

# GAIA

## LA REVOLUCIÓN DEL BIG DATA LLEGA A LA GALAXIA

MERCÈ ROMERO-GÓMEZ, LOLA BALAGUER  
NÚÑEZ Y JORDI PORTELL DE MORA

EL SEGUNDO CATÁLOGO DE LA MISIÓN GAIA DE LA ESA HA REVOLUCIONADO TANTO LA FORMA DE HACER ASTRONOMÍA POR LA PRECISIÓN Y EL VOLUMEN DE SUS DATOS, COMO LA VISIÓN ACTUAL DE NUESTRA GALAXIA.

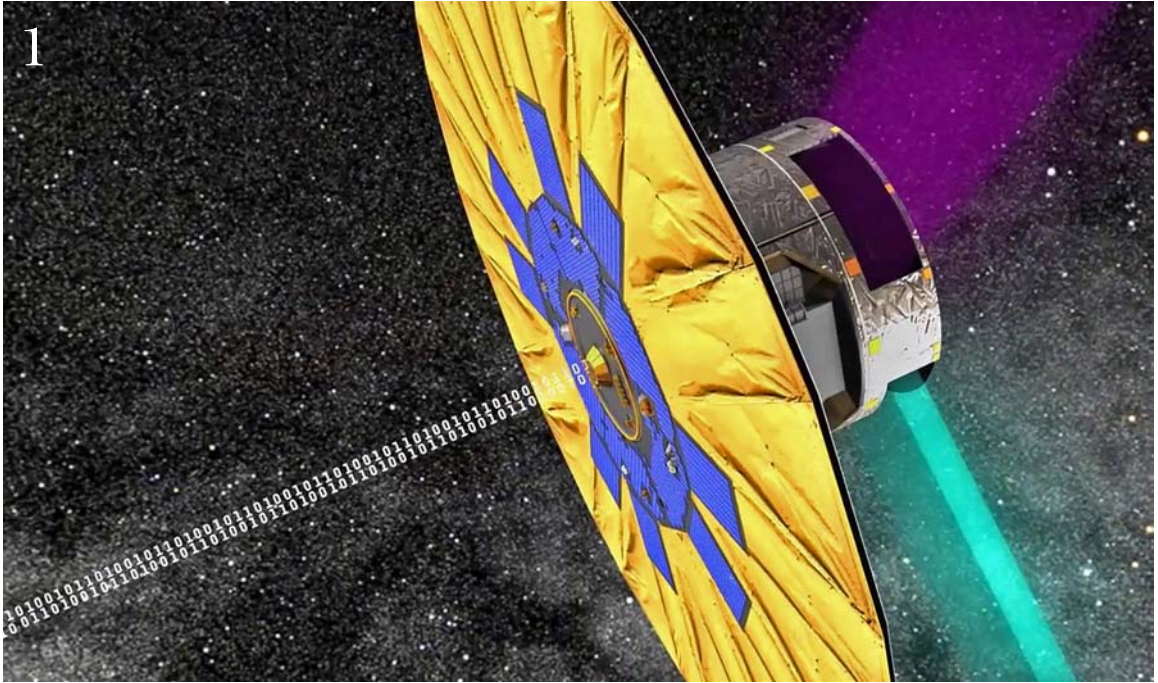


Artículo exclusivo  
colaboración  
de la Sociedad  
Española  
de Astronomía

El estudio de nuestra Galaxia se encuentra actualmente en una posición similar a la geografía terrestre en el siglo XV: grandes partes de la Tierra eran desconocidas para los científicos de la época y solo existían mapas rudimentarios de la mayoría de las partes conocidas

del planeta. Hoy en día, grandes partes de nuestra Vía Láctea están escondidas bajo densas capas de polvo, pero estamos empezando a descubrir y trazar mapas mucho más completos y precisos gracias a proyectos a gran escala como la misión espacial Gaia.

Gaia es una ambiciosa misión de la Agencia Espacial Europea (ESA en sus siglas en inglés) que está realizando el mapa tridimensional de nuestra Galaxia para revelar así su composición, formación y evolución. Gaia mide posiciones y movimientos de



**FIGURA 1.** Representación artística del satélite Gaia, observando continuamente la Galaxia con sus dos telescopios y descargando los datos a las estaciones terrestres. (ESA/Gaia)

las estrellas con una precisión extraordinaria para un censo de alrededor de 1700 millones de objetos. Esto significa un uno por ciento del total de las estrellas de nuestra Galaxia. La anterior misión, *Hipparcos*, que ya fue un hito sin precedentes, estudió

100 000 estrellas con una precisión mucho menor.

Tras cinco años escaneando el cielo, el satélite Gaia completó su misión nominal el 16 de julio de este año, iniciando así la fase de extensión de operaciones científicas que está tentati-

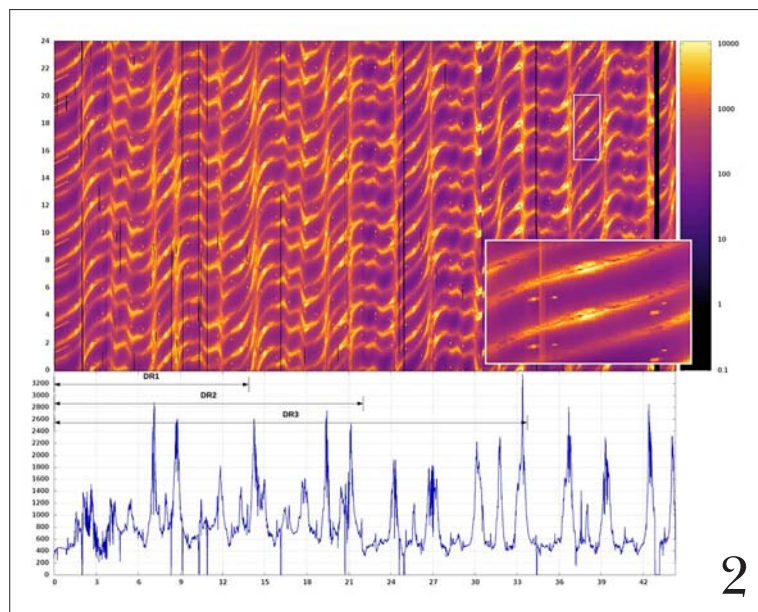
vamente aprobada por la ESA hasta finales de 2022. La misión podría alargarse hasta finales de 2024 gracias a la cantidad de combustible que aún conserva, lo que supondría duplicar la vida del satélite respecto a los cinco años originales y aumentar

así enormemente la precisión de sus resultados.

De momento ya tenemos disponibles los resultados de los primeros veintidós meses de misión que han supuesto una revolución en todos los campos de la astrofísica. Desde la publicación del segundo catálogo de Gaia (DR2) en abril de 2018, los usuarios del archivo han hecho más de 62 millones de búsquedas de datos con más de 100 000 Gigabytes descargados en el primer mes, todo un récord para los archivos de la ESA. Gaia ha producido ya más de 1900 publicaciones científicas de alto nivel, de las que más de 200 son de investigadores en centros españoles.

El éxito de los catálogos DR1 y DR2 de la misión Gaia (demostrado por su repercusión internacional y el número de publicaciones, presentaciones a congresos y medios de comunicación) ha hecho realidad la previsión de que Gaia iba a cambiar la forma de entender las estrellas y la Vía Láctea, así como su impacto en todos los ámbitos de la astrofísica.

La comunidad de astrónomos de todos los niveles se encuentra ante el reto de obtener información del gran volumen de datos fundamentales de alta precisión y sin sesgos observacionales para abordar el estudio de la estructura, cinemática y dinámica de nuestra Galaxia y de objetos situados en y fuera de ella, así como para la determinación precisa de los parámetros físicos de más de mil millones de estrellas y centenares de miles de objetos en el Sistema Solar. A nivel nacional, doscientos astrónomos españoles de más de treinta ins-



tituciones reunidos en la Red Española de explotación científica de Gaia (REG) ya se están coordinando para hacer frente al desafío que esta ingente cantidad de datos supone.

### LA REVOLUCIÓN DEL BIG DATA

Para generar el segundo catálogo de Gaia, con unos 1693 millones de estrellas, se usaron datos adquiridos durante veintidós meses de misión. Esto corresponde a 525 mil millones de observaciones astrométricas (pequeñas imágenes monocromo de hasta 18 por 12 píxeles), 105 mil millones de observaciones espectro-fotométricas, y 9 mil millones de observaciones espectroscópicas (del orden de mil veces el número de espectros tomados desde tierra en toda la historia). Todos estos datos se procesaron en distintos centros europeos, incluyendo ESAC (cerca de Madrid), el Centro Nacional de Supercomputación (el *MareNostrum* en Barcelona), el CNES (en Toulouse, Francia), el Institu-

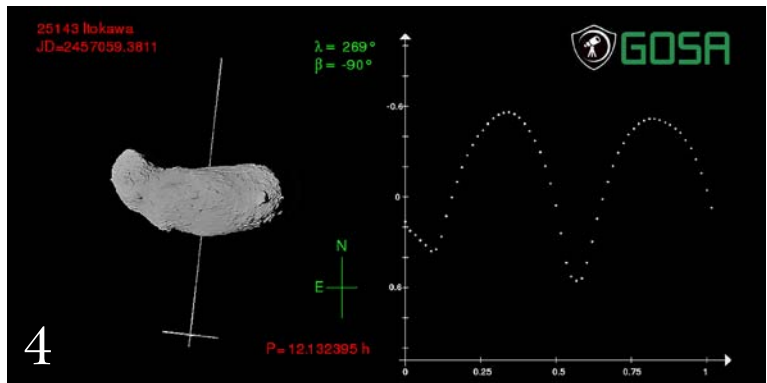
to de Astronomía de Cambridge (Inglaterra), el Observatorio Astronómico de Ginebra (Suiza), y el Observatorio Astronómico de Torino (Italia). Además, contamos con centros de computación adicionales, como por ejemplo el CESGA, donde el Grupo Gaia de Galicia implementa y ejecuta un algoritmo de redes neuronales para realizar una clasificación no supervisada de las estrellas de la Vía Láctea y detectar automáticamente casos atípicos. Todos estos centros, junto con las unidades que desarrollan los algoritmos y el software en sí, componen el consorcio para el procesado y análisis de datos de Gaia (DPAC, por sus siglas en inglés).

El grupo Gaia de Barcelona es responsable, entre otras cosas, del desarrollo y operación de uno de los sistemas más exigentes en DPAC, llamado IDU (siglas en inglés de Actualización de los Datos Intermedios). Para procesar los datos finales para DR2, este sistema usó

**FIGURA 2.** Ilustración de la densidad de estrellas observadas por unidad de tiempo a lo largo de la misión. La imagen superior muestra, con distintos colores, el número de estrellas por segundo en función del día [eje horizontal] y hora [eje vertical]. La imagen pequeña ilustra cómo esta representación revela la estructura galáctica. En la imagen inferior, rangos temporales de las distintas publicaciones [eje horizontal y flechas], y densidad de estrellas por segundo [eje vertical]. [ESA/Gaia/DPAC]

**FIGURA 3.** Supercomputador *MareNostrum* en el Barcelona Supercomputing Center. [BSC]

**FIGURA 4.** Asteroide Itokawa con su curva de luz. [Gaia/GOSA]



más de mil doscientos procesadores del supercomputador *MareNostrum* durante unas dos semanas. Para cada nueva publicación, IDU debe procesar todos los datos en bruto acumulados hasta la fecha, determinando los niveles de sesgo electrónico del instrumento, el fondo de cielo, la respuesta óptica y electrónica del instrumento, y las posiciones, brillos y calidad de cada una de los miles de millones de medidas. Así mismo, IDU se encarga del proceso llamado «Cross-Match», consistente en identificar las observaciones repetidas para cada estrella. Para el DR2, IDU recibió una lista de 52 mil millones de tránsitos estelares (cada uno con varias observaciones astrométricas, espectro-fotométricas y espectroscópicas), a partir de la cual determinó las características de cerca de dos mil millones de estrellas. Este sofisticado proceso requiere de complejos algoritmos de agrupación (o «clustering») y de limpieza de

detecciones falsas (causadas por efectos instrumentales).

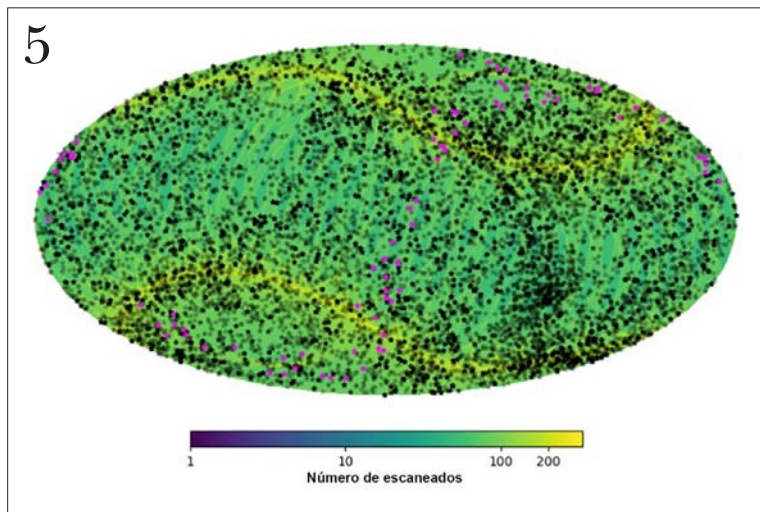
El consorcio DPAC lleva cerca de dos años trabajando en la preparación del tercer catálogo de Gaia (DR3) que se publicará en 2021, para el que se usarán unos 780 mil millones de observaciones astrométricas, 156 mil millones espectro-fotométricas y 14 mil millones espectroscópicas, adquiridas durante treinta y cuatro meses de misión.

Y ya estamos trabajando en la preparación de la cuarta publicación (DR4), que contará con datos adquiridos durante cinco años y medio de misión, y más de un billón y medio de observaciones astrométricas. El volu-

men de datos descargados desde el satélite hasta el momento es de unos 68 TB (1 TB es un millón de millones de bytes). Para generar el DR3 se están analizando 300 TB de datos, mientras que para DR4 ya llegaremos al PB (mil millones de millones de bytes).

#### CIENCIA CIUDADANA CON GAIA

Todos los datos que proporciona Gaia se publican con acceso abierto en el archivo general de la ESA, [archives.esac.esa.int/gaia](http://archives.esac.esa.int/gaia), que, además, proporciona herramientas de visualización, cruces con otros catálogos, tutoriales para aprender a hacer búsquedas y abundante documentación



**FIGURA 5.** Cobertura en el cielo de las Alertas de Gaia a fecha de 1 de agosto de 2019. En rosa las alertas de los últimos siete días. (Gaia/ESA/DPAC/Cambridge)

**FIGURA 6.** La Vía Láctea vista por Gaia. Se señalan las componentes principales de nuestra Galaxia. (ESA/Gaia)

**FIGURA 7.** Alabeado del disco galáctico, ver texto para más detalles. (Romero-Gómez et al., 2019)

complementaria. Todos los contenidos están a disposición del público en general (eso sí, en inglés), pero hay dos proyectos de Gaia que cuentan especialmente con la colaboración ciudadana: GOSA para el estudio de asteroides, y las Alertas de Gaia:

— Gaia GOSA es un servicio web abierto a toda la comunidad de astrónomos, tanto profesionales como aficionados, para planificar el seguimiento y la observación fotométrica de asteroides en base a las efemérides proporcionadas por Gaia. Los datos recogidos por la comunidad GOSA se usan dentro del consorcio Gaia DPAC para mejorar los resultados de los objetos del Sistema Solar. [www.gaiagosa.eu](http://www.gaiagosa.eu).

— Las «Alertas de Gaia» es un catálogo fotométrico de todo el cielo basado en las medidas repetidas y de alta precisión de Gaia. Esta repetición de medidas es ideal para descubrir variaciones de brillo en el cielo y descubrir fenómenos transitorios: desde nuevas fuentes a variaciones muy grandes en estrellas ya conocidas. Estos datos se hacen públicos inmediatamente (tan-

to las curvas de luz de Gaia como su espectrofotetría) para facilitar que cualquier astrónomo profesional o aficionado (o incluso grupos escolares) puedan participar recogiendo y compartiendo datos. [gsaweb.ast.cam.ac.uk/alerts](http://gsaweb.ast.cam.ac.uk/alerts).

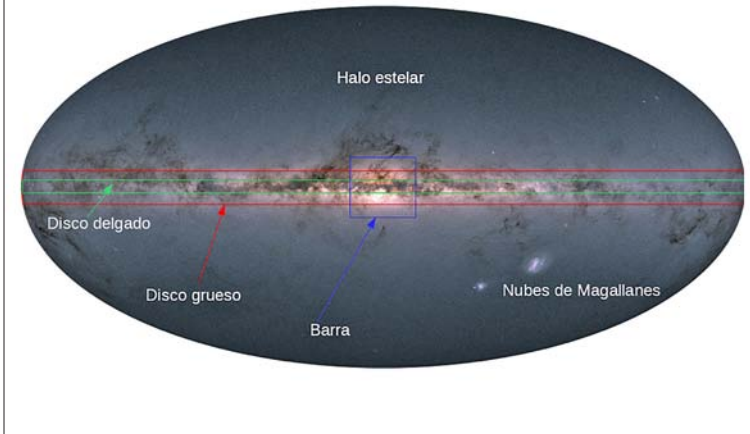
#### REVELANDO LA HISTORIA DE FORMACIÓN DE LA GALAXIA

La forma de ver y entender la Vía Láctea ha cambiado radicalmente desde la publicación del segundo catálogo de datos de Gaia. La ingente cantidad de paralajes y su precisión ha permitido afrontar retos que hasta entonces se tenían que descartar por el hecho de no disponer de la distancia a suficientes estrellas. Uno de los temas que recientemente ha visto un gran avance es el estudio de la tasa o ritmo de formación de estrellas en nuestra Galaxia. Los astrónomos intentamos utilizar datos observados actualmente para deducir el pasado, de ahí viene el término de Arqueología Galáctica. El ritmo de formación de estrellas y su historia de formación estelar son dos de los ingredientes cla-

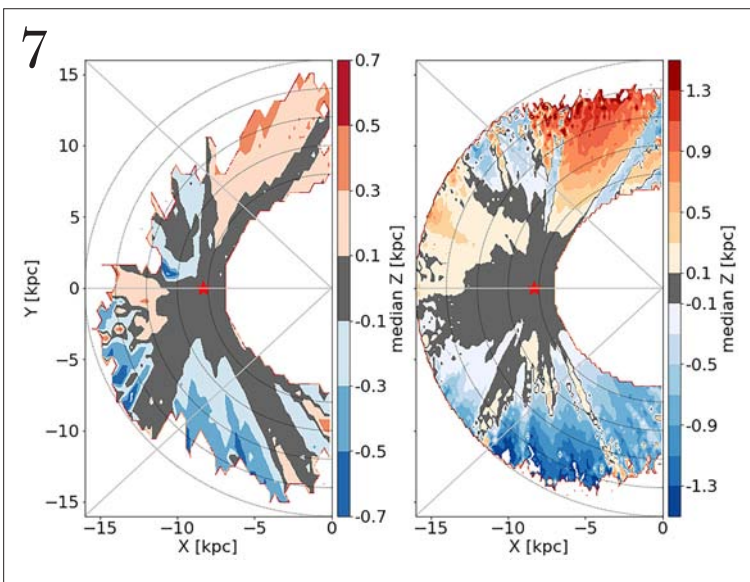
ve para entender el nacimiento y evolución de una galaxia, incluyendo información muy valiosa sobre la fusión y acreción entre otras galaxias y la Vía Láctea. El ritmo de formación estelar nos proporciona la cantidad de masa total de estrellas formadas por año, mientras que la historia de formación estelar desvela cómo se formaron las estrellas a lo largo del tiempo y el espacio. Puede ser que una galaxia se haya formado rápidamente o durante largos periodos de tiempo. Hasta la llegada de los datos de Gaia, se debatía que la Vía Láctea ha sufrido diferentes brotes de formación estelar en el pasado y que la tasa de formación estelar actual es del orden de una masa solar por año. No ha sido hasta tener disponible el segundo catálogo de datos de Gaia que se ha afrontado el reto de desvelar con precisión la historia de formación estelar de nuestra Galaxia.

Cuando dos galaxias se fusionan, este suceso deja una impronta en la velocidad de las estrellas y en su composición química al ser de un origen distinto. El año pasado se realizó un

6



7



estudio dinámico-químico con una selección de estrellas del entorno solar que encontró dos grupos de estrellas con velocidades y química diferentes. Un grupo está compuesto por estrellas que ya formaban parte de nuestra Galaxia primigenia mientras que el que ofrecía una cinemática y química esencialmente diferenciada se atribuye a una galaxia menor que se fusionó con nuestra Galaxia, y a la que se ha bautizado como Gaia-Encélado.

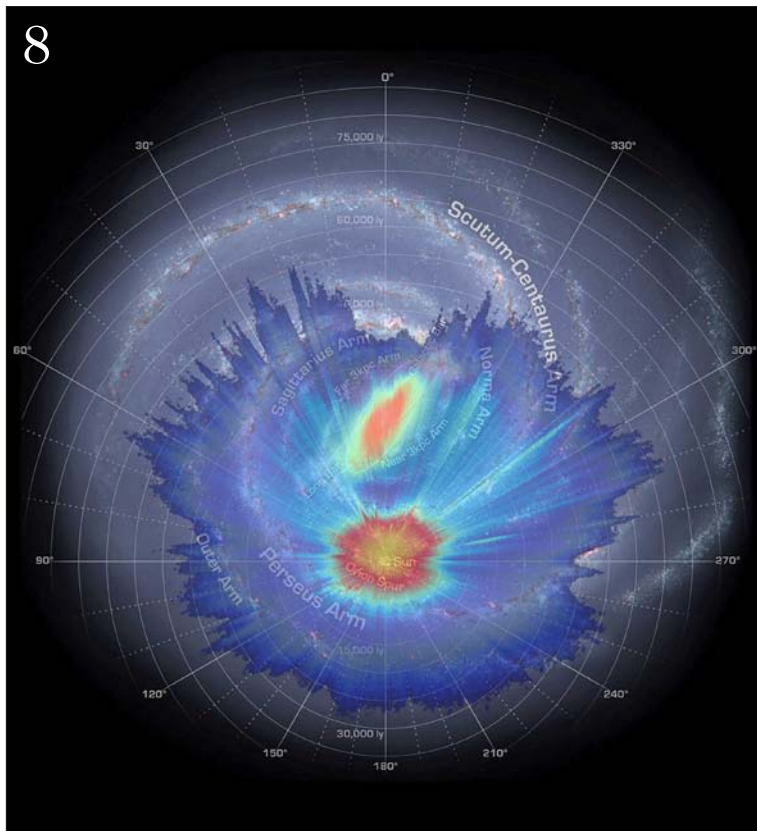
Este estudio daba un rango amplio de cuándo pudo ocurrir esta fusión, hace entre 10 000 y 13 000 millones de años. Recientemente, investigadores españoles liderados por Carme Gallart en el Instituto de Astrofísica de Canarias han sido capaces de fijar una edad más precisa a esta fusión: exactamente hace 10 000 millones de años. Para ello, además de utilizar la cinemática y química de las estrellas han utilizado modelos del estado evolutivo

de una población de estrellas. En los modelos se conocen exactamente la edad de las estrellas. El estudio de Gallart ha sido capaz de obtener simulaciones cosmológicas de tal calidad que se comparan exactamente con los datos de Gaia. Analizando la distribución de edades de las diferentes poblaciones, se ha podido poner fecha al momento en que la galaxia Gaia-Encélado quedó engullida por nuestra Galaxia. Esta fusión hizo que el ritmo de formación de estrellas aumentara de manera significativa hace 9000-10 000 millones de años y luego disminuyese.

Pero ese no ha sido el único revés para nuestra Vía Láctea. Otro estudio liderado por Roger Mor de la Universidad de Barcelona, gracias a los datos de Gaia, ha desvelado que nuestra Galaxia sufrió otro encuentro violento, pero con una compañera mucho menor. No se han descubierto todavía las estrellas que formaban parte de la compañera, pero sí que el ritmo de formación estelar volvió a aumentar significativamente hace unos 6000 millones de años, tuvo su máximo hace 2000-3000 millones de años, donde se cree que se formó la mitad de las estrellas del disco delgado de nuestra Galaxia, y ha ido bajando hasta la actualidad.

#### GAIA REVELA EL ALABEO DEL DISCO DE LA GALAXIA

La cantidad y calidad de los datos de Gaia han permitido establecer que el disco galáctico se deforma en las partes externas de manera que si pudiéramos verlo de perfil tendría una forma de S, asciende por un lado mientras que desciende por



**FIGURA 8.** Mapa de la densidad de estrellas en la Vía Láctea sobre una vista artística de nuestra Galaxia. Muchas de las estrellas en Gaia DR2 se encuentran cerca del Sol (el área roja en la parte inferior de la imagen), pero también se distingue una estructura elongada poblada por muchas estrellas en la parte central de la Vía Láctea: la primera indicación astrométrica de la barra galáctica. (Datos: ESA/Gaia/DPAC, A. Khalatyan -AIP-, F. Anders -ICCUB- y StarHorse team; Imagen de la Galaxia: NASA/JPL-Caltech/R. Hurt -SSC/Caltech-)

### Bibliografía

- «Gaia. Astrometría al microsegundo de arco», *Astronomía* 136, oct. 2010.
- «El primer mapa 3D de la Vía Láctea», *Investigación y Ciencia*, mar. 2019.
- «The merger that led to the formation of the Milky Way's inner stellar halo and thick disk» (Helmi *et al.* *Nature* 563: 7729, p. 85, 2018).
- «Gaia reveals a complex lopsided and twisted Galactic disc warp» (Romero-Gómez *et al.*, *Astronomy & Astrophysics* 627, id. A150 2019).
- «Photo-astrometric distances, extinctions, and astrophysical parameters for Gaia DR2 stars brighter than  $G = 18$ » (Anders *et al.*, *Astronomy & Astrophysics* 628, id. A94 2019).
- «Gaia DR2 reveals a star formation burst in the disc 2-3 Gyr ago» (Mor *et al.*, *Astronomy & Astrophysics* 624, id. L1 2019).
- «Uncovering the birth of the Milky Way through accurate stellar ages with Gaia» (Gallart *et al.*, *Nature Astronomy*, Advanced Online Publication 2019).
- **Para más información:** *Gaia Mission* es una app para Android e iPhone que te permite descubrir los detalles de Gaia desde tu móvil.

el otro. Gaia ha permitido, además, afrontar el reto de estudiar si esta deformación depende de la edad de las estrellas. En la Figura 7, vemos a la izquierda una muestra de estrellas jóvenes tipo OB, y a la derecha una muestra de estrellas gigantes. La altitud respecto al plano galáctico se dibuja en color rojo por encima y azul por debajo del plano. La posición solar está marcada con una estrella roja mientras que el centro galáctico está en el origen de coordenadas.

### GAIA NOS MUESTRA LA BARRA DE LA VÍA LÁCTEA

Combinando los datos de Gaia con otros catálogos fotométricos en el óptico e infrarrojo, se ha podido determinar la distancia con mayor precisión y el en-

rojecimiento debido al polvo a aproximadamente 150 millones de estrellas de la Vía Láctea. Esta estrategia ha permitido cartografiar las regiones más distantes de nuestra Galaxia y visualizar de manera clara la barra en su centro. La Figura 8 muestra la densidad de estrellas donde el origen de coordenadas y la posición solar son los mismos que en la figura anterior. Los datos muestran una sobre-densidad alargada y en forma de barra centrada en el centro galáctico.

El trabajo de los investigadores españoles en el consorcio Gaia DPAC está financiado por el Ministerio de Ciencia, Innovación e Universidades a través de las ayudas ESP2016-80079-C2-R y RTI2018-095076-B-C (MICIU/AEI/FEDER, UE). (A)



Mercè Romero-Gómez, Lola Balaguer-Núñez y Jordi Portell de Mora, son investigadores del Instituto de Ciencias del Cosmos de la Universidad de Barcelona-Instituto de Estudios Espaciales de Catalunya (ICCUB-IEEC), y miembros de la misión Gaia (DPAC).