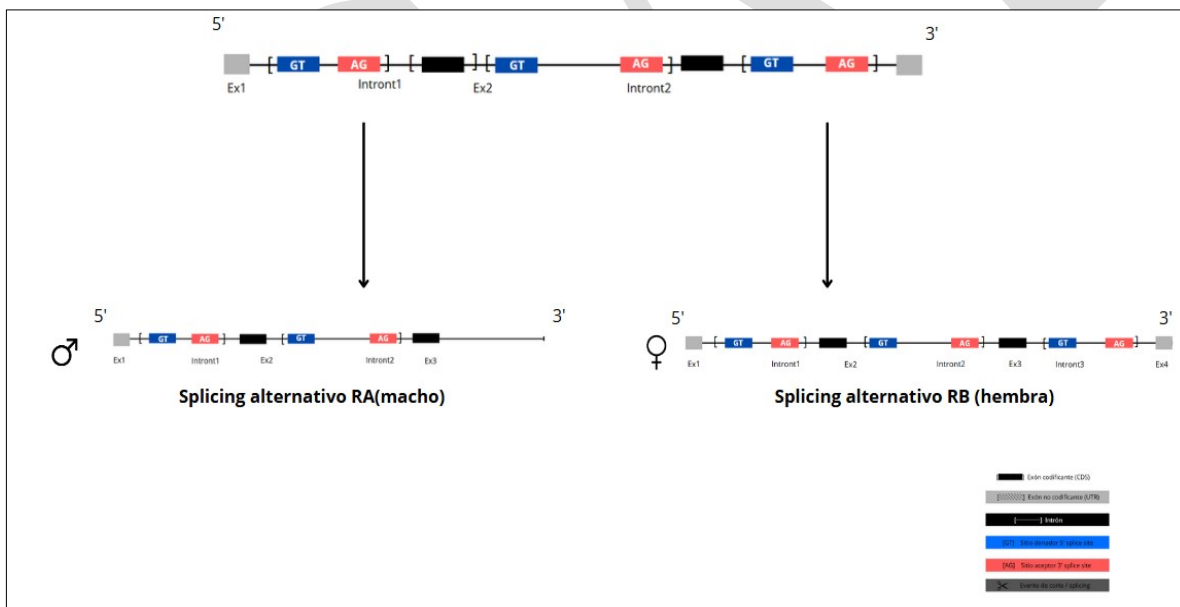
113 **Aspectos éticos:** Se basan en su naturaleza *in silico*. Dado que no se trabajó con sujetos
114 humanos ni con animales, no fue necesario obtener aprobación de un comité de ética.
115 No obstante, el estudio se desarrolló cumpliendo estrictamente con las normativas

116 nacionales e internacionales vigentes en investigación científica, garantizando la
117 integridad y transparencia de los procedimientos.

118

RESULTADOS

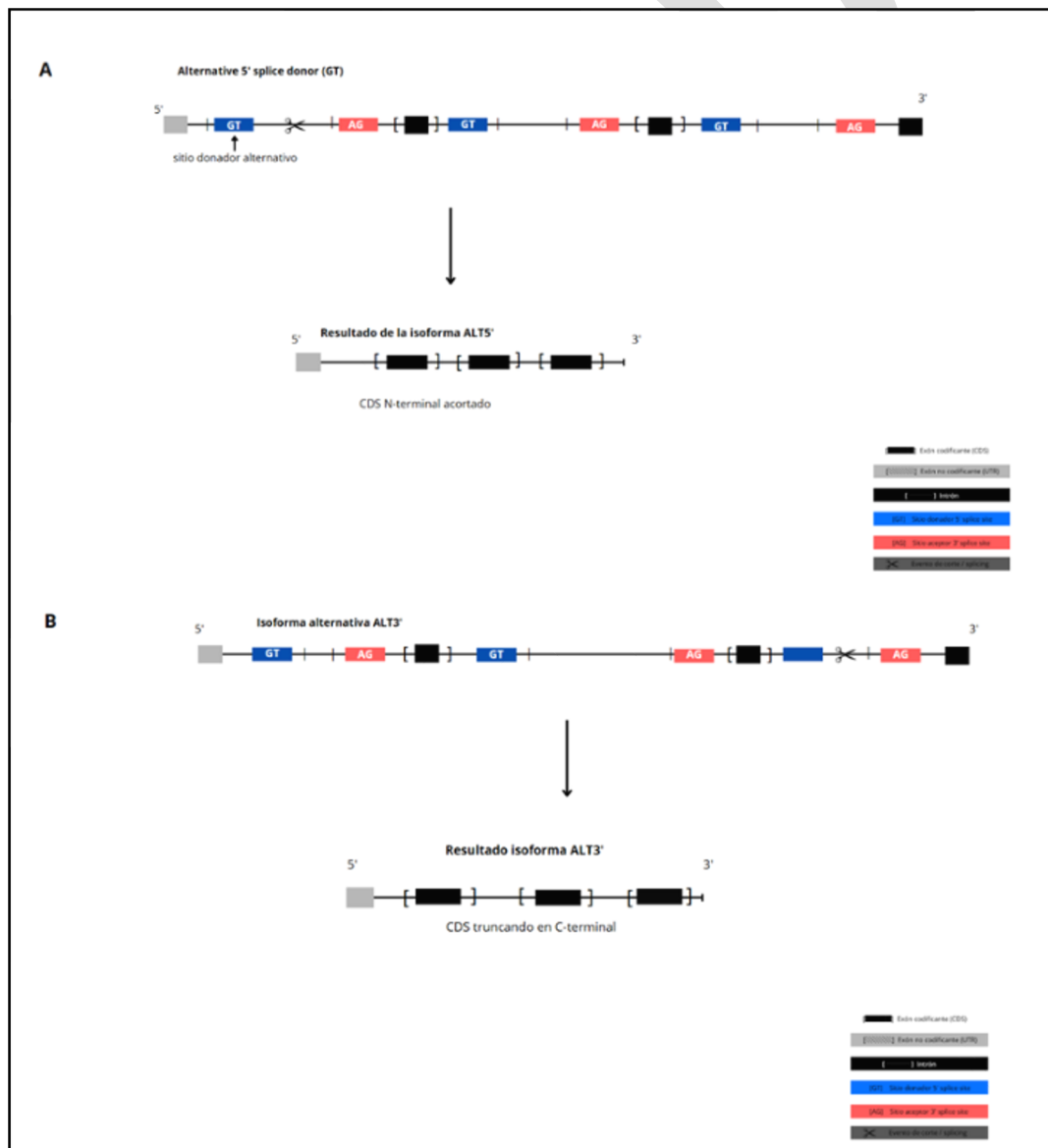
119 A partir del modelo génico de AGAP004050 (*dsx*), se identificaron dos isoformas
120 principales, AGAP004050-RA y AGAP004050-RB, ambas localizadas en la hebra
121 negativa (strand -). La isoforma RA presenta 6 exones y 4 segmentos CDS, con una
122 longitud codificante de 1902 nt, mientras que RB contiene 7 exones y 4 segmentos CDS,
123 con 798 nt. Estas diferencias indican que ambas isoformas presentan arquitecturas
124 génicas distintas que impactan directamente la extensión del producto proteico (Figura
125 1).



126

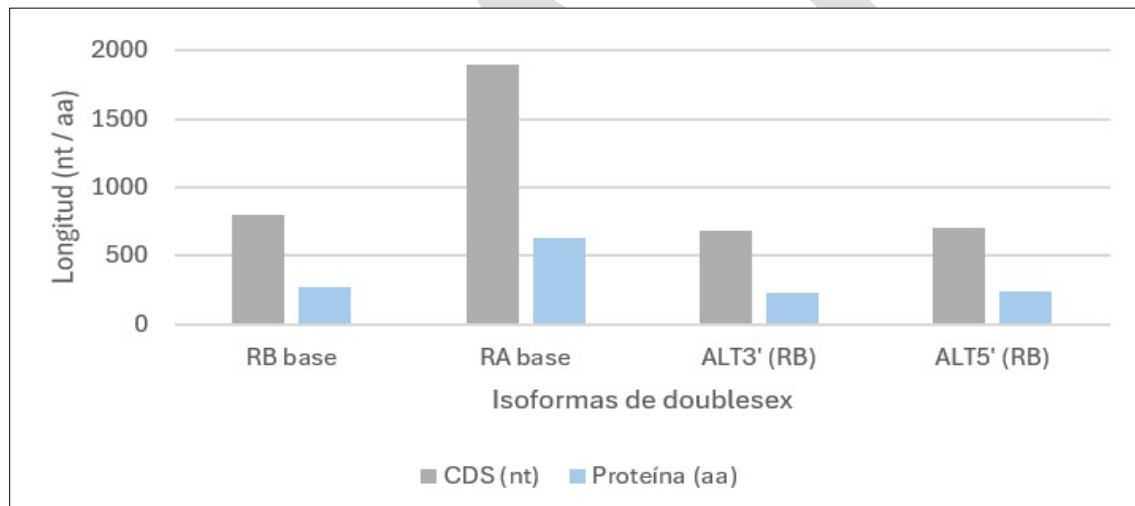
127 **Figura 1.** Arquitectura génica del gen *dsx* (AGAP004050) y generación de isoformas
128 sexo-específicas por *splicing* alternativo en *A. gambiae sensu stricto*.

129 La concordancia entre las anotaciones estructurales del GFF y las secuencias CDS y
130 proteicas en FASTA se confirmó para ambas isoformas: RA (1902 nt; 634 aa) y RB (798
131 nt; 266 aa), ambas con marco de lectura íntegro y sin codones de parada internos.
132 Con fines exploratorios, se generaron variantes teóricas de RB mediante truncaciones
133 controladas, conservando el marco de lectura: ALT3'-B (678 nt; 226 aa) y ALT5'-A' (708
134 nt; 236 aa). Estas variantes no representan transcritos biológicos observados (Figura 2).



151 **Figura 2.** Variantes teóricas del gen *dsx* generadas mediante simulación de eventos de
152 splicing alternativo. (A). Variante derivada de la modificación del extremo 5' del CDS. (B)
153 Variante derivada de la modificación del extremo 3' del CDS.

154 La comparación cuantitativa entre las isoformas analizadas evidenció diferencias
155 progresivas en longitud a nivel de CDS y proteína. El orden fue: RA base (1902 nt; 634
156 aa) > RB base (798 nt; 266 aa) > ALT5'-A' (708 nt; 236 aa) > ALT3'-B (678 nt; 226 aa).
157 Estos resultados indican que las variantes teóricas afectan principalmente regiones
158 terminales del CDS sin comprometer la integridad del marco de lectura.



159 **Figura 3.** Comparación de las longitudes del CDS y de las proteínas predichas de las
160 isoformas RA y RB del gen *dsx* y de las variantes teóricas ALT5'- A' y ALT3'-B derivadas
161 de RB.
162 de RB.

163 Se compararon las longitudes del CDS y de los productos proteicos predichos de las
164 isoformas RA y RB del gen *doublesex* (*dsx*, AGAP004050), así como de las variantes
165 teóricas ALT5' y ALT3' derivadas de RB. Se observaron diferencias en la extensión

166 codificante entre RA y RB, y las variantes truncadas conservaron un marco de lectura
167 íntegro, generando proteínas más cortas sin introducir cambios de fase (Figura 3).

168 **DISCUSIÓN**

169 La identificación de las isoformas AGAP004050-RA y RB en *A. gambiae* s.s. concuerda
170 con el modelo clásico de determinación sexual, donde el gen *dsx* actúa como regulador
171 terminal mediante *splicing* alternativo específico del sexo. Las diferencias en el número
172 de exones y la longitud de la región codificante (1902 nt para RA y 798 nt para RB)
173 demuestran arquitecturas transcripcionales distintas que impactan la composición del
174 producto proteico. Este patrón es consistente con lo descrito previamente para la
175 especie, donde los transcritos sexo-específicos conservan regiones centrales, pero
176 difieren en segmentos terminales responsables de funciones regulatorias (Scali *et al.*,
177 2005). Asimismo, la correspondencia entre modelos génicos y productos traducidos
178 confirmó marcos de lectura íntegros, evidenciando que la ausencia de un codón stop
179 terminal responde a la convención de los archivos CDS FASTA y no a productos
180 truncados, lo que refuerza la robustez del análisis. Desde una perspectiva funcional, la
181 isoforma RB (femenina) presenta una organización coherente con la variante DSX-F. En
182 *A. gambiae*, la integridad de esta isoforma es crítica para la diferenciación y fertilidad;
183 por ello, alteraciones específicas en sus regiones codificantes pueden generar efectos
184 poblacionales significativos (Kyrou *et al.*, 2018). Las variantes generadas mediante
185 truncaciones controladas sugieren que las regiones terminales modulan la función del
186 factor de transcripción sin comprometer su continuidad traduccional. Aunque son
187 predicciones *in silico*, estas variantes respetan el principio biológico de generar
188 diversidad funcional mediante la variación de extremos codificantes mientras se

189 preservan los dominios centrales de unión al ADN y dimerización. En relación con la
190 salud pública, el control vectorial mediante ITNs sigue siendo un pilar fundamental,
191 complementado por herramientas como los cebos ATSB que reducen la densidad de
192 vectores mediante alimentación inducida (Traore *et al.*, 2020; World Health Organization,
193 2025). Estrategias genéticas como las de Target Malaria buscan reducir la transmisión
194 mediante la liberación de mosquitos modificados, con o sin *gene drives* (Target Malaria,
195 s. f.). Dado que solo las hembras transmiten patógenos, la alteración de *dsx* es una línea
196 de interés clave para reducir la proporción de hembras funcionales. Sin embargo, esta
197 manipulación implica riesgos de bioseguridad: su alta sensibilidad funcional facilita una
198 rápida reducción poblacional, pero también supone un riesgo de dispersión no
199 intencionada y difícilmente reversible en poblaciones naturales (Kyrou *et al.*, 2018). Pese
200 a la baja variabilidad observada en sitios de splicing clave (intrón 4/exón 5), la detección
201 de SNPs sugiere que podrían surgir mecanismos de resistencia genética que afecten la
202 estabilidad de estas intervenciones (Kientega *et al.*, 2024). Adicionalmente, la supresión
203 de poblaciones de *A. gambiae* s.s. podría generar efectos ecológicos indirectos, como
204 cambios en la competencia interespecífica o alteraciones en redes tróficas locales
205 (Kormos *et al.*, 2023). Por tanto, el potencial de *dsx* debe valorarse junto a sus riesgos
206 genéticos y regulatorios (Connolly *et al.*, 2023; James & Santos, 2023) bajo marcos éticos
207 sólidos y alineados con la capacidad regulatoria local (Diéguez-Fernández *et al.*, 2021),
208 tal como se ha discutido en la República Dominicana para el control de *Aedes aegypti*
209 (Linnaeus, 1762) (Colomé-Hidalgo & Vallejo Degaudenzi, 2024). Finalmente, al ser un
210 análisis computacional, estos hallazgos deben interpretarse como una base conceptual

211 que requerirá validación funcional y evaluación en condiciones de campo (Frieß *et al.*,
212 2023).

213 El análisis *in silico* del gen *dsx* (AGAP004050) en *A. gambiae* s.s. permitió caracterizar
214 comparativamente las isoformas canónicas RA y RB, evidenciando diferencias en su
215 arquitectura génica y en la extensión de sus regiones codificantes. Estos hallazgos
216 indican que la diversidad funcional de *dsx* depende de una organización transcripcional
217 sexo-específica que impacta directamente el producto proteico.

218 La concordancia entre los modelos génicos y las secuencias traducidas respaldó la
219 integridad del marco de lectura de ambas isoformas, apoyando la confiabilidad del
220 enfoque computacional utilizado. Además, la generación exploratoria de variantes
221 teóricas de RB sugirió que modificaciones terminales pueden originar proteínas más
222 cortas sin alterar dicho marco.

223 En conjunto, los resultados refuerzan la relevancia de *dsx* como regulador central de la
224 determinación sexual y como posible diana molecular para estrategias de control
225 genético del vector. No obstante, por tratarse de un estudio exclusivamente
226 computacional, estas observaciones deben considerarse preliminares y requieren
227 validación experimental y transcriptómica futura.

228

229 **Author contributions: CRediT (Contributor Roles Taxonomy)**

230 **MB** = Mabel Serrano

231 **CM** = Cheryl Marte

232 **NG** = Nasla Gomez

233 **MCH** = Manuel Colomé-Hidalgo

- 234 **EF** = Edian Franco
- 235 **PMAE** = Pedro María Alarcón-Elbal
- 236 **AVD** = Alejandro Vallejo-Degaudenzi
- 237
- 238 **Conceptualization:** MB, AVD
- 239 **Data curation:** CM
- 240 **Formal Analysis:** MB, CM
- 241 **Funding acquisition:** MB, CM, NG, MCH, EF, PMAE, AVD
- 242 **Investigation:** NG, MCH
- 243 **Methodology:** MS, AVD
- 244 **Project administration:** AVD
- 245 **Resources:** PMAE, AVD
- 246 **Software:** MS, CM
- 247 **Supervision:** EF, AVD
- 248 **Validation:** NG, MCH
- 249 **Visualization:** MSM, AVD
- 250 **Writing – original draft:** MS, AVD
- 251 **Writing – review & editing:** CM, EF
- 252

253 **REFERENCIAS BIBLIOGRÁFICAS**

- 254 Beier, J. C., Keating, J., Githure, J. I., Macdonald, M. B., Impoinvil, D. E., & Novak, R. J.
255 (2008). Integrated vector management for malaria control. *Malaria Journal*,
256 7(Suppl 1), S4.

257 Chikami, Y., Kawamoto, M., Toyoda, A., Katsuma, S., & Abe, H. (2022). Evolutionary
258 history of sexual differentiation mechanisms in insects. *Molecular Biology and*
259 *Evolution*, 39(7), msac145.

260 Colomé-Hidalgo, M., & Vallejo Degaudenzi, A. (2024). Uso de mosquitos genéticamente
261 modificados para el control del dengue en la República Dominicana: análisis ético
262 y regulatorio. *Ciencia, Ambiente y Clima*, 7, 11–29.

263 Connolly, J. B., Romeis, J., Devos, Y., Glandorf, D. C. M., Turner, G., & Coulibaly, M. B.
264 (2023). Gene drive in species complexes: Defining target organisms. *Trends in*
265 *Biotechnology*, 41, 154–164.

266 Diéguez-Fernández, L., Monzón-Muñoz, M. V., Juárez-Sandoval, J. A., Barrios-Barrios,
267 D. Y., Barrientos-Juárez, M.E., Barrios, Y.H., Rodríguez-Sosa, M. A., Vásquez-
268 Bautista, Y. E., & Alarcón-Elbal, P. M. (2021). Caracterización de la infestación de
269 viviendas por tres especies de mosquitos (Diptera: Culicidae) de importancia
270 médica en países de bajos y medios ingresos: recomendaciones para su control
271 doméstico. *Ciencia y Salud*, 5, 49–66.

272 Frieß, J. L., Lalyer, C. R., Giese, B., Simon, S., & Otto, M. (2023). Review of gene drive
273 modelling and implications for risk assessment of gene drive organisms.
274 *Ecological Modelling*, 478, 110285.

275 James, S., & Santos, M. (2023). The promise and challenge of genetic biocontrol
276 approaches for malaria elimination. *Tropical Medicine and Infectious Disease*,
277 8(4), 201.

278 Kientega, M., Morianou, I., Traoré, N., Kranjc, N., Kaboré, H., Zongo, O. N., Millogo, A.
279 A., Epopa, P. S., Yao, F. A., Belem, A. M. G., Burt, A., & Diabaté, A. (2024).
280 Genomic analyses revealed low genetic variation in the intron-exon boundary of
281 the *doublesex* gene within the natural populations of *Anopheles gambiae* s.l. in
282 Burkina Faso. *BMC Genomics*, *25*, 1207.

283 Killeen, G. F., Fillinger, U., Kiche, I., Gouagna, L. C., & Knols, B. G. (2002). Eradication
284 of *Anopheles gambiae* from Brazil: lessons for malaria control in Africa? *The*
285 *Lancet Infectious Diseases*, *2*, 618–627.

286 Kormos, A., Dimopoulos, G., Bier, E., Lanzaro, G. C., Marshall, J. M., & James, A. A.
287 (2023). Conceptual risk assessment of mosquito population modification gene-
288 drive systems to control malaria transmission: Preliminary hazards list workshops.
289 *Frontiers in Bioengineering and Biotechnology*, *11*, 126112.

290 Krzywinska, E., Ferretti, L., Li, J., Li, J. C., Chen, C. H., & Krzywinski, J. (2021).
291 Femaleless Controls Sex Determination and Dosage Compensation Pathways in
292 Females of *Anopheles* Mosquitoes. *Current Biology*, *31*, 1084–1091.

293 Kyrou, K., Hammond, A. M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A. K., Nolan, T., &
294 Crisanti, A. (2018). A CRISPR–Cas9 gene drive targeting *doublesex* causes
295 complete population suppression in caged *Anopheles gambiae* mosquitoes.
296 *Nature Biotechnology*, *36*, 1062–1066.

297 Messenger, L. A., Furnival-Adams, J., Chan, K., Pelloquin, B., Paris, L., & Rowland, M.
298 (2023). Vector control for malaria prevention during humanitarian emergencies: A
299 systematic review and meta-analysis. *The Lancet Global Health*, *11*, e534–e545.

300 Price, D. C., Egizi, A., & Fonseca, D. M. (2015). The ubiquity and ancestry of insect
301 *doublesex*. *Scientific Reports*, 5, 13068.

302 Scali, C., Catteruccia, F., Li, Q., & Crisanti, A. (2005). Identification of sex-specific
303 transcripts of the *Anopheles gambiae doublesex* gene. *Journal of Experimental*
304 *Biology*, 208, 3701–3709.

305 Sinka, M. E., Rubio-Palis, Y., Manguin, S., Patil, A. P., Temperley, W. H., Gething, P. W.,
306 Van Boeckel, T., Kabaria, C. W., Harbach, R. E., & Hay, S. I. (2010). The dominant
307 *Anopheles* vectors of human malaria in the Americas: occurrence data,
308 distribution maps and bionomic précis. *Parasites & Vectors*, 3, 72.

309 Target Malaria. (s.f.). *A Vector Control Research Alliance*. <https://targetmalaria.org/>

310 Traore, M. M., Junnila, A., Traore, S. F., Doumbia, S., Revay, E. E., Kravchenko, V. D.,
311 Schlein, Y., Arheart, K. L., Gergely, P., Xue, R. D., Hausmann, A., Beck, R.,
312 Prozorov, A., Diarra, R. A., Kone, A. S., Majambere, S., Bradley, J., Vontas, J.,
313 Beier, J. C., & Müller, G. C. (2020). Large-scale field trial of attractive toxic sugar
314 baits (ATSB) for the control of malaria vector mosquitoes in Mali, West Africa.
315 *Malaria Journal*, 19, 72.

316 Wexler, J. R., Plachetzki, D. C., & Kopp, A. (2019). Hemimetabolous insects elucidate
317 the origin of sexual differentiation mechanisms. *eLife*, 8, e47490.

318 World Health Organization. (2021). *Global technical strategy for malaria 2016–2030,*
319 *2021 update*. Geneva: World Health Organization.
320 <https://www.who.int/publications/i/item/9789240031357>

321 World Health Organization. (2025). *World malaria report 2025: Addressing the threat of*
322 *antimalarial drug resistance*. Geneva: World Health Organization.
323 <https://www.who.int/publications/i/item/9789240117822>

324 Received May 5, 2026.

325 Accepted May 29, 2026.

326

ASAP