



**Aplicación de métodos
genómicos para el
seguimiento y evaluación
fauna bentónica de
macroinvertebrados en
aguas de transición y
costeras**

Fundación AZTI Fundazioa

ura

URAREN
EUSKAL
AGENTZIA

AGENCIA
VASCA
DEL AGUA



ELISKO JAURLARITZA
GOBIERNO VASCO

TIPO DE DOCUMENTO:	Informe
TÍTULO DEL DOCUMENTO:	Aplicación de métodos genómicos para el seguimiento y evaluación fauna bentónica de macroinvertebrados en aguas de transición y costeras
ELABORADO POR:	Fundación AZTI Fundazioa
AUTORES:	Ángel Borja, Anders Lanzén, Iñigo Muxika
FECHA:	Diciembre 2023

Índice

Aplicación de métodos genómicos para el seguimiento y evaluación fauna bentónica de macroinvertebrados en aguas de transición y costeras

1. Introducción	4
1.1. Antecedentes	4
1.2. Objetivos	7
2. Métodos	8
2.1. Estaciones de control	8
2.2. Protocolo de muestreo	9
2.3. Protocolo de análisis. Identificación morfológica.....	10
2.4. Protocolo de análisis. Identificación genómica.....	10
2.5. Evaluación de los datos.....	13
3. Resultados	14
3.1. Comparación de variables simples	14
3.2. Comparación de M-AMBI y M-gAMBI	16
3.3. Análisis de similaridad	18
4. Discusión.....	25
4.1. Variables simples	25
4.2. Comparación entre M-AMBI y M-gAMBI.....	27
4.3. Análisis de similaridad	28
5. Conclusiones	29

1.

Introducción

Este informe se presenta como parte de los trabajos realizados en el marco del Convenio de Colaboración entre la Agencia Vasca del Agua y la Fundación AZTI para realizar la “*Ejecución de programas de seguimiento de las aguas de transición y costeras de la Comunidad Autónoma del País Vasco y determinación de su estado. Campañas 2018-2020 y 2021-2023*” suscritos en sendos convenios con fecha de enero de 2018 y 2021.

1.1. ANTECEDENTES

Entre los objetivos de la Directiva Marco del Agua (DMA)¹ está conseguir el buen estado y la adecuada protección de los sistemas acuáticos. Los Planes Hidrológicos deben elaborarse para cada demarcación hidrográfica y deben incluir el conjunto de actuaciones dirigidas a alcanzar dichos objetivos de la DMA. En el caso de las aguas de transición y costeras de la CAPV resulta de aplicación el Plan Hidrológico de la Demarcación Hidrográfica del Cantábrico Oriental².

La DMA establece la necesidad de disponer de programas de seguimiento de las masas de agua que sirvan como base de información sobre el estado de las masas de agua y el grado de cumplimiento de los objetivos ambientales, y para evaluar el grado de ejecución y de efectividad de los programas de medidas que se planteen en el marco de la Planificación Hidrológica.

La Administración Hidráulica de la Comunidad Autónoma del País Vasco lleva desde 1994 obteniendo información relevante sobre el estado de las aguas de transición y costeras de la CAPV. Prácticamente desde su inicio, el planteamiento de control de esta red fue similar a los requerimientos que posteriormente exigió la DMA. El diseño de esta red se concibió con el objetivo de contar con un instrumento imprescindible para llevar a cabo una correcta planificación y gestión del recurso hídrico y éste ha sido el referente que ha ido marcando todas y cada una de las mejoras que se han ido incorporando a las redes de vigilancia para adaptarse a las exigencias de control de la legislación estatal y europea.

De la combinación de obligaciones indicadas anteriormente y en el desarrollo de su competencia, la Agencia Vasca del Agua mantiene el considerable esfuerzo realizado en los últimos años para avanzar tanto en el conocimiento de las aguas continentales, de transición y litorales, como en poner en marcha

¹ Directiva 2000/60/CE del Parlamento Europeo y del Consejo, de 23 de octubre de 2000, por la que se establece un marco comunitario de actuación en el ámbito de la política de aguas, también denominada Directiva Marco del Agua.

² Real Decreto 35/2023, de 4 de enero, por el que se aprueba la revisión de los Planes Hidrológicos de las demarcaciones hidrográficas del Cantábrico Occidental, Guadalquivir, Ceuta, Melilla, Segura y Júcar, y de la parte española de las demarcaciones hidrográficas del Cantábrico Oriental, Miño-Sil, Duero, Tago, Guadiana y Ebro.

mecanismos útiles para su control y vigilancia.

Actualmente el Real Decreto 817/2015³ (en adelante RDSE) es la norma que determina los criterios básicos y homogéneos para el diseño y la implantación de los programas de seguimiento del estado de las masas de agua superficiales y para el control adicional de las zonas protegidas. También establece elementos clave en la evaluación de estado de masas de agua superficial, tales como, las condiciones de referencia y los límites de clases de estado de los indicadores de los elementos de calidad biológicos, fisicoquímicos e hidromorfológicos para clasificar el estado o potencial ecológico; y las normas de calidad ambiental (NCA) para las sustancias prioritarias y para otros contaminantes, así como para las sustancias preferentes.

El artículo 92 bis del Texto Refundido de la Ley del Agua (TRLA)⁴ establece que: para conseguir una adecuada protección de las aguas, se deberán alcanzar determinados **objetivos medioambientales** para las aguas superficiales que se concretan en prevenir el deterioro del estado de las masas de agua superficiales; proteger, mejorar y regenerar todas las masas de agua superficial con el objeto de alcanzar un buen estado de las mismas; y reducir progresivamente la contaminación procedente de sustancias prioritarias y eliminar o suprimir gradualmente los vertidos, las emisiones y las pérdidas de sustancias peligrosas prioritarias.

El cumplimiento de esos objetivos, en el caso de **estado ecológico** de aguas de transición y costeras, implica que los diferentes indicadores del estado basados en determinados elementos de calidad no deben apartarse significativamente de las condiciones naturales, es decir, el grado de distorsión o desviación de las condiciones inalteradas o **condiciones de referencia** debe ser tal que permita la consecución de un buen estado ecológico o un buen potencial ecológico.

Así en el artículo 3 del RDSE se define **buen estado ecológico** como el estado de una masa de agua superficial en la que los valores de los indicadores de los elementos de calidad biológicos correspondientes al tipo de masa de agua superficial muestran valores bajos de distorsión causada por la actividad humana, desviándose ligeramente de los valores normalmente asociados con el tipo de masa de agua superficial en condiciones inalteradas. Los indicadores hidromorfológicos son coherentes con la consecución de dichos valores y los indicadores químicos y fisicoquímicos cumplen con los rangos o límites que garantizan el funcionamiento del ecosistema específico del tipo y la consecución de los valores de los indicadores biológicos. Las concentraciones de los contaminantes específicos cumplen las NCA pertinentes.

En los artículos 12 y 13 del RDSE se indican los elementos de calidad para la clasificación del estado o potencial ecológico para las masas de agua de la categoría aguas de transición y costeras, encontrándose entre los elementos biológicos la “*Composición y abundancia de la fauna bentónica de invertebrados*”.

El RDSE establece que el índice M-AMBI o *multivariate AZTI's Marine Biotic Index* (Muxika *et al.*, 2007) es un indicador para el elemento de calidad fauna bentónica de invertebrados en masas de agua de transición y costeras de aplicación en el caso de masas de agua de transición y costeras de la Demarcación Hidrográfica del Cantábrico Oriental. El índice M-AMBI⁵ incorpora en su cálculo tres

3 Real Decreto 817/2015, de 11 de septiembre, por el que se establecen los criterios de seguimiento y evaluación del estado de las aguas superficiales y las normas de calidad ambiental.

4 Real Decreto Legislativo 1/2001, de 20 de julio, por el que se aprueba el Texto Refundido de la Ley de Aguas

5 URA, 2021. Protocolo de muestreo, análisis y evaluación de la fauna bentónica de macroinvertebrados de sustrato blando en masas de agua de transición y costeras. https://www.uragentzia.euskadi.eus/contenidos/informacion/protocolos_estado_aguas/es_def/adjuntos/01_TW_CW_MACROINVERTEBRADOS_URA_V_2-1.pdf

métricas: el AMBI (AZTI Marine Biotic Index), la riqueza y el índice de diversidad de Shannon.

El índice M-AMBI responde a las definiciones normativas de clasificación del estado indicadas en el anexo V de la Directiva 2000/60/CE, es decir, evalúa la composición y abundancia taxonómica, el cociente entre taxones sensibles a las perturbaciones y taxones insensibles, y el grado de diversidad de taxones y se han completado íntegramente todas las etapas del proceso de intercalibración expuestas en los documentos de guía⁶.

El cálculo del M-AMBI se realiza con un programa (<https://ambi.azti.es>), para cuyo uso se recomienda seguir las guías correspondientes (Borja y Muxika, 2005; Borja et al., 2008, 2012a). Dicho cálculo se basa en el uso de Análisis Factorial, es decir, análisis de las distancias del conjunto de valores que identifican a una estación, respecto a dos condiciones virtuales:

- condiciones virtuales de muy mal estado (situación azoica que se corresponde con $S = 0$; $H' = 0$ ind. bits⁻¹; y AMBI = 6);
- condiciones virtuales de muy buen estado o condiciones de referencia (Tabla 1) establecidas para las masas de agua de transición y costeras de las tipologías presentes en la Demarcación Hidrográfica del Cantábrico Oriental en función de los tramos salinos y de las comunidades dominantes existentes, que fueron calculados mediante análisis multivariante de datos históricos de estaciones poco o nada alteradas, para comunidades dominantes en la costa vasca, modelos y juicio de experto (Muxika et al., 2007).

Tabla 1 Condiciones virtuales de referencia para S (riqueza), H' (diversidad) y AMBI para tipologías de aguas de transición y costeras presentes en la Demarcación Hidrográfica del Cantábrico Oriental (Real Decreto 817/2015).

Categoría	Tipología	Tramo salino/Profundidad	Comunidad asociada	S	H'	AMBI
Aguas de transición	AT-T08, AT-T09, AT-T10	0-18	<i>Cerastoderma edule-Scrobicularia plana</i>	13	2,5	2,8
		18-30	<i>Venus fasciata</i>	32	3,8	2,0
		30-34	<i>Abra alba</i>	40	3,5	2,1
Aguas costeras	AC-T12	20-50 m	<i>Tellina tenuis-Venus fasciata</i>	42	4	1
		70-120 m	<i>Amphiura</i>	130	5,7	1

Los **límites de clase entre estados ecológicos** para aguas de transición y costeras de la Demarcación Hidrográfica del Cantábrico Oriental consideradas como masas de agua naturales según lo indicado en el Real Decreto 817/2015 y derivado del resultado del ejercicio de intercalibración se presentan en la Tabla 2.

Tabla 2 Límites de cambio de clase de M-AMBI para masas naturales expresados como valores del Ratio de Calidad Ecológica (RCE)

Categoría	Tipo	Tramo salino	Límites de cambio de clase (RCE)			
			muy bueno/bueno	bueno/moderado	moderado/deficiente	deficiente/malo
Aguas costeras	AC-T12	>34	≥0,77	≥0,63 ⁷	0,38	0,20
Aguas de transición	AT-T08, AT-T09 y AT-T10	0-18	≥0,77	≥0,53	0,38	0,20
		18-30				
		30-34				

⁶ Decisión (UE) 2018/229 de la Comisión, de 12 de febrero de 2018, por la que se fijan, de conformidad con la Directiva 2000/60/CE del Parlamento Europeo y del Consejo, los valores de las clasificaciones de los sistemas de seguimiento de los Estados miembros a raíz del ejercicio de intercalibración, y por la que se deroga la Decisión 2013/480/UE

⁷ Para la tipología AC-T12 en el Real Decreto 817/2015 establece el límite bueno/moderado en 0,53. En la tabla se indica el valor obtenido posteriormente en el ejercicio de intercalibración de 2018⁶.

1.2. OBJETIVOS

En el marco de los programas de seguimiento de las aguas de transición y costeras de la Demarcación Hidrográfica del Cantábrico Oriental, las comunidades de macroinvertebrados bentónicos se controlan y evalúan conforme a los **estándares morfológicos** recogidos en los protocolos desarrollados por URA⁵. No obstante, esta metodología tradicional basada en estándares morfológicos implica ciertos inconvenientes tales como dificultades en la identificación morfológica, elevada inversión de tiempo por muestra, necesidad de expertos taxónomos y elevado coste de análisis por muestra.

Actualmente se han desarrollado de **métodos genómicos** para el monitoreo y evaluación de comunidades de macroinvertebrados bentónicos que podrían hacer que el coste de análisis se reduzca de manera importante y que los resultados se obtengan más rápidamente.

En relación con el desarrollo de métodos genómicos para el monitoreo y evaluación de comunidades biológicas, en lo que respecta a las comunidades bentónicas, hasta ahora se ha demostrado que el AMBI genómico (gAMBI) proporciona resultados similares a los del AMBI convencional (Aylagas *et al.*, 2014, 2018). Sin embargo, los métodos genómicos proporcionan valores de riqueza y diversidad muy diferentes para el cálculo de M-AMBI, por lo que el uso de un M-AMBI genómico (M-gAMBI) requiere del desarrollo de condiciones de referencia específicas, que, eventualmente, permitieran intercalibrar los resultados con el análisis tradicional basado en estándares morfológicos.

Por todo ello, se propuso a URA estudiar la posibilidad de aplicar dichos métodos genómicos (M-gAMBI) en los programas de seguimiento de las aguas de transición y costeras que ejecuta URA, con el **objetivo** de determinar las condiciones de referencia de M-gAMBI y explorar la posibilidad de intercalibrar en un futuro M-gAMBI con M-AMBI.

El reto en este estudio consiste pues en determinar las condiciones de referencia para un M-gAMBI, que puedan ser equivalentes a las establecidas para M-AMBI en la Tabla 1, y determinar si los límites de clase de la Tabla 2 son válidos o deben establecerse otros específicos.

Los autores queremos agradecer el trabajo al equipo de muestreo de AZTI, por el esfuerzo adicional de toma de muestras extra de réplicas para genómica, a los analistas de laboratorio de Pasaia y Sukarrieta, en especial a Iñaki Mendibil, por la preparación de muestras, y a la Sociedad Cultural INSUB, por la identificación morfológica de las especies

2.

Métodos

2.1. ESTACIONES DE CONTROL

Con objeto de tener un cierto gradiente de presión, para este estudio se escogieron para el muestreo 26 estaciones que se controlan en el marco de los programas de seguimiento de las aguas de transición y costeras de la Demarcación Hidrográfica del Cantábrico Oriental, cubriendo tanto aguas de transición (20 estaciones), como costeras (4 estaciones de aguas costeras y 2 de plataforma litoral), así como las diferentes tipologías descritas anteriormente. Sin embargo, no todas las estaciones se muestrearon todos los años (Tabla 3).

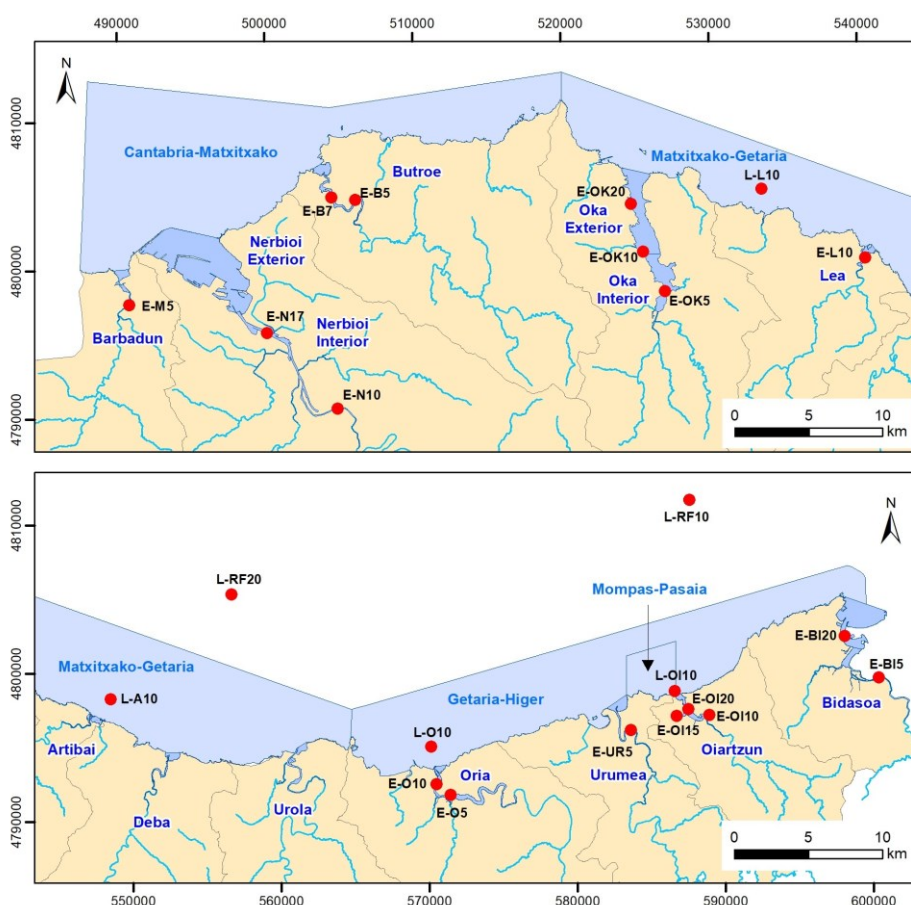


Figura 1. Ubicación de las estaciones de control de la Red de seguimiento del estado ecológico de las aguas de transición y costeras de la CAPV seleccionadas para este estudio.

Tabla 3 Estaciones de muestreo en aguas de transición y en aguas costeras, muestreadas.

Masa de agua	Código estación	UTMX ETRS89	UTMY ETRS89	Estación	Tipo	Años muestreo
Getaria-Higer	L-O10	570105	4795093	Litoral de Orio	AC-T12 somero	2018-2022
	L-O110	586537	4798855	Litoral de Pasaia	AC-T12 somero	2018-2022
Matxixako-Getaria	L-L20	541347	4802354	Litoral de Lekeitio	AC-T12 somero	2018-2022
	L-A10	548439	4798291	Litoral de Ondarroa	AC-T12 somero	2018-2022
Bidasoa	E-BI20	598024	4802583	Hondarribia (Txingudi)	AT-T09	2018-2022
	E-BI5	600337	4799756	Irún (Behobia)	AT-T08	2018-2022
Oiartzun	E-OI15	586667	4797168	Pasaia San Pedro (Dársena Herrera)	AT-T10	2018-2022
	E-OI10	588878	4797244	Lezo	AT-T10	2022
	E-OI20	587465	4797618	Pasaia (San Pedro)	AT-T10	2018-2022
Urumea	E-UR5	583597	4796227	Donostia (Loiola)	AT-T08	2022
Oria	E-O5	571392	4791824	Orio (rampa)	AT-T08	2018-2022
	E-O10	570456	4792569	Orio (puente de la autopista)	AT-T09	2018-2022
Urola	E-U8	561250	4793514	Zumaia (puente del ferrocarril)	AT-T08	2018-2021
Deba	E-D10	552145	4793494	Deba (puente)	AT-T08	2018-2021
Artibai	E-A5	545136	4796732	Ondarroa (Erreterria)	AT-T08	2018-2021
Lea	E-L10	540602	4800938	Lekeitio (molino)	AT-T08	2018-2022
Oka Interior	E-OK5	527059	4798683	Gernika (salida de la depuradora)	AT-T08	2021-2022
Oka Exterior	E-OK10	525598	4801359	Murueta (astillero)	AT-T09	2022
	E-OK20	524758	4804573	Sukarrieta (Txatxarramendi)	AT-T10	2018-2022
Butroe	E-B5	506146	4804824	Plentzia (Abanico)	AT-T08	2018-2022
	E-B7	504518	4805004	Plentzia (campo de fútbol)	AT-T09	2018-2022
Nerbioi Interior	E-N10	504948	4790762	Bilbao (puente de Deusto)	AT-T08	2018-2022
	E-N17	500185	4795862	Leioa (Lamiako)	AT-T09	2018-2022
Barbadun	E-M5	490876	4797710	Muskiz (Petronor)	AT-T08	2018-2022
Plataforma	L-RF10	587545	4811735	Litoral Oiartzun - plataforma	AC-T12 profundo	2018-2022
	L-RF20	556693	4805474	Litoral Deba - plataforma	AC-T12 profundo	2018-2022

2.2. PROTOCOLO DE MUESTREO

Como el objetivo del estudio es tratar de replicar lo más posible los resultados de M-AMBI usando M-gAMBI, se decidió seguir el mismo protocolo de muestreo⁵. Así:

- El periodo de muestreo fue en invierno (enero-marzo), coincidiendo con las abundancias mínimas y fuera de la época principal de reclutamiento, entre 2018 y 2023, ambos años incluidos.
- Las estaciones de muestreo válidas se posicionan en zonas sedimentarias, y se muestrea el tipo sedimentario mayoritario en la masa de agua, minimizando los cambios en las características sedimentarias debidos a la posible heterogeneidad espacial.
- El muestreo en zona intermareal se realiza en la zona más cercana al 0 de marea y se utiliza un cuadrado metálico de 0,5 x 0,5 m. La profundidad de la muestra obtenida es de 15 cm.
- El muestreo en zona submareal se realiza preferentemente sobre la cota de 25-30 m, evitando las zonas inestables sometidas a fuertes corrientes o cambios por oleaje (transporte intenso de sedimentos), así como a fuertes gradientes batimétricos. Los muestreos se realizan independientemente de las condiciones de marea, pero con altura de ola inferior a 1,0 m, para evitar muestras incompletas por anomalías en el cierre de la draga. En las zonas submareales se muestrea con una draga Van Veen (0,1 m²).

Tanto en los muestreos intermareales como los submareales se tomaron seis **réplicas**, tres se dedicaron al análisis tradicional morfológico y tres al análisis genómico. En todas ellas el **pretratamiento de la muestra** consistió en:

- Tamizado de inmediato, a través de una malla de 1 mm. Las piedras y los animales grandes fueron separados para evitar daños al resto de la muestra. El tamiz se lavó con agua de mar tras cada operación de tamizado por réplica, eliminando todas las impurezas.
- Submuestreo: Si el volumen de la muestra tamizada fue superior a 3 L, se realizó un cuarteo, de modo que se tomaron las fracciones de octavos (p.ej.: 1/8, 2/8, etc.) necesarias hasta alcanzar el volumen total más próximo a 3 L, sin superarlo. Este cuarteo se tuvo en cuenta, en los trabajos de laboratorio, para el cálculo de densidades por estación.
- Todo el material retenido se recogió mediante suspensión en agua desde la zona inferior del tamiz, y se introdujo en un recipiente para cada réplica, convenientemente identificado y rotulado.
- Una vez tamizado el material e introducido en botes, con objeto de conservar las muestras, se procedió a la fijación en una solución de formol (1 parte de formaldehído al 40% y 9 partes de agua de mar), neutralizado con tetraborato de sodio (bórax), para las muestras con identificación morfológica. Sin embargo, las muestras para análisis molecular fueron conservadas en -20 °C hasta su procesamiento.

2.3. PROTOCOLO DE ANÁLISIS. IDENTIFICACIÓN MORFOLÓGICA

El procesado de las muestras para **identificación morfológica** implica el lavado de estas con agua dulce y la separación de los organismos bentónicos del resto de material (partículas de sedimento, detritos, etc.), mediante pinzas, primero a simple vista y luego en lupas con el material en placas de Petri.

Las muestras se trían con lupa, y se realiza la identificación hasta nivel de especie, en aquellos taxones en los que es posible. La identificación de especies se apoya en el Registro Europeo de Especies Marinas (www.marbef.org/data/erms.php) y en los códigos NODC o ITIS (<http://www.itis.usda.gov/>). Todos los individuos de cada taxón son contabilizados, excepto los coloniales, que se anotan como presencia.

Los datos de abundancia se transforman a datos de densidad (en número de individuos por metro cuadrado), en función del área de superficie muestreada, bien sea con draga o con cuadrado y, en su caso, de los cuarteos realizados.

2.4. PROTOCOLO DE ANÁLISIS. IDENTIFICACIÓN GENÓMICA

Por su parte, para las muestras destinadas a **identificación genómica**, hay un trabajo de **preprocesado**, en el que las muestras obtenidas se preprocesan previamente antes de realizar la extracción de ADN dependiendo del tipo de sedimento. Las muestras se dividen en dos tipos:

- Sedimento de tipo medio a grueso (arena o piedra): Se realiza una decantación para separar la fracción orgánica del sedimento en varios pasos, y se homogeneiza la muestra para su extracción en batidora o mortero dependiendo de su volumen.
- Sedimento fangoso: Se realiza una homogeneización completa de la muestra utilizando una batidora de gran volumen (batidora de vaso) y se recoge una muestra representativa de la muestra (12 – 15 g).

En ambos casos, se guarda una fracción de la muestra en etanol al 96-100% y se almacenan a -20°C

para su conservación hasta posterior extracción de ADN.

Para la **extracción de ADN**, las muestras preservadas en etanol se centrifugan a 2000×g durante 5 min, retirando el líquido.

El ADN del material sólido se extrae utilizando el kit comercial DNEasy PowerMax Soil (Qiagen, Cat. No. / ID:12988-10), siguiendo las recomendaciones del fabricante, con una modificación durante el proceso de lisis; una fracción de muestra (hasta 10 g) se digiere añadiendo la enzima Proteinasa K e incubando 16-18 horas a 56 °C en movimiento.

Al día siguiente, se continúa con diferentes pasos de purificaciones hasta obtener el ADN, específicamente utilizando el kit DNeasy PowerClean Pro Cleanup Kit (Qiagen, Cat. No. / ID: 12997-50) siguiendo las instrucciones del fabricante, para eliminar inhibidores de PCR. Se cuantifican los extractos de ADN mediante el kit de ensayo dsDNA BR utilizado en un fluorómetro Qubit.

Al final, se preparan alícuotas normalizadas a una concentración de 10 ng μL^{-1} , diluyendo con agua Ultrapure. Los extractos se guardan a -80°C.

Una vez concluido ese paso, es necesario **preparar las librerías genéticas para secuenciación**. Las librerías se preparan siguiendo el protocolo de Illumina “16S Metagenomic Sequencing Library Preparation”. Brevemente, las alícuotas de cada muestra se amplifican mediante una primera PCR usando los cebadores universales para la región “Leray fragment” del gen mitocondrial Citocromo Oxidasa I (COI) (Leray *et al.*, 2013) a los que se les añade una secuencia “linker” específica que permite una segunda amplificación con adaptadores Illumina.

La amplificación se llevó a cabo bajo las siguientes condiciones: 98°C durante 3 minutos; 30 ciclos de 98°C durante 15s, 46°C durante 30s y 72°C durante 45s; y al final 72°C durante 5 minutos.

En la segunda amplificación (98°C 3 min. y 8 ciclos de 98°C 30s, 55°C 30s y 72°C 30s; más 72°C 5 min) cada muestra se etiqueta con una combinación diferente de cebadores específicos para el “linker” y con un índice específico para cada muestra (secuencia “index” o “tag”), lo que permite secuenciar centenares de muestras a la vez y su posterior separación en el proceso de análisis de los datos de secuencia.

Después de cada amplificación se purificó el producto mediante beads magnéticos (Ampure XP, Beckman Coulter) para quitar los cebadores restantes.

El resultado de la amplificación de cada muestra se cuantifica utilizando Qubit, y basado en esta cuantificación se prepara una mezcla equimolar, es decir con una cantidad igual de amplicons de ADN de cada muestra.

Las librerías generadas se envían a un servicio externo (CNAG, Centro Nacional de Análisis Genómico, Barcelona) para su secuenciación en el MiSeq con reactivos v3 con lecturas Paired-End de 300 ciclos.

A partir de los datos sin procesar obtenidos de la instalación de secuenciación (en formato FASTQ), se llevaron a cabo los siguientes pasos para producir lecturas correspondientes al amplicon COI completo, eliminando secuencias problemáticas y agrupando las restantes según la clasificación taxonómica:

- Se utilizó la herramienta FASTQC para visualizar la calidad a lo largo de los *reads* y controlar la presencia de secuencias de linkers o adapters restantes.

- El solapamiento de pares de lectura, la eliminación del cebador y el filtrado de calidad inicial se realizaron utilizando vsearch (v2.17.0; Rognes et al., 2016) y cutadapt (v3.5; Martin, 2011). Se aceptó un máximo de 40 discrepancias al solapar *reads* emparejados. Se descartaron las secuencias con cebadores incompletos o con más de dos bases de diferencia entre la secuencia y los cebadores. Se descartaron también *reads* solapados con un tamaño de menos de 274 o más de 333 bases.
- La agrupación de *reads* por similitud (sin tener en cuenta la taxonomía) se realizó usando SWARM (v2.2.1; Mahé et al., 2015), según la configuración predeterminada, seguido de la eliminación de "singletons" (secuencias únicas con un solo *read*) y filtrado de secuencias quiméricas (*de novo* y basado ambos en secuencias de referencia) usando vsearch. Además, se llevó a cabo una corrección posterior al agrupamiento utilizando LULU (Frøslev et al., 2017) con un límite de similitud del 97%.
- La clasificación taxonómica se realizó mediante CREST (v4.3.6; Lanzén et al., 2012) y la base de datos Midori v253 + DARN (<https://github.com/xapple/crest4>). Se requirió al menos un 97% de similitud con la secuencia de referencia para la clasificación en el rango de especie y un 95% para la clasificación en el mismo género. Los contaminantes probables se identificaron y eliminaron en función de los perfiles de abundancia y los blancos de extracción, en forma de placa, utilizando decontam (Davis et al., 2018), eliminando 37 secuencias únicas, representando 894.000 *reads* (1,8% del total). Los *reads* restantes se agruparon en función de su clasificación taxonómica en lugar de secuencias únicas corregidas por SWARM y LULU.
- Se eliminaron del listado todos los taxones sin clasificación taxonómica o clasificados como otro reino que no fuera Metazoa (24% de los *reads*). Además, se llevó a cabo una filtración basada en asignaciones taxonómicas para eliminar taxones con origen probable de medio pelágico, parasítico, terrestre o de agua dulce. Así se eliminaron aproximadamente 11 millones de *reads* adicionales (29% del total).
- Las réplicas con menos de 1.000 *reads* después de este paso de filtración (20 de 401) se descartaron por no tener una resolución suficiente para estimar la composición, estructura y diversidad de la comunidad estudiada.

Una vez que los datos de metabarcoding fueron obtenidos, en número de *reads* (como aproximación a la abundancia, en el sentido de biomasa), por réplica y nivel taxonómico, se procedió a hacer las tablas de datos más comparables con las existentes para el análisis morfológico. Esto se debe a que en metabarcoding se obtienen datos de multitud de taxa, a diferentes niveles taxonómicos, desde especie a phylum, alcanzándose 1.289 taxa (26,9 millones de *reads*).

Para ello se eliminaron del listado todos los niveles taxonómicos superiores a género, excepto en el caso de los phyla Nematoda y Nemertea, la clase Insecta, la subclase Oligochaeta y el orden Diptera. Asimismo, se sumaron todos los *reads* de taxones de cada uno de los grupos indicados en el punto anterior, para hacerlos comparables con los resultados de la identificación morfológica, en la que no se alcanza mayor resolución para dichos grupos. Si había duplicados (especies o géneros con algún error de escritura, etc.), se suman los *reads* y se deja un taxón. Finalmente, se eliminaron todos los taxa que sumaban menos de 10 *reads* en el total de las muestras y réplicas (líneas horizontales que sumadas tienen valores entre 0 y 9), ya que se considera que no son representativos de la fauna presente en las muestras.

Una vez hechas estas correcciones, el número de taxones disponibles identificados por metabarcoding fue de 310 (22,9 millones de *reads*).

2.5. EVALUACIÓN DE LOS DATOS

Tanto los datos basados en morfología, como los de metabarcoding, fueron usados para calcular AMBI, riqueza, diversidad y M-AMBI, y sus respectivas versiones basadas en genómica (g-). Las condiciones de referencia de M-AMBI para cada tipología son las de la Tabla 1, mientras que para M-gAMBI se determinarán según los valores máximos de riqueza y diversidad y mínimo de AMBI correspondientes a cada tipo de los cinco investigados, tal y como se ha hecho en algunos lugares cuando no hay un número abundante de datos (Borja et al., 2019). Para las condiciones de referencia de mal estado se han usado: AMBI = 6, riqueza = 0 y diversidad = 0; para cada tipo.

En unas pocas estaciones y años (E-N10 y E-U8 en 2019, E-L10 en 2020, y E-O10 en 2021), por razones diversas, solo quedó disponible una réplica. Como esto puede afectar a los datos de riqueza y diversidad, estas estaciones se eliminaron del análisis.

Por otro lado, los datos de metabarcoding de E-OK20 en 2020, 2021 y 2023, resultan poco consistentes con los morfológicos, por lo que también se han suprimido.

Como análisis estadístico, se han realizado regresiones y determinación de la correlación entre los análisis morfológicos y de metabarcoding, considerándose como significativas las correlaciones con una $p < 0,05$.

Con el fin de evaluar si las diferencias en los valores de diversidad y de riqueza vienen causadas por diferencias en la determinación de los taxones presentes en las muestras, según se realice tal determinación por métodos genómicos o morfológicos tradicionales, se llevaron a cabo análisis multivariantes tipo *cluster* basados en el índice de similitud de Bray-Curtis (Bray y Curtis, 1957).

A nivel de réplica se trabajó tomando los datos de presencia/ausencia de los taxones identificados (tras aplicar los filtros arriba indicados), como base para la elaboración de las matrices triangulares de similitudes. En cambio, a nivel de muestra, la matriz triangular se elaboró a partir de los datos de abundancia relativa (número de individuos de cada taxón dividido entre el total de individuos identificados en la respectiva muestra, en el caso de los datos provenientes de la identificación tradicional; y número de *reads* de cada taxón dividido entre el total de *reads* de la respectiva muestra, para los datos procedentes del análisis genómico).

Las matrices triangulares así elaboradas se emplearon para generar dendrogramas jerárquicos, usando la media de grupo como método de agrupamiento (Clarke y Warwick, 2001). Los resultados de cada una de las matrices fueron sometidos a sendas rutinas SIMPROF Ltd (Clarke y Gorley, 2006), que permiten comprobar qué agrupaciones son estadísticamente significativas.

Por último, se llevaron a cabo análisis SIMPER para determinar los taxones que contribuían en mayor medida a la ordenación obtenida, bien por su peso relativo en (o aportación relativa a) la similitud intragrupal, o bien por peso relativo en la disimilitud intergrupala (Clarke, 1993). Este análisis permite identificar, por un lado, los taxones que sistemáticamente son identificados por una de las técnicas (morfología o taxonomía) y no por la otra. Por otro lado, permite identificar taxones que pueden ser características de algunos de los grupos.

Todos estos análisis fueron llevados a cabo con el *software* Primer v6.1.12 de Primer e-Ltd.

3.

Resultados

3.1. COMPARACIÓN DE VARIABLES SIMPLES

En la Figura 2 se pueden observar los resultados de comparar AMBI, diversidad y riqueza, calculados usando métodos tradicionales y genómicos.

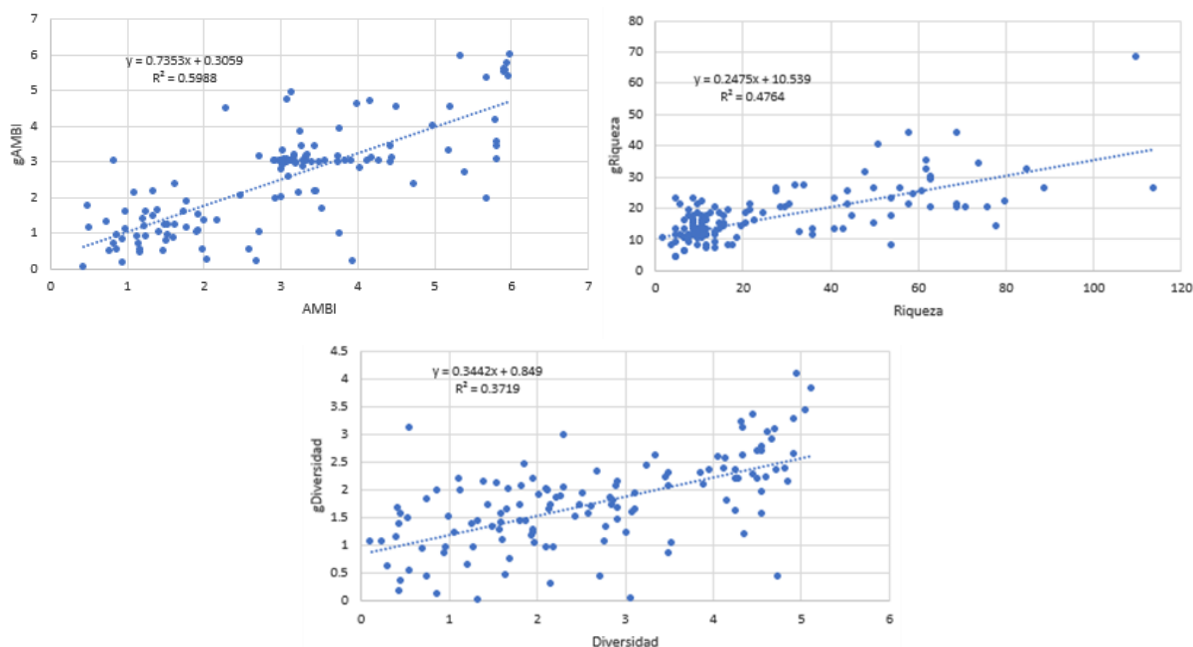


Figura 2. Regresión entre AMBI, riqueza y diversidad calculados con el método morfológico y con el genómico (gAMBI, gRiqueza y gDiversidad), incluyendo todos los datos y muestras obtenidas entre 2018 y 2023.

En los tres casos la correlación es altamente significativa ($p < 0,001$), con r^2 de 0.6 en AMBI, 0.48 en riqueza, y 0.37 en diversidad. Aunque estos resultados resultan esperanzadores, de cara a calcular un M-AMBI genómico, hay que hacer notar que estos datos resultan de juntar toda la información de estuario y costa, cuando la comparación debería hacerse por tipos (Figura 3).

En este caso, AMBI, riqueza y diversidad de los tipos AT-T08 y AT-T10 presentan una correlación significativa ($p < 0,05$) entre los datos morfológicos y los de metabarcoding. En el tipo AT-T09 solo la riqueza es significativa, mientras que en AC-T12 someras lo es la diversidad. En el caso de AC-T12 profundas las tres variables están cerca de ser significativas, pero, posiblemente, el bajo número de datos lo impide.

Uno de los problemas que se observa en la Figura 3 es que el **gradiente de presión** (si consideramos su medida con AMBI) no es completo en todos los tipos estudiados. Así, en el AT-T08, que es típico del interior de estuarios, no hay casi datos con valores bajos de AMBI, mientras que en los dos tipos costeros (AC-T12 somero y profundo), no hay datos de AMBI con valores que muestren disturbancia. Esto puede explicar que sea difícil encontrar una buena regresión entre las variables en estos tipos, porque su gradiente es muy corto (Figura 3).

En el caso de la **riqueza**, parece que los valores morfológicos tienden a ser menores que los obtenidos con metabarcoding en la parte interna de los estuarios (AT-T08) y esto se va atenuando hacia el exterior de éstos (AT-T09, AT-T10) y se invierte en la zona costera, donde los valores morfológicos son bastante mayores que los de metabarcoding en ambos tipos (Figura 3). En el caso de la **diversidad**, parece que en todos los tipos la calculada de manera tradicional es mayor que la calculada con metabarcoding (Figura 3). Algo similar, pero en menor medida, parece darse en el caso de **AMBI** (Figura 3).

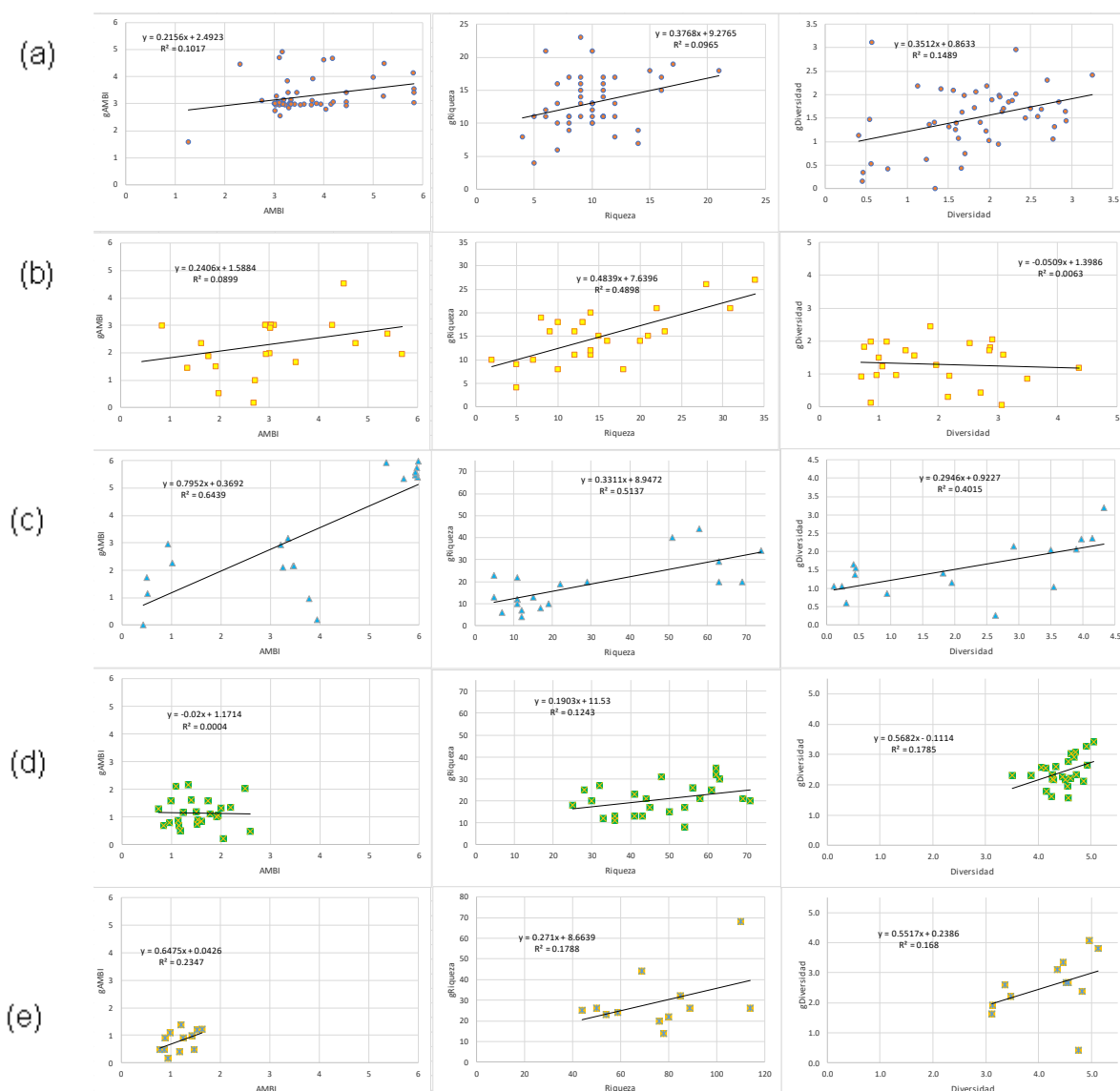


Figura 3. Regresión entre AMBI, riqueza y diversidad calculados con el método morfológico y con el genómico (g), por tipo, incluyendo todos los datos y muestras obtenidas entre 2018 y 2023. Tipos de aguas de transición: (a) AT-T08, *Cerastoderma edule-Scrobicularia plana*; (b) AT-T09, *Venus fasciata*; (c) AT-T10, *Abra alba*; Tipos de aguas costeras: (d) AC-T12, someras, *Tellina tenuis-Venus fasciata*; (e) AC-T12, profundas, *Amphiura*.

3.2. COMPARACIÓN DE M-AMBI Y M-GAMBI

Como M-gAMBI requiere condiciones de referencia para cada tipo, como se ha mencionado en la metodología, se han usado los valores más altos de riqueza y diversidad, y los menores de AMBI, en cada tipo (Tabla 4).

Si se comparan la Tabla 1 y Tabla 4, se observa que las condiciones de referencia de **riqueza** para el M-AMBI presentan un gradiente creciente desde el interior de los estuarios (13 especies) hasta la parte más profunda de la costa (130 especies). En el caso de metabarcoding, también hay un gradiente creciente, que se rompe en la costa somera (Tabla 4). En todo caso, el rango es mucho menor, ya que comienza con 23 especies y acaba en 68.

La **diversidad** morfológica también aumenta desde el interior del estuario (2,5) hasta la zona profunda (5,7), aunque con un pequeño cambio entre AT-T09 y AT-T10 (Tabla 1). En metabarcoding sucede igual, aunque también con un rango menor, entre 3,11 y 4,08 (Tabla 4).

Finalmente, las condiciones de referencia de **AMBI** morfológico presentan un fuerte gradiente desde el interior del estuario (2,8) a las zonas profundas (1), algo que no sucede tanto en metabarcoding (Tabla 4).

Tabla 4 Condiciones de referencia utilizadas para calcular M-gAMBI, para masas naturales. Los datos son los usados en la intercalibración. gS: riqueza, gH': diversidad de Shannon. La "g" indica que son datos calculados con genómica.

Categoría	Tipología	Tramo salino/ Profundidad	Comunidad asociada	gS	gH'	gAMBI
Aguas de transición	AT-T08, AT-T09, AT-T10	0-18	<i>Cerastoderma edule-Scrobicularia plana</i>	23	3,11	1,58
		18-30	<i>Venus fasciata</i>	27	2,44	0,18
		30-34	<i>Abra alba</i>	44	3,20	0,00
Aguas costeras	AC-T12	20-50 m	<i>Tellina tenuis-Venus fasciata</i>	35	3,42	0,22
		70-120 m	<i>Amphiura</i>	68	4,08	0,16

Una vez calculado M-gAMBI, usando las condiciones de referencia de la Tabla 4, en la Figura 4 se observa la regresión entre M-AMBI y M-gAMBI, usando todos los datos conjuntamente.

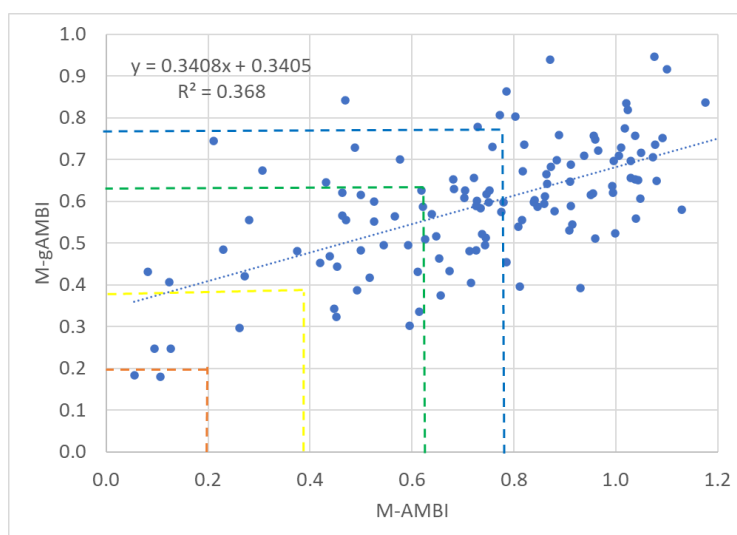


Figura 4. Regresión entre M-AMBI y M-gAMBI, incluyendo todos los datos y muestras obtenidas entre 2018 y 2023. Las líneas indican los límites de clase: Naranja: entre mal y deficiente estado ecológico, Amarillo: entre deficiente y moderado, Verde: entre moderado y bueno, Azul: entre bueno y muy bueno. Hay que hacer notar que la verde se ha puesto con el límite de costa, mientras que el de estuario es menor.

La correlación es significativa ($p < 0,001$), siendo la r^2 de 0,37 pero el intercepto en M-gAMBI es en 0,34, lo que indica que los valores de M-gAMBI son superiores a los de M-AMBI, al menos para los valores más bajos. Esto se observa bien con las líneas que determinan el estado ecológico, si bien hay que hacer notar que en esta figura están mezclados los datos de estuario y litoral y los límites de clase de moderado-bueno son diferentes en ambos.

En la Figura 5 se observan los mismos datos, pero segregados según correspondan a muestras estuáricas o litorales. En ambos casos, la correlación es significativa ($p < 0,001$). En estuario se observa todo el gradiente de impacto, con una correlación algo menor que la que se obtiene con los resultados no segregados. Sin embargo, en el litoral el grado de correlación es inferior ($r^2 = 0,15$), y el gradiente de impacto se limita a los niveles de buen y muy buen estado ecológico.

Finalmente, en la Figura 6 se han representado las regresiones entre M-AMBI y M-gAMBI para cada uno de los tipos estudiados. Los tipos con mayor gradiente de impacto (como AT-T08, entre moderado y muy bueno, y AT-T10, entre malo y muy bueno) tienden a tener mejor correlación, siendo los únicos casos significativos ($p < 0,05$). Como ya se ha indicado anteriormente, en los tipos litorales no se cubre todo el gradiente (habiendo solo estados bueno o muy bueno), y las correlaciones son bajas.

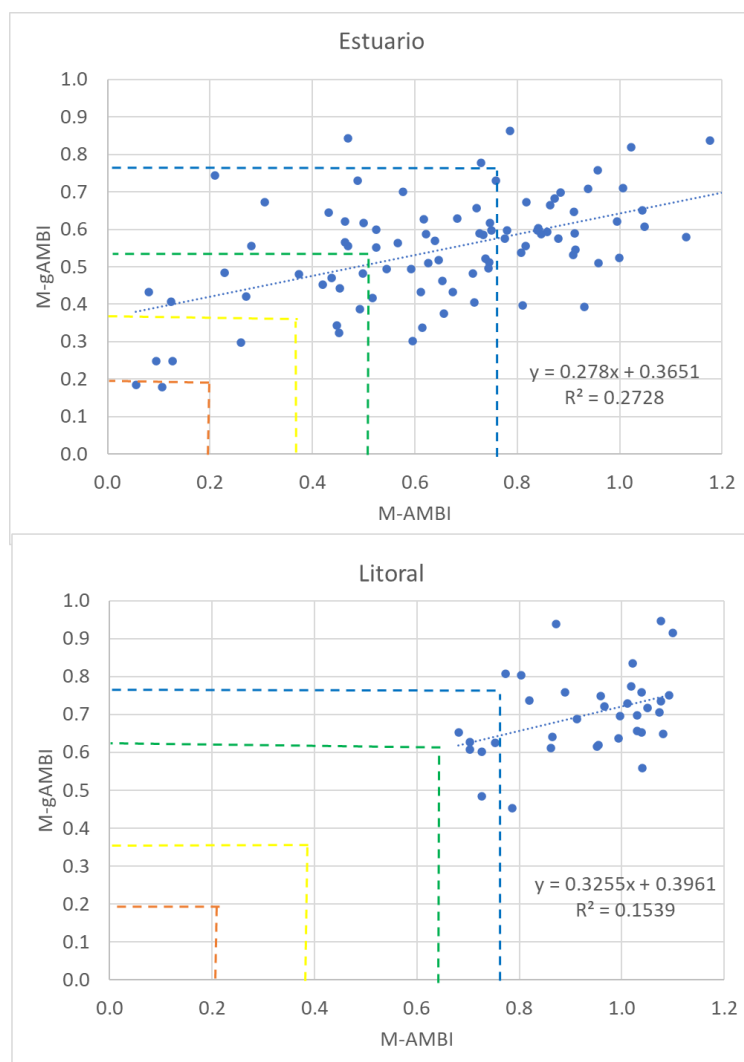


Figura 5. Regresión entre M-AMBI y M-gAMBI, incluyendo todos los datos y muestras obtenidas entre 2018 y 2023, separados entre hábitats estuáricos y litorales. Las líneas indican los límites de clase: Naranja: entre mal y deficiente estado ecológico, Amarillo: entre deficiente y moderado, Verde: entre moderado y bueno, Azul: entre bueno y muy bueno. Hay que hacer notar que los límites entre bueno y moderado son diferentes en estuario y litoral.

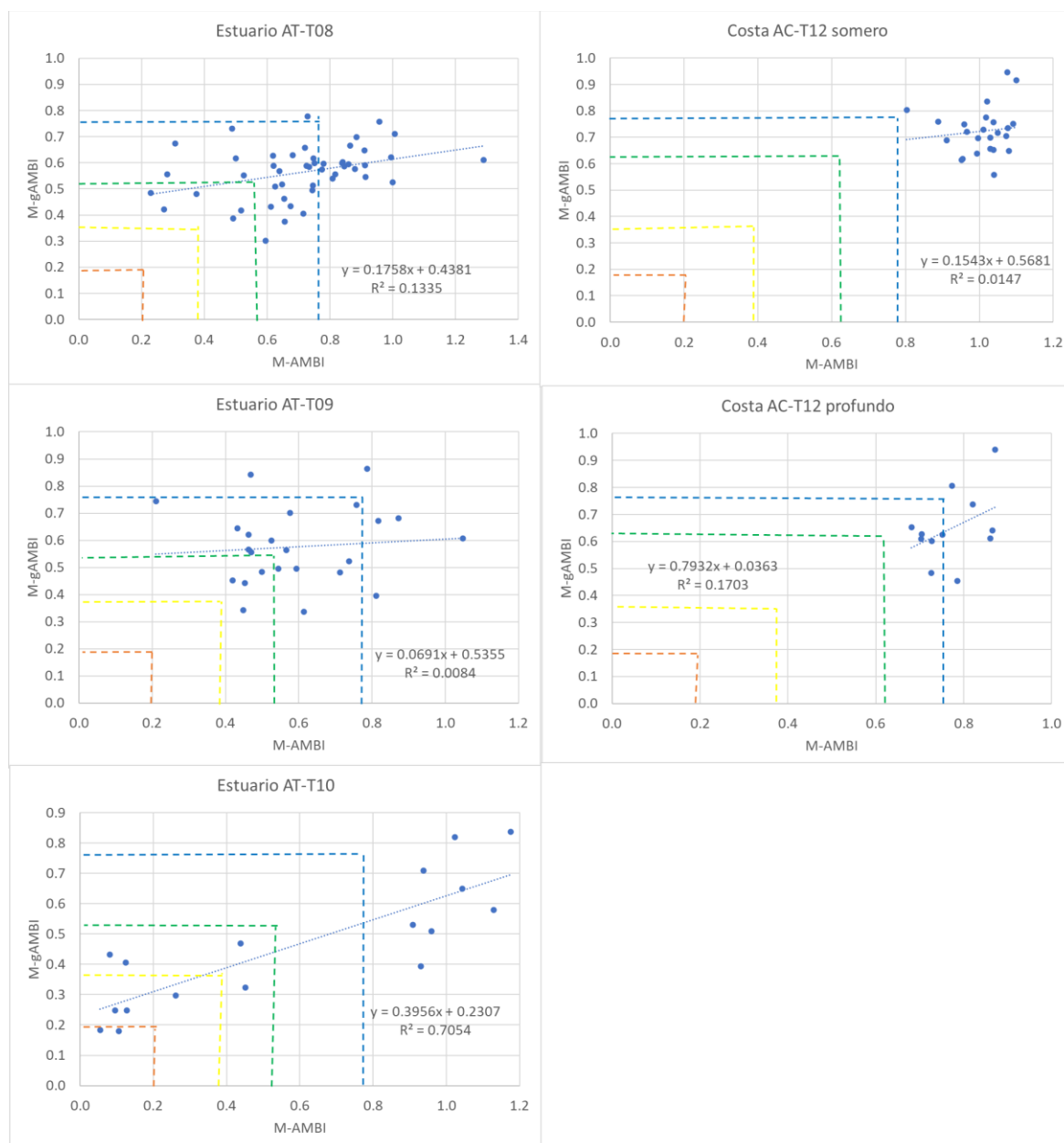


Figura 6. Regresión entre M-AMBI y M-gAMBI, incluyendo todos los datos y muestras obtenidas entre 2018 y 2023, en cada tipología estudiada. Las líneas indican los límites de clase: Naranja: entre mal y deficiente estado ecológico, Amarillo: entre deficiente y moderado, Verde: entre moderado y bueno, Azul: entre bueno y muy bueno. Hay que hacer notar que los límites entre bueno y moderado son diferentes en estuario y litoral.

3.3. ANÁLISIS DE SIMILARIDAD

En la Figura 7 se representa la ordenación de las réplicas llevadas a cabo con los datos de presencia y ausencia de los taxones identificados por métodos tradicionales y por métodos genómicos, basada en la similitud entre las muestras.

Dicha ordenación muestra una clara segregación entre las réplicas analizadas por métodos tradicionales y por métodos genómicos. De hecho, ningún grupo estadísticamente significativo contiene simultáneamente muestras procedentes de una misma estación en la que la determinación taxonómica se haya realizado por métodos tradicionales y por métodos genómicos, a excepción de sendos grupos que incluyen muestras procedentes de las estaciones E-BI5 y E-OK20.

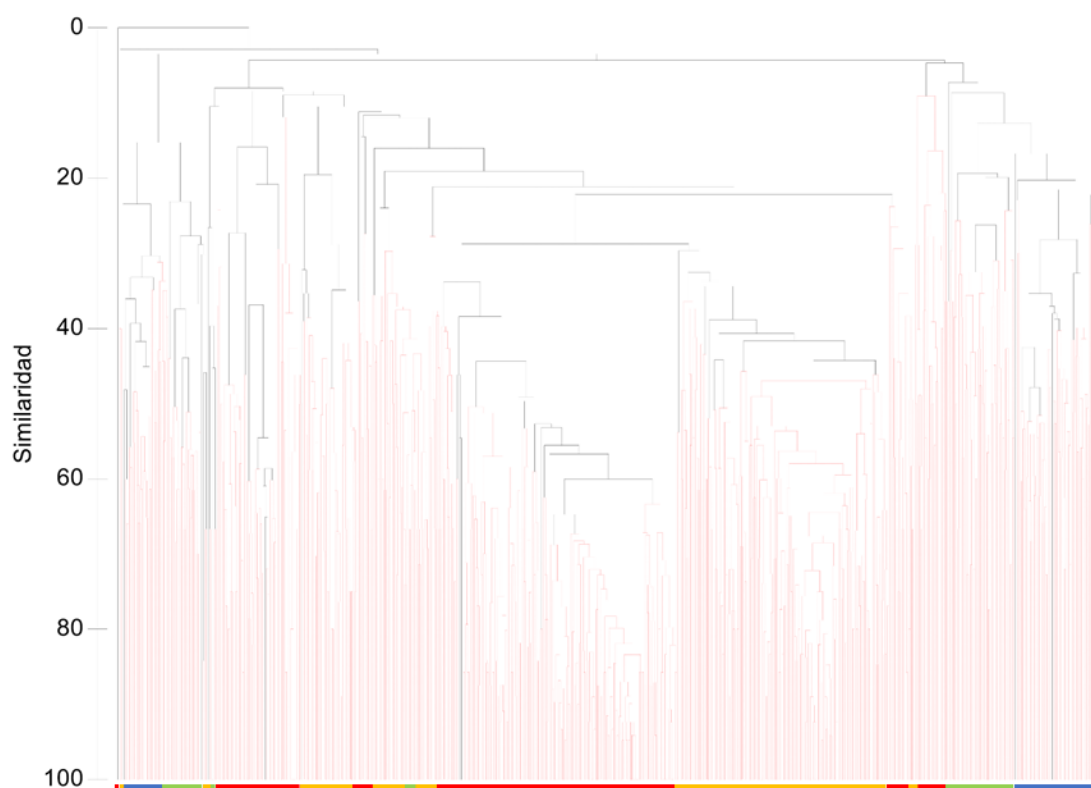


Figura 7. Dendrograma resultado de la ordenación obtenida del análisis de similaridad entre réplicas basado en los datos de presencia o ausencia de los taxones determinados por métodos tradicionales y genómicos, en las respectivas muestras. En rojo se indican las muestras estuáricas analizadas por medios tradicionales; en naranja se indican las muestras estuáricas analizadas por métodos genómicos; en azul se representan las muestras litorales analizadas por medios tradicionales; en verde se representan las muestras litorales analizadas por métodos genómicos. Las agrupaciones en rojo del dendrograma corresponden a agrupaciones significativamente distintas

El análisis SIMPER llevado a cabo evidencia una importante disimilaridad entre los principales grupos de muestras identificados en el análisis *cluster*.

En lo referido a las **muestras estuáricas**, la disimilaridad media entre los grupos que incluyen más de 20 muestras analizadas por métodos tradicionales o por técnicas genómicas (dos grupos que contienen exclusivamente muestras que se han tratado con técnicas tradicionales, uno que contiene exclusivamente muestras tratadas por técnicas genómicas, y un tercer grupo que contiene mayoritariamente muestras analizadas por métodos tradicionales) es del 64-69%.

Entre las especies con mayor contribución a dicha disimilaridad (>5%), llama la atención en las muestras analizadas por métodos genómicos la ausencia de algunas especies que se consideran características de comunidades que habitan los estuarios de la Demarcación Hidrográfica del Cantábrico Oriental, caso de los artrópodos *Alkmaria romijni* y *Cyathura carinata*, y del poliqueto *Streblospio eunatae*.

Por el contrario, destaca la ausencia (o presencia residual) en las muestras analizadas por métodos tradicionales de algunos taxones identificados a nivel de género en las muestras analizadas por métodos genómicos, como *Hediste* o *Peringia*. Estos géneros, sí se encuentran presentes en las muestras analizadas por métodos tradicionales y alcanzan, además, una presencia importante, pero fueron determinadas hasta nivel de especie (*Hediste diversicolor* y *Peringia ulvae*).

De modo similar, la disimilaridad media entre los principales grupos que incluyen más de 10 **muestras litorales** analizadas por métodos tradicionales o por métodos genómicos (dos grupos formados

exclusivamente por muestras tratadas con técnicas tradicionales y dos grupos formados exclusivamente por muestras tratadas con técnicas genómicas) es del 84-92%.

Al igual que lo indicado para las muestras procedentes de estaciones estuáricas, entre las especies con mayor contribución a la disimilaridad entre grupos ($>2,5\%$), destaca la ausencia, en las muestras tratadas por métodos genómicos, de algunas especies características del infralitoral de la Demarcación Hidrográfica del Cantábrico Oriental: los anélidos *Prionospio cristaventralis* y *Myriochele danielsseni*, y los artrópodos *Urothoe pulchella* y *Centraloecetes kroyeranus*, entre otros.

Por el contrario, destaca la ausencia, en las muestras tratadas por métodos tradicionales de algún taxón identificado a nivel de género en las muestras analizadas por métodos genómicos, como es *Diastylis*. Este género, en cambio, sí se encuentra representado en las muestras analizadas por técnicas tradicionales por taxones determinados a nivel de especie.

Por último, destaca también la identificación por métodos genómicos de algunos taxones que son característicos de zonas estuáricas, caso de poliquetos del género *Hediste* y del bivalvo *Scrobicularia plana*. Estos taxones no fueron identificados en las muestras litorales tratadas por métodos tradicionales. De hecho, las muestras estuáricas y litorales, tratadas por métodos tradicionales, no comparten grupo (con la única excepción de la muestra tomada en 2020 en la estación L-A10, que comparte grupo con 12 muestras de la estación E-OK20 y una muestra de la estación E-BI20, además de con 12 muestras tratadas por métodos genómicos, tres de ellas litorales), mientras que, en el caso de las tratadas por métodos genómicos, hasta ocho grupos incluyen tanto muestras estuáricas como litorales.

Con el fin de comprobar si las diferencias en la resolución taxonómica tenían un efecto importante sobre la alta disimilaridad entre las muestras analizadas por métodos tradicionales y las analizadas por métodos genómicos, se agruparon a nivel de género aquellos taxones que aportaban más de un 5% a la disimilaridad entre los grandes grupos estuáricos y más de un 2,5% en el caso de las estaciones litorales, es decir: *Ampelisca*, *Centraloecetes*, *Corophium*, *Diastylis*, *Eurydice*, *Hediste*, *Magelona*, *Nephtys*, *Peringia* (junto con *Hydrobia*), *Prionospio*, *Streblospio* y *Urothoe*. A continuación, se repitió el análisis.

Esta agrupación no afectó de manera relevante a la ordenación previamente obtenida y, una vez más, se observa una clara segregación entre las réplicas analizadas por métodos tradicionales y por métodos genómicos (Figura 8). De hecho, el análisis SIMPER llevado a cabo sigue evidenciando una importante disimilaridad entre los principales grupos de muestras identificados en el nuevo análisis *cluster*. También se observa que las muestras estuáricas y litorales, tratadas por métodos genómicos, se segregan en diferentes grupos (con la excepción de un grupo que contiene una muestra tomada en L-A10 en 2020, junto con 16 muestras tomadas en las estaciones E-OK20 y E-BI20, así como con una muestra tomada también en la estación E-OK20 y tratada por métodos genómicos). En cambio, las muestras estuáricas y litorales, tratadas por métodos genómicos, siguen coincidiendo hasta en 8 grupos.

En este análisis, la disimilaridad media entre los grupos que incluyen más de 20 **muestras estuáricas** analizadas por métodos tradicionales o por técnicas genómicas (tres grupos que contienen exclusivamente muestras que se han tratado con técnicas tradicionales, uno que contiene mayoritariamente muestras analizadas también por métodos tradicionales y uno que contiene mayoritariamente muestras tratadas por técnicas genómicas) es del 57-86%.

Entre las especies con mayor contribución a dicha disimilaridad ($>5\%$), sigue destacando la contribución de algunos taxones que se consideran característicos de los estuarios de la Demarcación Hidrográfica

del Cantábrico Oriental, identificados en las muestras analizadas por métodos tradicionales, pero ausentes de las muestras trabajadas por métodos genómicos, caso de: *A. romijni*, *C. carinata*, *Eurydice* sp. o *Ruditapes* sp. Por otro lado, algunos de los taxones agrupados para este análisis, siguen contribuyendo de manera importante a la disimilaridad entre grupos (p.e., *Corophium*, *Hediste*, *Peringia* o *Streblospio*), por su distinta frecuencia en las muestras según se trataran con métodos tradicionales o genómicos.

De igual modo, la disimilaridad media entre los principales grupos que incluyen más de 10 **muestras litorales** analizadas por métodos tradicionales o por métodos genómicos (dos grupos formados exclusivamente por muestras tratadas con técnicas tradicionales y tres grupos formados exclusivamente por muestras tratadas con técnicas genómicas) sigue siendo también alta (79-89%). Como ya se ha indicado para las muestras procedentes de estaciones estuáricas, entre las especies con mayor contribución a la disimilaridad entre grupos (>2,5%), siguen destacando algunos de los taxones que se han agrupado a nivel de género (*Ampelisca*, *Centraloecetes*, *Diastylis*, *Eurydice*, *Magelona*, *Nephtys*, *Prionospio*, y *Urothoe*). Además, sigue destacando la presencia entre los taxones que más contribuyen a la disimilaridad entre géneros, de poliquetos característicos de ambientes estuáricos (género *Hediste* y *Scrobicularia plana*).

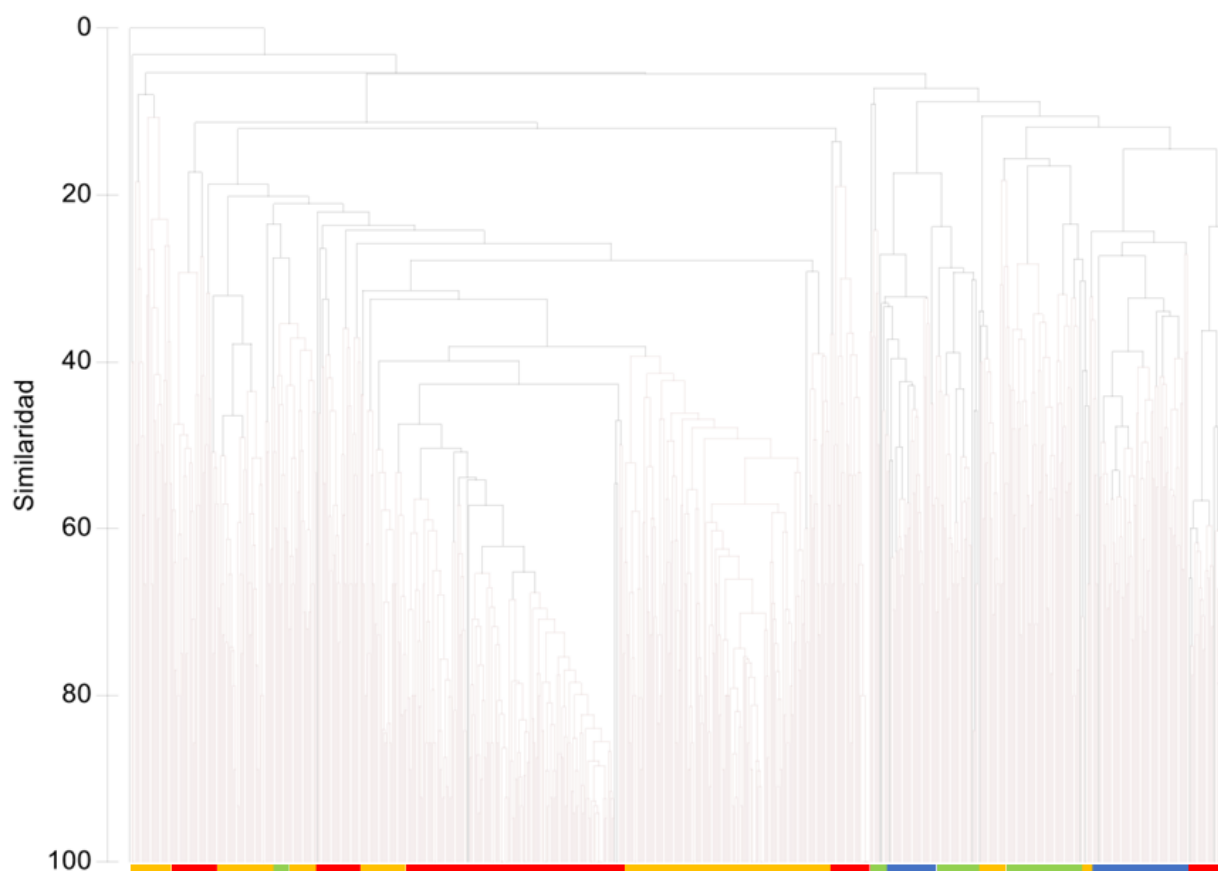


Figura 8. Dendrograma resultado de la ordenación obtenida del análisis de similaridad entre réplicas basado en los datos de presencia o ausencia de los taxones determinados por métodos tradicionales y genómicos, en las respectivas muestras, agrupando a nivel de género los taxones: *Ampelisca*, *Centraloecetes*, *Corophium*, *Diastylis*, *Eurydice*, *Hediste*, *Magelona*, *Nephtys*, *Peringia* (junto con *Hydrobia*), *Prionospio*, *Streblospio* y *Urothoe*. En rojo se indican las muestras estuáricas analizadas por medios tradicionales; en naranja se indican las muestras estuáricas analizadas por métodos genómicos; en azul se representan las muestras litorales analizadas por medios tradicionales; en verde se representan las muestras litorales analizadas por métodos genómicos. Las agrupaciones en rojo del dendrograma corresponden a agrupaciones significativamente distintas

En vista de los altos valores de disimilaridad entre réplicas, se procedió a repetir el análisis promediando las composiciones de las réplicas de cada estación para cada una de las campañas, y manteniendo las agrupaciones a nivel de género usadas en el análisis previo (por medio de la suma de abundancias en el caso de las réplicas tratadas por métodos tradicionales y la suma de *reads* para las muestras trabajadas por métodos genómicos). Además, en lugar de los datos de presencia o ausencia de los taxones identificados, el análisis se realizó en base a las abundancias relativas (abundancia o número de *reads* específico, dividido por la abundancia o número de *reads* total de la muestra promediada) (Figura 9).

La ordenación así obtenida también muestra una clara segregación entre las muestras analizadas por métodos tradicionales y por métodos genómicos. Sin embargo, sí se observa algún grupo estadísticamente significativo que contiene simultáneamente muestras procedentes de una misma estación y campaña en la que la determinación taxonómica se haya realizado por métodos tradicionales y por métodos genómicos, aunque siempre junto con otras muestras. Es el caso de algunas muestras tomadas en las estaciones: E-M5 (en 2019); E-OK10 (en 2023); E-BI5 (en 2020 y 2023); y E-BI20 (en 2019 y en 2020) y E-OK20 (en 2018, 2019 y 2023).

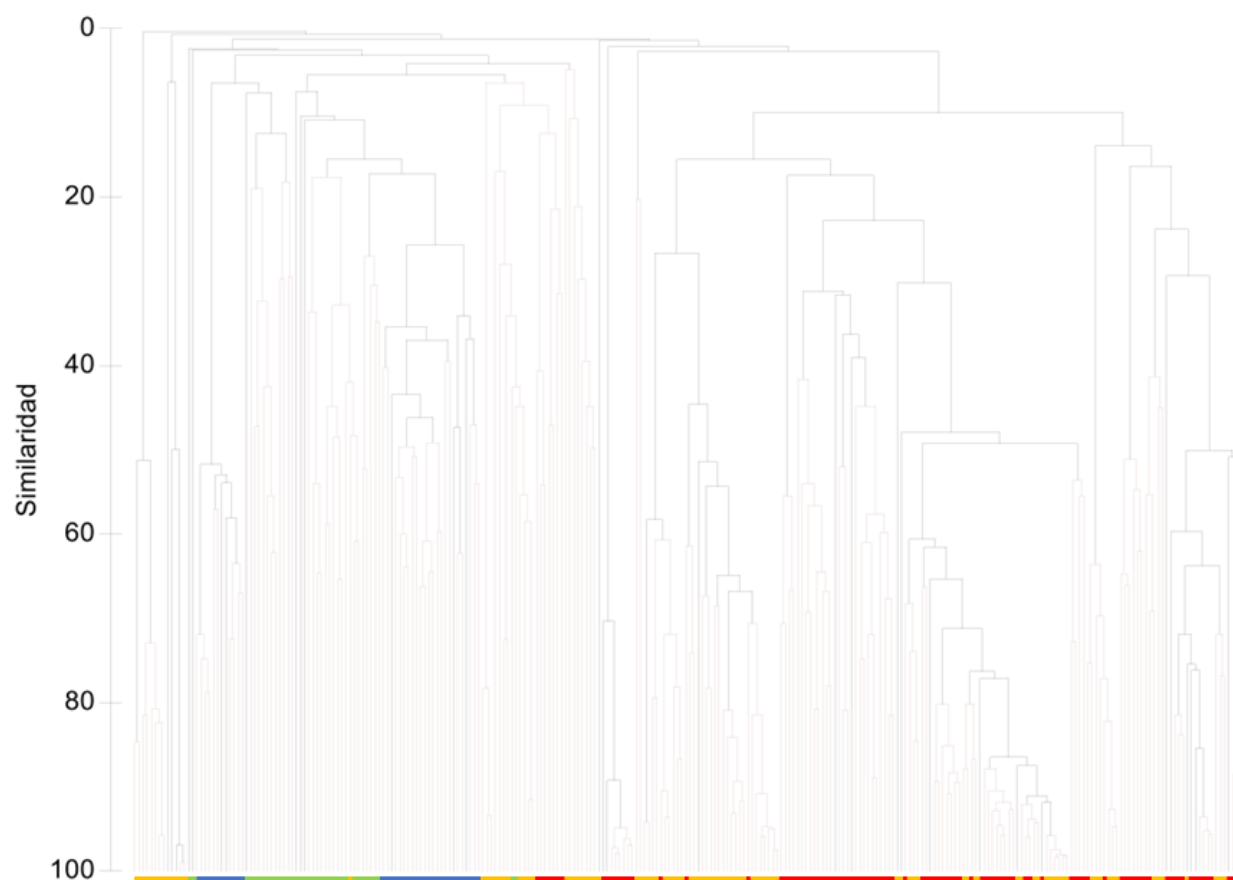


Figura 9. Dendrograma resultado de la ordenación obtenida del análisis de similaridad entre muestras (consistentes en la composición promediada de las réplicas), basado en los datos de abundancia o número de *reads* relativos de los taxones determinados por métodos tradicionales y genómicos, respectivamente. Se han agrupado a nivel de género los taxones: *Ampelisca*, *Centraloecetes*, *Corophium*, *Diastylis*, *Eurydice*, *Hediste*, *Magelona*, *Nephtys*, *Peringia* (junto con *Hydrobia*), *Prionospio*, *Streblospio* y *Urothoe*. En rojo se indican las muestras estuáricas analizadas por medios tradicionales; en naranja se indican las muestras estuáricas analizadas por métodos genómicos; en azul se representan las muestras litorales analizadas por medios tradicionales; en verde se representan las muestras litorales analizadas por métodos genómicos. Las agrupaciones en rojo del dendrograma corresponden a agrupaciones significativamente distintas.

El análisis SIMPER llevado a cabo indica que el grupo que incluye las muestras tomadas en 2019 en E-M5 presenta una similaridad promedio del 85%, siendo los poliquetos del género *Hediste* los que más contribuyen a dicha similaridad (98% de la similaridad total).

Por otro lado, el grupo que incluye las muestras tomadas en 2023 en la estación E-OK10 presenta una similaridad del 83%, siendo los gasterópodos del género *Peringia* los taxones con mayor contribución a la similaridad (84%). En cambio, los artrópodos del género *Corophium* son los que más contribuyen (93%) a la similaridad del grupo que incluye las muestras tomadas en 2020 y 2023 en la estación E-BI5.

Por último, la similaridad del grupo que incluye las muestras tomadas en 2019 y en 2020 en E-BI20, y las muestras tomadas en 2018, 2019 y 2023 en E-OK20 es de tan sólo el 17% (aunque la agrupación es significativa), siendo los poliquetos del género *Nephtys* los principales contribuyentes a dicha similaridad (76%).

Además, el análisis SIMPER evidencia una importante disimilaridad entre los principales grupos de muestras identificados en el análisis *cluster*.

En lo referido a las **muestras estuáricas**, la disimilaridad media entre los dos grupos que incluyen más muestras analizadas por métodos tradicionales o por técnicas genómicas (un grupo que contiene exclusivamente 10 muestras tratadas con técnicas tradicionales, y uno que contiene 7 muestras tratadas por métodos tradicionales y 13 tratadas por métodos genómicos, incluidas dos muestras litorales) es del 95%.

Además, entre las especies con mayor contribución a dicha disimilaridad (>5%), sigue llamando la atención en el grupo que incluye muestras analizadas por métodos genómicos la baja abundancia promedio (en comparación con el grupo formado exclusivamente por muestras tratadas por métodos tradicionales) de algunas especies que se consideran características de comunidades que habitan los estuarios de la Demarcación Hidrográfica del Cantábrico Oriental, caso del artrópodo *C. carinata*, y de los moluscos del género *Peringia*.

Por el contrario, destaca la ausencia en el grupo formado exclusivamente por muestras tratadas por métodos tradicionales, de taxones con alta abundancia relativa en el otro grupo, como son los poliquetos del género *Nephtys* y el molusco *Tritia pellucida*.

De modo similar, la disimilaridad media entre los principales grupos que incluyen más de 5 **muestras litorales** analizadas por métodos tradicionales o por métodos genómicos (dos grupos formados exclusivamente por 6 muestras tratadas con técnicas tradicionales, y dos grupos formados exclusivamente por 7 y 13 muestras tratadas con técnicas genómicas) es del 81-97%.

De igual modo a lo indicado para las muestras procedentes de estaciones estuáricas, entre las especies con mayor contribución a la disimilaridad entre grupos (>5%), destaca la ausencia o la presencia residual, en las muestras tratadas por métodos genómicos, de algunas especies características del infralitoral de la Demarcación Hidrográfica del Cantábrico Oriental: los anélidos *Paradoneis armata* y *Mediomastus fragilis*, el molusco *Fabulina fabula*, entre otros.

Por el contrario, destaca la ausencia, en las muestras tratadas por métodos tradicionales de algunos taxones con abundancia promedia alta en las muestras analizadas por métodos genómicos, como: el anélido *Maldane glebifex*, el molusco *Tritia reticulata*, o los equinodermos *Amphiura filiformis* y *Brissopsis lyrifera*.

Por último, destaca también la identificación por métodos genómicos de algunos taxones que son característicos de zonas estuáricas, caso de poliquetos del género *Hediste* y del bivalvo *Scrobicularia plana*. Estos taxones no fueron identificados en las muestras litorales tratadas por métodos tradicionales.

Sin embargo, con este último análisis se reduce el número de grupos que contiene simultáneamente muestras estuáricas y litorales, tratadas por métodos genómicos, a dos que contienen dos muestras litorales frente a 11 muestras estuáricas (más otras siete muestras estuáricas tratadas por métodos tradicionales), y una muestra estuárica frente a 12 litorales. Por el contrario, no se observa agrupación que contenga simultáneamente muestras estuáricas y litorales tratadas por métodos tradicionales.

4.

Discusión

4.1. VARIABLES SIMPLES

Desde hace unos años se ido imponiendo en la comunidad científica el uso de la genómica, especialmente el metabarcoding, para el monitoreo y la evaluación del estado del medio acuático, desarrollando índices bióticos que, de alguna manera, repliquen lo que hacen los índices tradicionales (Bourlat et al., 2013; Pawlowski et al., 2018).

Al comienzo, las variables estudiadas eran riqueza y diversidad, pero pronto comenzaron los primeros intentos de desarrollo de índices univariantes de evaluación del estado del medio, como **AMBI** (Ranasinghe et al., 2012). Poco después, esto dio lugar a un índice basado en genómica, gAMBI, aunque calculado solo usando la presencia-ausencia y no la abundancia (Aylagas et al., 2014).

Sin embargo, enseguida se consiguió relacionar la abundancia, basada en el número de *reads* de gAMBI, con la abundancia de individuos de AMBI. Así, Aylagas et al. (2018), encontraron una correlación entre ambos índices muy similar a la encontrada en el presente trabajo (0,84 y 0,77, respectivamente), si bien en nuestro caso el número de muestras es mucho mayor.

A través de estos resultados, se demuestra que esta relación es robusta, al igual que han demostrado otros estudios (Lejzerowicz et al., 2015; Lanzén et al., 2021; Días y Sukumaran, 2023), aunque otros autores (Willassen et al., 2022; Duarte et al., 2023) han encontrado algunas diferencias, pero usando solo presencia-ausencia y no abundancia y número de *reads*. Además, Cordier et al. (2020) han revisado los estudios que comparan monitoreo por metabarcoding y la identificación basada en la morfología en una variedad de entornos, resumiendo las diferencias fundamentales entre las dos metodologías.

Estos resultados tan esperanzadores incluyen datos conjuntos de los cinco tipos de masas de agua de la Demarcación Hidrográfica del Cantábrico Oriental, con datos de estuario y de costa agregados. Sin embargo, cuando se separan los datos por tipos, solo son significativas las relaciones en los tipos del interior y del exterior de los estuarios (AT-T08 y AT-T10). En este caso, coinciden con aquellos tipos en los que contamos con todo el gradiente de presión o disturbancia, que es importante cuando se quiere evaluar el estado ecológico (Borja et al., 2009), aunque en el caso de AT-T08 la pendiente no es muy elevada.

En los tipos costeros (AC-T12 somero y profundo), el gradiente de presión es pequeño y el estado ecológico en general es bueno, debido a la recuperación de la costa de la Demarcación Hidrográfica del Cantábrico Oriental (Borja et al., 2016), lo que hace que no haya buena correlación, pero sí una clasificación similar de no disturbado o ligeramente disturbado, medidos con ambos métodos.

En relación con **la riqueza**, es sabido que los métodos moleculares pueden incrementar el número de especies identificadas, puesto que algunos grupos difíciles de identificar morfológicamente (e.g. Nematodos, Nemertinos, Oligoquetos) pueden verse sobrerrepresentados, mientras que especies crípticas pueden ser también identificadas (Gibson et al., 2013; Carugati et al., 2015; Lobo et al., 2017; Duarte et al., 2023). Aunque, también es cierto que, tamizando las especies antes de extraer el ADN, puede llevar a una infraestimación de la riqueza de organismos que tengan estructuras más delicadas (Carugati et al., 2015).

Sin embargo, no parece haber sido el caso en nuestro estudio, donde inicialmente se encontraron 1.289 taxones que, tras agrupar algunos taxa y eliminar los menos abundantes, quedaron en 310, que es un número más parecido a los que se encuentran en el análisis morfológico, habiendo encontrado también algunos autores dicha congruencia (Keck et al., 2022).

En cambio, estos números son diferentes dependiendo de la tipología, encontrándose más riqueza con métodos moleculares en el interior de los estuarios y menos en las zonas costeras. Quizá esto pueda deberse a que especies con concha o caparazón, que tienden a ser más abundantes en costa, sean más difíciles de identificar con metabarcoding (Hajibabaei et al., 2019). Otros autores, usando protocolos estandarizados, encuentran riquezas similares entre diferentes laboratorios (van den Bulcke et al., 2023), usando muestras totales tamizadas por una malla de 1 mm, como se ha hecho en nuestro estudio.

Chariton et al. (2015), estudiando estuarios australianos con métodos moleculares, sugieren que la riqueza total no es un indicador sensible a la condición ecológica, considerando que la riqueza de grupos taxonómicos clave parece más sensible a las condiciones ambientales. Esto reforzaría el papel y la respuesta encontrada con AMBI, puesto que sus grupos ecológicos (equivalentes a grupos funcionales) responden a las presiones humanas.

A pesar de ello, la respuesta encontrada, al comparar la riqueza total con ambos métodos, resulta altamente significativa ($p < 0,001$), incluso también para cada uno de los tres tipos de estuario. Esto significa que posiblemente su uso en el cálculo de M-gAMBI puede ser considerado adecuado.

De las tres variables individuales estudiadas, **la diversidad** es la que presenta una menor correlación, aunque sigue siendo altamente significativa ($p < 0,001$). En este caso, también es significativa para dos tipos de estuario (AT-T08 y AT-T10) y para uno de costa (AC-T12 somero). También van den Bulcke et al. (2023) encontraron buena concordancia entre laboratorios para la diversidad de Shannon, asegurando que esta variable es reproducible y robusta especialmente cuando aparecen especies abundantes y de tamaño grande, aunque en nuestro caso no ha sido así en todos los tipos estudiados.

La necesidad de utilizar protocolos estandarizados, para asegurar la reproducibilidad de los resultados, es destacada por numerosos investigadores (Ransome et al., 2017; van den Bulcke et al., 2023), aunque no siempre es posible debido a los continuos cambios en las metodologías moleculares, con mejoras, diferentes extracciones, actualizaciones en las librerías de datos, etc., que hacen complicada la comparación a largo plazo.

De hecho, se ha demostrado que los cambios en los protocolos pueden llevar a grandes cambios en la composición específica y, por tanto, a cambios en la riqueza y diversidad (Deiner et al., 2015). Para minimizar los cambios que esto puede producir sería conveniente seguir las recomendaciones propuestas por van der Loos y Nijland (2021).

4.2. COMPARACIÓN ENTRE M-AMBI Y M-GAMBI

De los índices univariantes estudiados arriba, se pasó pronto a tratar de replicar los métodos multivariantes que han sido intercalibrados en Europa para la Directiva Marco del Agua (ver nota 6). Entre ellos, uno de los primeros en desarrollarse fue el **M-gAMBI** usando el número de *reads* en el cálculo (Aylagas et al., 2018). Sin embargo, la correlación encontrada por estos autores entre M-AMBI y M-gAMBI, no fue significativa ($p > 0,05$), achancando este resultado a las diferencias encontradas en diversidad y a la necesidad de determinar condiciones de referencia adecuadas para M-gAMBI.

La determinación de condiciones de referencia para este tipo de índices es clave para obtener una evaluación robusta del estado ecológico, existiendo diversos métodos para su determinación (Borja et al., 2012b).

En el caso de M-AMBI, normalmente se ha utilizado una mezcla de datos históricos, juicio de experto y modelización en la determinación de dichas condiciones (Muxika et al., 2007), para lo que se requiere una buena cantidad de datos disponibles (Santibañez-Aguascalientes et al., 2020).

En el caso de los índices genómicos no es posible la disponibilidad de datos históricos, por lo que se debe tratar de obtener una cantidad suficiente de información que permita hacer una aproximación a las condiciones naturales de las tipologías estudiadas. Es por ello por lo que se propuso realizar este trabajo, si bien es verdad que para algunas tipologías es difícil obtener un número suficiente de información. Así, aunque en total contamos con 125 pares de datos, estos se distribuyen de manera diferente en las cinco tipologías estudiadas: AT-T08 cuenta con 49 pares, AT-T09 con 24, AT-T10 con 17, AC-T12 somero con 24 y AC-T12 profundo con 12.

El método escogido para establecer las condiciones de referencia, usando los valores más elevados de riqueza y diversidad y los más bajos de AMBI en el total de datos, no se considera el más adecuado, aunque sí puede usarse especialmente cuando se conoce que los lugares objeto de estudio no se encuentran afectados por presiones importantes (Borja et al., 2019).

En el caso de la Demarcación Hidrográfica del Cantábrico Oriental, el hecho de que la calidad del medio bentónico haya mejorado ostensiblemente en los últimos 20 años (Borja et al., 2016), hace que sea imposible encontrar todo el gradiente de presión o disturbancia en todas las tipologías.

En este sentido, las estaciones costeras, tanto las referidas a zonas someras como profundas, se encuentran en un buen o muy buen estado ecológico, según las medidas de M-AMBI (Borja et al., 2023), por lo que es muy plausible que las condiciones de referencia establecidas sean cercanas a las que se podrían determinar en un monitoreo a largo plazo.

En el caso de los tres tipos de estuario, posiblemente también sea así, puesto que algunas de las estaciones utilizadas presentan valores de M-AMBI correspondientes a estado muy bueno (incluso con valores superiores a 1), por lo que las condiciones de referencia asociadas a dichas estaciones deben estar también cercanas a las que se determinarían a largo plazo.

Asumiendo que estas condiciones de referencia son las adecuadas, cuando se tiene un gradiente de presión completo o casi completo, como es en los tipos AT-T08 y AT-T10, se tienen correlaciones significativas entre los datos obtenidos por M-AMBI y M-gAMBI, que mejoran considerablemente los datos obtenidos por Aylagas et al. (2018). Según nuestro mejor conocimiento, no existe ninguna publicación más en que se intente aplicar M-gAMBI, posiblemente debido a la falta de condiciones de referencia, por lo que nuestros resultados no se pueden poner en un contexto más amplio.

4.3. ANÁLISIS DE SIMILARIDAD

Los análisis de similaridad llevados a cabo, tanto a nivel de réplica, como a nivel de muestra, arrojan resultados que indican que las muestras analizadas por métodos genómicos y por métodos tradicionales difieren.

Las agrupaciones significativas son homogéneas en cuanto al método empleado en el análisis de las muestras que incluyen, es decir: las principales agrupaciones que se obtienen contienen casi exclusivamente muestras analizadas por uno u otro método.

En parte, esta segregación de muestras en función de la metodología empleada para su tratamiento parece obedecer a la ausencia, en las muestras tratadas por métodos genómicos, de algunas especies características en diferentes comunidades descritas para los estuarios y para la costa del País Vasco (Borja *et al.*, 2004). Es el caso, por ejemplo, de los poliquetos *A. romijni* y *Streblospio* sp. (que incluye las especies *Streblospio shrubsolii* y *Streblospio benedicti*, entre otras), y el crustáceo *C. carinata*, característicos de la comunidad de *Scobicularia plana*-*Cerastoderma edule*, así como los bivalvos del género *Ruditapes*, que incluyen a la almeja fina *Ruditapes decussatus*, característica también de dicha comunidad.

A esta desagregación contribuye también la presencia más o menos importante, en las muestras tratadas por métodos genómicos, de taxones que no pudieron ser identificados o presentaban abundancias prácticamente despreciables en las tratadas por métodos tradicionales, como es el caso de los nematodos.

A pesar de la baja comparabilidad entre los dos conjuntos de datos (el método tradicional vs. el genómico, es decir *metabarcoding*), el patrón en los *conjuntos* de datos, es decir la similitud o disimilitud compartida entre muestras y réplicas individuales, fue altamente congruente entre los dos conjuntos de datos, según un test de Mantel ($R = 0,47$, $p < 0,001$).

Por último, el alto grado de similaridad que presentan las agrupaciones que incluyen muestras analizadas por ambos métodos en idénticas estaciones y campañas de muestreo (salvo la agrupación que incluye muestras tomadas en E-BI20 en 2019 y 202, y en E-OK20 en 2018, 2019 y 2023), viene explicado por las altas dominancias de taxones individuales.

En este sentido, es importante señalar que, por ejemplo, no existen secuencias de referencia taxonómica en las bases de datos para *Alkmaria*, mientras que existen secuencias de COI para especies de los géneros *Cyathura* y *Streblospio* no identificadas en el presente trabajo. En total, de entre las especies identificadas por métodos tradicionales, 152 secuencias estarían ausentes en la base de datos empleada (Midori v253).

En total, de los 733 taxones identificados utilizando el método tradicional basado en la morfología, 311 también se encontraron mediante metabarcoding, así como una rica diversidad de taxones fuera del alcance de la identificación taxonómica tradicional, y no considerados en este estudio.

5.

Conclusiones

En relación con el objetivo propuesto, de si es posible usar métodos genómicos en la evaluación del estado ecológico, existen diversas revisiones que abogan por el uso de metabarcoding en el monitoreo y la evaluación del estado (Bourlat et al., 2013; Leese et al., 2016; Hering et al., 2018; Pawlowski et al., 2018), pero pocos desarrollan índices complejos para realizarlo (Keeley et al., 2018).

Entre los criterios para considerar que un índice basado en genómica es adecuado, Hering et al. (2018) proponen los siguientes:

Criterio de representatividad que implica evaluar la **aplicabilidad** de los medios actualmente disponibles y de métodos alternativos para obtener material biológico; la **detección de errores** en los métodos genómicos y similaridad con los morfológicos; y evaluar la **abundancia** en los métodos genómicos. En nuestro caso, se cumple el criterio de **aplicabilidad** puesto que se han usado los mismos métodos en el morfológico que en el genómico, en el muestreo realizado. En cuanto a la **detección de errores**, en este trabajo se ha sido capaz de detectar errores en la asignación de especies y corregirlos. Sin embargo, para estudiar este tema de detección de errores, hemos realizado un estudio de similaridad entre muestras genómicas y morfológicas, que ha detectado importantes diferencias en la composición. Los problemas asociados podrían venir determinados, al menos en parte, por la ausencia de secuencias de referencia válida en la base de datos empleada. Esto se podría solucionar secuenciando las especies más frecuentes y abundantes e incorporándolas a la base de datos. Por último, en este estudio hemos constatado que el número de *reads* se puede asimilar a la **abundancia** en el morfológico, tal y como se demostró en Aylagas et al. (2018).

Criterio de sensibilidad, es decir, la capacidad de los métodos genómicos para muestrear especies sensibles. Tanto en este estudio como en anteriores (Aylagas et al., 2014, 2018), que se presentan especies de los cinco grupos ecológicos de AMBI, incluyendo el I, que son especies sensibles a la presión.

Reads no asignados: El fragmento “Leray” del gen COI mitocondrial tiene una entropía suficiente para permitir la clasificación de especies para la gran mayoría de los metazoos. Sin embargo, esto depende también de la disponibilidad de secuencias de referencia para las especies en la comunidad estudiada, que aún faltan en una gran parte de las comunidades estudiadas aquí (y para la mayoría del meiobentos; Leray y Knowlton, 2016). Por esta razón, varios de los taxones dominantes, según un enfoque basado en la morfología, faltaban en el conjunto de datos moleculares, y un 12% de los *reads* de metazoos se eliminaron (con excepciones) debido a que no se clasifican a nivel de género o especie. Sin embargo, hemos demostrado que las lecturas y taxones restantes fueron suficientes para estimar M-gAMBI de una manera comparable a M-AMBI, en la mayoría de las muestras. Esta situación puede mejorarse aún más mediante la secuenciación específica de especies localmente dominantes.

Criterio de precisión, que implica conocimiento sobre la incertidumbre de la identificación del método genómico. Los cebadores utilizados también amplifican secuencias no dirigidas, incluidas bacterias y no metazoos, por lo que se eliminó un 24% de las lecturas, que no tenían relación con macroinvertebrados. Una limitación de la precisión de metabarcoding es la posibilidad de introducción de ADN contaminante en las muestras estudiadas desde fuentes externas o entre muestras (contaminación cruzada), lo que sesga especialmente el cálculo de la riqueza y puede afectar especialmente a muestras con baja biomasa. Este problema puede agravarse aún más debido al intercambio de secuencias de índice, durante la secuenciación, dependiendo del protocolo y la tecnología utilizada, a menos que se tomen medidas para controlar y compensarlo. Aquí hemos utilizado una tecnología menos afectada por este problema (Illumina MiSeq) y se han incluido controles negativos, para compensar estos problemas, y hemos descubierto que no ha influido en los resultados en una manera importante. Además, por cobertura insuficiente de secuencias de referencia (ver Sensibilidad arriba), se pueden generar falsos positivos, o la clasificación de secuencias como especies ausentes en el conjunto de datos (Keck et al., 2022).

Criterio de comparabilidad. Consideramos se cumple con este criterio usando un *Ecological Quality Ratio*: en función de los datos encontrados con M-gAMBI, aunque hay diferencias con el método tradicional que se podrían mejorar mediante cambios en los límites de clase.

El objetivo de este trabajo no era llegar a una **intercalibración**, pero creemos que en un futuro próximo podemos establecer una primera intercalibración, que permita ajustar los límites de clase. Mientras se realiza esta intercalibración, es importante constatar que, a pesar de la baja comparabilidad entre abundancias o presencia/ausencia de taxones, entre el método genómico y el tradicional, los patrones en los conjuntos de datos resultantes fueron altamente congruentes entre los dos conjuntos de datos.

Criterio de coste-beneficio. En este trabajo no se ha estudiado el coste comparado con el método tradicional ni una comparativa de la velocidad de procesado, pero Aylagas et al. (2018) lo hicieron y vieron que era aproximadamente el 50% del tradicional.

Criterio de impacto ambiental, es decir, si el método genómico reduce el impacto. En nuestro caso debemos considerar que este impacto es similar, puesto que utiliza muestras destructivas al igual que el morfológico.

Atendiendo a los resultados de este estudio, el **índice M-gAMBI se podría considerar como un índice adecuado para la evaluación del estado ecológico** puesto que reúne la mayor parte de los criterios para considerar si un índice basado en genómica es adecuado (Hering et al. 2018). Sin embargo, en las próximas campañas se plantea como necesario avanzar en al menos los siguientes aspectos:

- En cuanto al **criterio de representatividad** se plantea mejorar la detección de errores en los métodos genómicos y similaridad con los morfológicos. Para ello, estudiaremos las listas actualmente disponibles en GenBank y otros, para ver qué especies presentes en la Demarcación Hidrográfica del Cantábrico Oriental no están en las listas y qué especies en las listas son las más cercanas a las que sí aparecen en la Demarcación Hidrográfica del Cantábrico Oriental. Esto permitirá reducir el número de errores y, eventualmente, aumentar la similaridad entre las muestras tratadas por métodos tradicionales y las tratadas por métodos genómicos.
- En cuanto al **criterio de comparabilidad** se plantea como necesario realizar una intercalibración que permita ajustar los límites de clase y determinar el estado ecológico de manera equivalente utilizando ambos métodos.

Referencias

- Aylagas, E., Á. Borja, N. Rodríguez-Ezpeleta, 2014. Environmental Status Assessment Using DNA Metabarcoding: Towards a Genetics Based Marine Biotic Index (gAMBI). *Plos ONE*, 9: e90529.
- Aylagas, E., Á. Borja, I. Muxika, N. Rodríguez-Ezpeleta, 2018. Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecological Indicators*, 95: 194-202.
- Borja, A., I. Muxika, 2005. Guidelines for the use of AMBI (AZTI's Marine Biotic Index) in the assessment of the benthic ecological quality. *Marine Pollution Bulletin*, 50: 787-789.
- Borja, A., Mader, J., Muxika, I., Rodríguez, J.G., Bald, J., 2008. Using M-AMBI in assessing benthic quality within the Water Framework Directive: Some remarks and recommendations. *Marine Pollution Bulletin* 56, 1377-1379.
- Borja, A., I. Muxika, J. G. Rodríguez, 2009. Paradigmatic responses of marine benthic communities to different anthropogenic pressures, using M-AMBI, within the European Water Framework Directive. *Marine Ecology*, 30: 214-227.
- Borja, Á., J. Mader, I. Muxika, 2012a. Instructions for the use of the AMBI index software (Version 5.0). *Revista de Investigación Marina*, 19: 71-82. https://ambi.azti.es/download/RIM19_3.pdf
- Borja, Á., D. M. Dauer, A. Grémare, 2012b. The importance of setting targets and reference conditions in assessing marine ecosystem quality. *Ecological Indicators*, 12: 1-7.
- Borja, Á., G. Chust, J. G. Rodríguez, J. Bald, M. J. Belzunce-Segarra, J. Franco, J. M. Garmendia, J. Larreta, I. Menchaca, I. Muxika, O. Solaun, M. Revilla, A. Uriarte, V. Valencia, I. Zorita, 2016. 'The past is the future of the present': Learning from long-time series of marine monitoring. *Science of The Total Environment*, 566-567: 698-711.
- Borja, A., G. Chust, I. Muxika, 2019. Chapter Three - Forever young: The successful story of a marine biotic index. *Advances in Marine Biology*, 82: 93-127.
- Borja, A., J. Bald, M.J. Belzunce, M.C. Uyarra, A. Fontán, J. Franco, J.M. Garmendia, A. Lanzén, J. Larreta, I. Menchaca, I. Muxika, S. Pouso, M. Revilla, J.G. Rodríguez, Y. Sagarmínaga, O. Solaun, A. Uriarte, I. Zorita, I. Adarraga, F. Aguirrezabalaga, J.C. Sola, I. Cruz, M.A. Marquegui, J. Martínez, J.M. Ruiz, M. Cano, A. Laza-Martínez, A. Manzanos, 2023. Red de seguimiento del estado ecológico de las aguas de transición y costeras de la Comunidad Autónoma del País Vasco. Informe de resultados. Campaña 2022. Informe de la Fundación AZTI para URA. 386 pp.
- Bourlat, S. J., A. Borja, J. Gilbert, M. I. Taylor, N. Davies, S. B. Weisberg, J. F. Griffith, T. Lettieri, D. Field, J. Benzie, F. O. Glöckner, N. Rodríguez-Ezpeleta, D. P. Faith, T. P. Bean, M. Obst, 2013. Genomics in marine monitoring: New opportunities for assessing marine health status. *Marine Pollution Bulletin*, 74: 19-31.
- Bray, J.R.; Curtis, J.T. 1957. An ordination of the upland forest communities of Southern Wisconsin. *Ecological Monographs*, 27: 325-349.

- Carugati, L., C. Corinaldesi, A. Dell'Anno, R. Danovaro, 2015. Metagenetic tools for the census of marine meiofaunal biodiversity: An overview. *Mar. Genomics*, 24: 11-20.
- Chariton, A. A., S. Stephenson, M. J. Morgan, A. D. L. Steven, M. J. Colloff, L. N. Court, C. M. Hardy, 2015. Metabarcoding of benthic eukaryote communities predicts the ecological condition of estuaries. *Environmental Pollution*, 203: 165-174.
- Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, 18: 117-143.
- Clarke, K.R.; Gorley, R.N. 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E: Plymouth, 190 pp.
- Clarke, K.R.; Warwick, R.M. 2001. Change in marine communities: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E: Plymouth. 174 pp.
- Cordier, T., L. Alonso-Sáez, L. Apothéoz-Perret-Gentil, E. Aylagas, D. A. Bohan, A. Bouchez, A. Chariton, S. Creer, L. Frühe, F. Keck, N. Keeley, O. Laroche, F. Leese, X. Pochon, T. Stoeck, J. Pawlowski, A. Lanzén, 2020. Ecosystems monitoring powered by environmental genomics: a review of current strategies with an implementation roadmap. *Molecular Ecology*, 30: 2937–2958.
- Davis, N. M., D. Proctor, S. P. Holmes, D. A. Relman, B. J. Callahan, 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, 6: 226.
- Deiner, K., J.-C. Walser, E. Mächler, F. Altermatt, 2015. Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation*, 183: 53-63.
- Dias, H. Q., S. Sukumaran, 2023. Are genomic indices effective alternatives to morphology based benthic indices in biomonitoring studies? Perspectives from a major harbour and marine protected area. *Marine Pollution Bulletin*, 187: 114586.
- Duarte, S., P. E. Vieira, B. R. Leite, M. A. L. Teixeira, J. M. Neto, F. O. Costa, 2023. Macrozoobenthos monitoring in Portuguese transitional waters in the scope of the water framework directive using morphology and DNA metabarcoding. *Estuarine, Coastal and Shelf Science*, 281: 108207.
- Frøslev, T. G., R. Kjølner, H. H. Bruun, R. Ejrnæs, A. K. Brunbjerg, C. Pietroni, A. J. Hansen, 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications*, 8 (1): 1188
- Gibson, J. F., S. Shokralla, C. Curry, D. J. Baird, W. A. Monk, I. King, M. Hajibabaei, 2015. Large-Scale Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing. *Plos ONE*, 10: e0138432.
- Hajibabaei, M., T. M. Porter, C. V. Robinson, D. J. Baird, S. Shokralla, M. T. G. Wright, 2019. Watered-down biodiversity? A comparison of metabarcoding results from DNA extracted from matched water and bulk tissue biomonitoring samples. *Plos ONE*, 14: e0225409.
- Hering, D., A. Borja, J. I. Jones, D. Pont, P. Boets, A. Bouchez, K. Bruce, S. Drakare, B. Hänfling, M. Kahlert, F. Leese, K. Meissner, P. Mergen, Y. Reyjol, P. Segurado, A. Vogler, M. Kelly, 2018. Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research*, 138: 192-205.

- Keck, F., R. C. Blackman, R. Bossart, J. Brantschen, M. Couton, S. Hürlemann, D. Kirschner, N. Locher, H. Zhang, F. Altermatt, 2022. Meta-analysis shows both congruence and complementarity of DNA and eDNA metabarcoding to traditional methods for biological community assessment. *Molecular Ecology*, 6: 1820-1835.
- Keeley, N., S. A. Wood, X. Pochon, 2018. Development and preliminary validation of a multi-trophic metabarcoding biotic index for monitoring benthic organic enrichment. *Ecological Indicators*, 85: 1044-1057.
- Lanzén, A., S. L. Jørgensen, D. H. Huson, M. Gorfer, S. H. Grindhaug, I. Jonassen, L. Øvreås, T. Urich, 2012. CREST - Classification Resources for Environmental Sequence Tags. *PLoS ONE*, 7 (11); e49334
- Lanzén, A., T. G. Dahlgren, A. Bagi, J. T. Hestetun, 2021. Benthic eDNA metabarcoding provides accurate assessments of impact from oil extraction, and ecological insights. *Ecological Indicators*, 130: 108064.
- Leese, F., F. Altermatt, A. Bouchez, T. Ekrem, D. Hering, P. Mergen, J. Pawlowski, J. Piggott, K. Abarenkov, P. Beja, L. Bervoets, P. Boets, A. Bones, Á. Borja, K. Bruce, J. Carlsson, E. Coissac, F. Costa, M. Costache, S. Creer, Z. Csabai, K. Deiner, Á. DelValls, S. Duarte, S. Fazi, W. Graf, Y. Hershkovitz, B. Japoshvili, J. Jones, M. Kahlert, B. Kalamujic Strojil, M. Kelly-Quinn, E. Keskin, E. Mächler, A. Mahon, M. Marečková, M. Mejdandzic, M. Montagna, C. Moritz, V. Mulk, I. Navodaru, S. Pálsson, K. Panksep, L. Penev, A. Petrusek, M. Pfannkuchen, B. Rinkevich, A. Schmidt-Kloiber, P. Segurado, M. Strand, S. Šulčius, M. Traugott, X. Turon, A. Valentini, B. van der Hoorn, M. Vasquez Hadjilyra, J. Viguri, A. Vogler, B. Zegura, 2016. DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes*, 2: e11321.
- Lejzerowicz, F., P. Esling, L. Pillet, T. A. Wilding, K. D. Black, J. Pawlowski, 2015. High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. *Scientific Reports*, 5: 13932.
- Leray, M., Yang, Y. J., Meyer, P. C., Mills, C. S., Agudelo, N., Ranwez, V., et al. (2013). New versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* 10:34. doi: 10.1186/ 1742-9994-10-34
- Leray, M. & Knowlton, N. 2016. Censusing marine eukaryotic diversity in the twenty-first century. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 371: 1702.
- Lobo, J., S. Shokralla, M. H. Costa, M. Hajibabaei, F. O. Costa, 2017. DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. *Scientific Reports*, 7: 15618.
- Mahé, F., T. Rognes, C. Quince, C. de Vargas, M. Dunthorn, 2015. SWARM v2: highly-scalable and high-resolution amplicon clustering. *PeerJ*, 3: e1420.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17: 10--12.
- Muxika, I., Á. Borja, J. Bald, 2007. Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water

Framework Directive. Marine Pollution Bulletin, 55: 16-29.

- Pawlowski, J., M. Kelly-Quinn, F. Altermatt, L. Apothéloz-Perret-Gentil, P. Beja, A. Boggero, A. Borja, A. Bouchez, T. Cordier, I. Domaizon, M. J. Feio, A. F. Filipe, R. Fornaroli, W. Graf, J. Herder, B. van der Hoorn, J. Iwan Jones, M. Sagova-Mareckova, C. Moritz, J. Barquín, J. J. Piggott, M. Pinna, F. Rimet, B. Rinkevich, C. Sousa-Santos, V. Specchia, R. Trobajo, V. Vasselon, S. Vitecek, J. Zimmerman, A. Weigand, F. Leese, M. Kahlert, 2018. The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of The Total Environment*, 637–638: 1295-1310.
- Ranasinghe, J. A., E. D. Stein, P. E. Miller, S. B. Weisberg, 2012. Performance of Two Southern California Benthic Community Condition Indices Using Species Abundance and Presence-Only Data: Relevance to DNA Barcoding. *Plos ONE*, 7: e40875 EP.
- Ransome, E., J. B. Geller, M. Timmers, M. Leray, A. Mahardini, A. Sembiring, A. G. Collins, C. P. Meyer, 2017. The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia. *Plos ONE*, 12: e0175066.
- Rognes, T., T. Flouri, B. Nichols, C. Quince, F. Mahé, 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4: e2584.
- Santibañez-Aguascalientes, N. A., Á. Borja, J. L. Montero-Muñoz, M. T. Herrera-Dorantes, P.-L. Ardisson, 2020. Setting reference conditions to assess the ecological status of the sublittoral and bathyal benthic communities of the southern Gulf of Mexico. *Ecological Indicators*, 111: 105964.
- Van den Bulcke, L., A. De Backer, J. Wittoeck, K. Beentjes, S. Maes, M. Christodoulou, P. Martinez Arbizu, R. Sapkota, B. Van der Hoorn, A. Winding, K. Hostens, S. Derycke, 2023. DNA metabarcoding on repeat: Sequencing data of marine macrobenthos are reproducible and robust across labs and protocols. *Ecological Indicators*, 150: 110207.
- van der Loos, L. M., R. Nijland, 2021. Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Molecular Ecology*, 30: 3270-3288.
- Willassen, E., J.-I. Westgaard, J. A. Kongsrud, T. Hanebrekke, P. Buhl-Mortensen, B. Holte, 2022. Benthic invertebrates in Svalbard fjords—when metabarcoding does not outperform traditional biodiversity assessment. *PeerJ*, 10: e14321.