

Andalucía-Biotec Salud

Coordinador: Mario Delgado Mora, Instituto de Parasitología y Biomedicina López-Neyra (IPBLN-CSIC)

Breve descripción de la propuesta y objetivos principales

El proyecto de Biotecnología Aplicada a la Salud (BAS) tiene como foco el desarrollo de tecnologías biomédicas y bioinformáticas que contribuyan a la implantación efectiva de la Medicina de Precisión, una nueva medicina centrada en el individuo y que hará que la diagnosis, tratamiento y prognosis de la enfermedad sea más eficiente en los sistemas de salud nacionales. El objetivo principal de este Plan Complementario requiere de múltiples capacidades tecnológicas y de análisis, así como de su implantación dentro de los sistemas de salud de las Comunidades Autónomas (CCAA).

La propuesta **Andalucía-Biotec Salud** que presentamos como expresión de interés pretende complementar este proyecto desarrollando de forma integrativa las cuatro Líneas de Actuación (LA1-4) identificadas por la CA de Andalucía como prioritarias en el Plan Complementario a nivel nacional. La propuesta Andalucía-Biotec Salud está constituida por siete proyectos de I+D+I más un proyecto de coordinación, dirigidos por investigadores reconocidos a nivel internacional en sus respectivos campos de trabajo y que serán ejecutados en los institutos de investigación y universidades de mayor prestigio en el campo de la biotecnología y biomedicina en Andalucía. Además, Andalucía-Biotec Salud ha conseguido integrar en este proyecto a una gran mayoría de las instituciones (Centros de Investigación, Universidades, Hospitales, Fundaciones Públicas, Institutos Biosanitarios) con un papel relevante en Biotecnología Aplicada a la Salud de todas las provincias andaluzas.

A través del abordaje coordinado de las cuatro líneas de actuación, el objetivo principal de este proyecto es la implementación de nuevas plataformas biotecnológicas transversales, pioneras y singulares enfocadas al desarrollo de herramientas para diagnóstico y terapias avanzadas o dirigidas en medicina personalizada, y a las que tengan acceso en igualdad de oportunidades los distintos centros de investigación, hospitales, universidades y empresas biotecnológicas, para que puedan integrarlas de manera rutinaria en sus investigaciones y en su toma de decisiones (diagnóstico, pronóstico, tratamiento de enfermedades), confiriéndoles ventaja competitiva e innovadora. El objetivo final de Andalucía-Biotec Salud es democratizar la tecnología de vanguardia y que se garantice una ciencia abierta y de calidad en áreas prioritarias para Andalucía como BAS.

A continuación se procede a una descripción breve de los proyectos y actuaciones que se van a llevar a cabo para abordar las distintas líneas de actuación del Plan Complementario Andaluz en Biotecnología de la Salud, y que se detallarán posteriormente con la descripción de los proyectos individualizados en sus distintos apartados.

LÍNEA DE ACTUACIÓN 1: Armonización de bases de datos del SSPA para su integración y análisis conjuntos en la iniciativa nacional.

La LA1 se va a abordar con el proyecto P1 titulado "IMPLEMENTACIÓN DE TECNOLOGÍAS PARA EL DESCUBRIMIENTO Y ANÁLISIS FEDERADO DE DATOS CLÍNICOS Y ÓMICOS", dirigido por el Dr. Joaquim Dopazo Blázquez, de la Fundación Andaluza Progreso y Salud (FPS), uno de los investigadores más relevantes a nivel nacional en campos como Bioinformática Clínica, Big data e inteligencia artificial aplicada a la salud. Como parte del equipo de investigación, el proyecto contará con la participación directa de: la Dra. M Carmen Durán, directora de los Servicios Centrales de Biomedicina de la Universidad de Cádiz (UCA), y responsable de la Unidad de proteómica asociada a la UCA e INIBICA; el Dr. Marcial García Rojo, jefe de servicio de Anatomía Patológica del Hospital Universitario de Jerez de la Frontera (HUJF); la Dra. Dolores Muñozero Muñoz, Subdirectora Técnica Asesora de Gestión de la Información del Servicio Andaluz de Salud, responsable de la gestión y explotación de diversos CMDB y de la Base Poblacional de Salud, y directora funcional del proyecto Big Data de Sistema Sanitario de Andalucía; el Dr. Carlos L Parra-Calderón, Jefe de Sección de Innovación Tecnológica del Hospital Universitario Virgen del Rocío (HUVR) y de Informática de la Salud Computacional del IBiS y miembro del Comité Director de IMPaCT Data; el Dr. Antonio J Pérez Pulido, responsable del grupo de Bioinformática de la Universidad Pablo de Olavide (UPO) y director del Centro de Cálculo Científico de la UPO (C3UPO); el Dr. M Ángel Armengol de la Hoz, responsable del Área de Big Data de la Plataforma de Medicina Computacional de FPS.

Con este proyecto se generará la posibilidad de usar datos clínicos y genómicos armonizados provenientes de la Base Poblacional de Salud andaluza en proyectos cooperativos con otras CCAA, y el alineamiento con el Proyecto Espacio de datos del Ministerio de Sanidad, la Secretaría de Estado de Digitalización e Inteligencia Artificial y la Secretaría General de Administración Digital (Data Lake Sanitario con estructura federada). Se pondrá a punto un sistema que permita localizar datos clínicos y genómicos de interés en las bases de datos, identificar la posibilidad de uso secundario de estos y cuando sea posible, su descarga en un sistema de análisis seguro que funcione al estilo de los Trusted Research Environments del NSH. Además se dotará a la base de datos la capacidad de descubrir de forma eficiente la existencia de pacientes con un perfil genómico compatible con ensayos clínicos y la integración en la base de datos de datos ómicos, incluidos los obtenidos mediante el uso de modelos no humanos. *El proyecto P1 se describe en detalle en las Páginas 7-9 de esta memoria.*

LÍNEA DE ACTUACIÓN 2: Generación, validación y análisis experimental de biomodelos, modelos computacionales y herramientas terapéuticas.

La LA2 se va a abordar con el proyecto P2 titulado "GENERACIÓN Y CONSOLIDACIÓN DE HERRAMIENTAS BIOTECNOLÓGICAS Y DE EDICIÓN GENÓMICA PARA EL DESARROLLO Y USO DE BIOMODELOS" dirigido por el Dr. Miguel Ángel Moreno Mateo, investigador de la UPO en el Centro Andaluz de Biología del Desarrollo (CABD) y experto en herramientas biotecnológicas de edición genética. El resto de participantes de esta propuesta va a estar integrado en una red de colaboración que se va a formar como actividad derivada de este proyecto compuesta por una multitud de investigadores multidisciplinares pertenecientes a varias instituciones de seis de las provincias andaluzas, todos ellos expertos en biomodelos en distintas especies. Este proyecto propone por un lado la creación de un catálogo de aproximaciones biotecnológicas mediante la recopilación, consolidación y obtención de herramientas y protocolos en edición genómica de última generación que sirvan de base para una estructuración interna de los recursos tecnológicos de producción de biomodelos dentro de la CA de Andalucía,

así como su integración a nivel supracomunitario, de forma que estos recursos estén no solo identificados, sino coordinados para dar respuesta al diseño de biomodelos de forma rápida, colaborando con las LA1 y LA3 con la posible generación una plataforma bioinformática de modelos computacionales. Por otro lado, propone la selección y validación de algunos de los biomodelos de enfermedades humanas, que se financiarán como prueba de principio, por propuestas internas de entre los participantes del consorcio de trabajo que se establecerá en paralelo. La selección de los biomodelos validados permitirá incorporar nuevas herramientas a las plataformas de descubrimiento de fármacos y terapias existentes en Andalucía y resto de CCAA desde etapas muy tempranas de la cadena de valor traslacional, reducirá los costes de los análisis (e.g. cribado de drogas) y los tiempos de desarrollo, e incrementarán su tasa de éxito. Además, permitirán reducir el uso de biomodelos animales siguiendo las recomendaciones de la UE. En coordinación con la LA1, se permitirá la inclusión de datos derivados de estos biomodelos en las bases de datos que se generarán de forma armonizada, de manera que contribuyan a generar modelos computacionales predictivos de mayor calidad. *El proyecto P2 se describe en detalle en las Páginas 10-14 de esta memoria.*

LÍNEA DE ACTUACIÓN 3: Caracterización ómica de muestras y biomodelos a escala de célula individual.

Debido a su complejidad y especialmente en base a las oportunidades que los distintos grupos de investigación y servicios de apoyo a la investigación de varias instituciones andaluzas brindan para el desarrollo de ómicas a escala de célula individual, el abordaje de la LA3 se va a realizar a través de cinco proyectos de investigación, íntimamente conectados entre ellos, y organizados en dos nodos geográficos y a la vez temáticos: un nodo situado en Granada focalizado en la generación y puesta a punto de dos plataformas biotecnológicas complementarias dedicadas a transcriptómica espacial de célula única, y otro nodo situado en Sevilla focalizado en la generación y puesta a punto de varias plataformas especializadas en transcriptómica, proteómica y epigenómica de célula individual.

En general, la ejecución de estos proyectos dentro de la LA3 resultará en el desarrollo de protocolos experimentales y bioinformáticos estandarizados, que serán aplicados a problemas identificados de forma coordinada dentro del consorcio BAS en forma de Proyectos Prueba de Concepto específicos que validen las diferentes aproximaciones. Desde un punto de vista organizativo, el desarrollo de estas actividades permitirá la coordinación, dentro de estos proyectos específicos, de las diferentes capacidades tecnológicas que se están implantando a gran velocidad en los centros de investigación de Andalucía. En particular, la generación de una unidad coordinada de análisis de "single cell omics" permitirá el desarrollo dinámico de protocolos de análisis y la extensión de su uso a un número creciente de centros de investigación biomédica/centros hospitalarios que implementen o demanden técnicas ómicas-sc. Los dos objetivos últimos son, por un lado, ofrecer una caracterización y análisis rápidos de muestras de pacientes que ayuden en la toma de decisiones clínicas en tiempo real y un procedimiento de iteración clínica/análisis que posibilite monitorizar la respuesta a esas decisiones, y por otro integrar el conocimiento generado para el descubrimiento y desarrollo de nuevas soluciones terapéuticas y mejorar su traslación a la clínica.

- El proyecto P3 (nodo Granada) titulado "PLATAFORMA DE TRANSCRIPTOMICA ESPACIAL PARA APLICACIONES BIOMÉDICAS" va a ser co-dirigido por la Dra. Fuencisla Matesanz del Barrio, Directora del Instituto de Parasitología y Biomedicina López-Neyra del CSIC (IPBLN-CSIC) y experta internacional en Genética Humana y la Dra. Alicia Barroso-del Jesus, responsable de la Unidad de Genómica del IPBLN-CSIC. Como parte del equipo de investigación multidisciplinar, el proyecto contará para su ejecución con la participación de: el Dr. Francisco O'Valle Ravassa, Catedrático de la Universidad de Granada (UGR) y experto nacional en Anatomía Patológica; el Dr. Pablo Galindo Moreno, Catedrático de la UGR y responsable de la plataforma de transcriptómica espacial recientemente adquirida y que será esencial para la ejecución de este proyecto; el Dr. Raúl Luque Huertas, Catedrático de la Universidad de Córdoba (UCO) y experto en biología de tumores cerebrales en el Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC); la Dra. Laura Terrón Camero de la Unidad de Bioinformática del IPBLN.

El proyecto contempla la implementación y puesta en marcha de una plataforma de transcriptómica espacial basada en la tecnología GeoMx®/nCounter®, infraestructura singular recientemente adquirida, para su utilización en investigación biomédica y clínica, tratando de dar una solución integral que abarca el tratamiento experimental de las muestras clínicas tanto frescas como de archivo y el procesamiento bioinformático de los datos e interpretación de resultados, suponiendo una innovadora herramienta para el desarrollo de procedimientos diagnósticos y terapias más personalizadas. El proyecto constará de tres fases consistentes en la implementación y puesta a punto inicial de la tecnología en un modelo experimental murino, la validación de la tecnología en muestras clínicas humanas de glioblastoma, y finalmente una prueba de concepto en el que la plataforma se abrirá a la comunidad científica desarrollando proyectos externos de interés. *El proyecto P3 se describe en detalle en la Página 15-18 de esta memoria.*

- El proyecto P4 (nodo Granada) titulado "IMPLEMENTACIÓN DE TECNOLOGÍAS DE TRANSCRIPTÓMICA ESPACIAL DE RESOLUCIÓN CELULAR CON APLICACIÓN BIOSANITARIA Y DESARROLLO DE MÉTODOS COORDINADOS DE ANÁLISIS BIOINFORMÁTICOS ESTANDARIZADOS" va a ser co-dirigido por el Dr. Javier Martín Ibáñez, profesor de investigación del IPBLN-CSIC y experto mundial en Genética de enfermedades complejas humanas, y por el Dr. Eduardo Andrés León, responsable de la Unidad de Bioinformática del IPBLN. El equipo de investigación lo compone los siguientes investigadores, esenciales para la ejecución del proyecto: el Dr. Norberto Ortego Centeno, Jefe de la Unidad de Enfermedades Autoinmunes Sistémicas del Hospital Universitario San Cecilio (HUSC, Granada) e investigador responsable del grupo de bases de fisiopatología y terapéutica médica del área de medicina de precisión del IBS-Granada; el Dr. José L. Callejas, médico internista en la unidad de Enfermedades Sistémicas del HUSC; el Dr. Ricardo Ruiz Villaverde, Jefe del Servicio de Dermatología del HUSC y responsable del grupo de Enfermedades Inmunomediadas de la Piel del IBS-Granada; la Dra. Luz Canet Antequera del Servicio de Genómica del IPBLN.

Esta propuesta es complementaria a la que desarrollará el otro proyecto de la LA3 del nodo de Granada y va dirigida a la implementación de una plataforma de transcriptómica espacial de célula única de alta resolución, que proporcione un mejor y más preciso diagnóstico clínico a través de un enfoque integral, en este caso focalizado en enfermedades autoinmunes, en concreto en muestras de piel de pacientes con esclerodermia. Con el objeto de alcanzar la adecuada potencia de análisis para la implementación de esta plataforma, el proyecto propone la adquisición de un sistema de secuenciación "ómica" de resolución espacial mejorada de StereoSeq de MGI®. El otro objetivo de la propuesta

es la creación de protocolos “sc-ómicos” de análisis bioinformáticos enfocados a dar soluciones a necesidades clínicas de forma transversal con las otras propuestas del proyecto Andalucía-Biotec Salud y del Plan Complementario BAS a nivel nacional. *El proyecto P4 se detalla en la Página 18-21 de esta memoria.*

- El proyecto P5 (nodo Sevilla) titulado “IMPLEMENTACIÓN DE TECNOLOGÍAS DE TRANSCRIPTÓMICA A NIVEL DE CÉLULA ÚNICA A PARTIR DE MUESTRAS CLÍNICAS” va a ser dirigido por el Dr. Javier López-Ríos Moreno, Investigador del CABD-CSIC, responsable de su Unidad de Genómica y experto en la tecnología de transcriptómica a nivel de célula única. El resto del equipo de investigación está formado por estrechos colaboradores del IP en el CABD: el Dr. Ildelfonso Cases, coordinador de la Unidad de Bioinformática del CABD y experto en biología computacional; y la Dra. Ana Fernández Miñán, responsable científica del Servicio de Genómica Funcional del CABD.

El objetivo principal de esta propuesta es la implementación de protocolos robustos que permitan acceder a estudios de tipo de transcriptómica de célula única a partir de muestras humanas de diferente origen (biopsias, piezas quirúrgicas, xenoinjertos derivados de pacientes), incluyendo muestras complejas de procesar, como las fijadas e incluidas en parafina, pero de alta importancia a nivel clínico. Para ello contempla la adquisición de un nuevo equipamiento (Chromium iX), único para este tipo de procesamiento. La prueba de concepto se realizará con xenoinjertos de sarcoma de Ewing, que serán proporcionados por colaboradores integrados en el proyecto P7 de este nodo. En línea y en coordinación con las propuestas del nodo Granada, otro objetivo principal de este proyecto es la generación de protocolos bioinformáticos de fácil implementación que permitan el procesamiento de datos de secuenciación masiva de genotecas de transcriptómica de célula única y otras ómicas y la generación de informes para la práctica clínica sobre el perfil molecular de la muestras de manera que puedan resultar informativos para la diagnosis, prognosis y elección de alternativa terapéutica. *El proyecto P5 se describe en detalle en la Página 22-24 de esta memoria.*

- El proyecto P6 (nodo Sevilla) titulado “IMPLEMENTACIÓN DE TECNOLOGÍAS DE ACCESIBILIDAD DE LA CROMATINA A NIVEL DE CÉLULA ÚNICA EN ESTRATEGIAS DE MEDICINA PERSONALIZADA” será dirigido por el Dr. José Carlos Reyes del Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER) y Jefe del Departamento de Biología del Genoma, que cuenta con una dilatada experiencia en regulación genética y epigenética del cáncer. El resto del equipo de investigación lo componen: el Dr. Andrés Aguilera, Catedrático de Genética de la Universidad de Sevilla, Director de CABIMER y responsable de su Unidad de Genómica; y la Dra. Vivian Capilla González, Investigadora de la FPS en CABIMER y especialista en terapias contra enfermedades neurológicas y glioblastomas.

El objetivo principal de este proyecto es la implementación de una plataforma singular basada en la tecnología denominada ATAC-seq de célula única (scATAC-seq) que permite el conocimiento de las regiones accesibles del genoma, en su mayoría regiones reguladoras, cuyo conocimiento es clave para comprender la heterogeneidad genética y epigenética de las células tumorales, origen a su vez de comportamientos metastáticos y resistencias a tratamientos. De nuevo, la elaboración de protocolos estandarizados en la aplicación de esta tecnología y el análisis bioinformático de los datos resultantes estará en el centro de la propuesta. Se cuenta con la infraestructura necesaria para llevar a cabo esta propuesta y se usará, como en el proyecto P5, xenoinjertos derivados de pacientes (sarcoma de Ewing y glioblastomas) así como como prueba de concepto. *El proyecto P6 se describe en detalle en la Página 25-27 de esta memoria.*

- El proyecto P7 (nodo Sevilla) titulado “IMPLEMENTACIÓN DE TECNOLOGÍAS DE PROTEÓMICA A NIVEL DE CÉLULA ÚNICA CON POTENCIAL APLICACIÓN EN MEDICINA PERSONALIZADA” va a ser dirigido por el Dr. Alberto Pascual del Instituto de Biomedicina de Sevilla (IBiS), responsable científico del Servicio de Proteómica del IBiS. El resto del equipo de investigación lo conforman: el Dr. Enrique de Álava, Jefe de Servicio de Anatomía Patológica del HUVR, investigador principal del IBiS, coordinador del nodo de Biobanco Andaluz del HUVR, y experto reconocido en patología molecular de sarcomas; y la Dra. Marisa Mateos Martín, responsable técnico del Servicio de Proteómica del IBiS.

El objetivo de la propuesta es desarrollar la potencialidad de la plataforma de proteómica de célula única del IBiS a partir de muestras humanas de diferente origen (biopsias, piezas quirúrgicas, xenoinjertos y avatares derivados de pacientes). A la finalización del proyecto, la plataforma tecnológica resultante dispondrá de las capacidades científico-técnicas para extender sus servicios a centros de investigación biomédica y hospitalarios que precisen de este tipo de tecnología en estrategias de medicina de precisión. *El proyecto P7 se describirá en detalle en la Página 27-29 de esta memoria.*

LÍNEA DE ACTUACIÓN 4: Unidad de Coordinación.

La coordinación de las propuestas de Andalucía-Biotec Salud va a ser llevada a cabo por una Unidad dirigida por el Dr. Mario Delgado Mora, Profesor de Investigación del IPBLN-CSIC, con amplia experiencia en el campo de la Biomedicina y en la dirección de grandes proyectos e iniciativas de transferencia en salud. Por ejemplo, ha dirigido el IPBLN durante ocho años, ha coordinado el Libro Blanco del CSIC enfocado a los retos en Biomedicina en los próximos 30 años, y fue uno de los responsables científicos de la llegada al mercado de una terapia avanzada basada en el uso de células madre alogénicas para el tratamiento de una enfermedad autoinmune (Alofisel). La Unidad de Coordinación se completará con la incorporación con cargo a este proyecto de una persona experta en manejo de base de datos que dará apoyo a la gestión, y de un experto en comunicación y difusión científica.

Los objetivos principales de la Unidad de Coordinación consistirán en:

- Coordinar las acciones de las instituciones Andaluzas participantes en Andalucía-Biotec Salud para hacer más eficiente el uso de los recursos asignados, dando cobertura a las actividades de las LA1-3, siguiendo periódicamente el progreso de las actividades programadas, y en su caso proponer actuaciones que vayan encaminadas a corregir posibles desviaciones del cronograma propuesto. Para ello la unidad de coordinación se reunirá de forma telemática trimestralmente con los responsables científicos de cada uno de los subproyectos de forma individual, y semestralmente con todos los responsables de forma coordinada (con al menos una reunión anual de forma presencial). Se elaborarán informes que recojan las conclusiones de dichas reuniones que serán difundidos entre los participantes en la propuesta y servirán de base para los informes intermedios de evolución de Andalucía-Biotec Salud que se emitirán semestralmente. Este objetivo de coordinación es esencial para establecer sinergias entre las distintas líneas de actuación, entre los distintos grupos que participan en un nodo, y entre los dos nodos de la LA3, en vistas a establecer protocolos unificados y coordinar actividades de formación conjunta.

- Actuar como interlocutor entre los agentes de las CCAA participantes en BAS y las instituciones de la CA de Andalucía que forman Andalucía-Biotec Salud. El coordinador científico acudirá como representante de Andalucía a las reuniones del Comité Científico dentro del Comité de Coordinación Estatal del Plan Complementario BAS, convocadas generalmente cada 4-6 meses. Con anterioridad y posterioridad a las reuniones del comité estatal, se celebrará una reunión de la unidad de coordinación con los responsables de los distintos proyectos para actualizar el estado de progresión de cada línea de actuación, identificar potenciales problemas y cambios en el desarrollo de las mismas, unificar posturas y discutir los asuntos que se llevan al comité de coordinación estatal concernientes a Andalucía-Biotec Salud, así como la información que sea necesaria transmitir al tejido científico y sanitario andaluz derivadas de esas reuniones (ej. convocatorias de actividades de otras CCAA donde puedan participar instituciones andaluzas, y viceversa). Además, la unidad de coordinación se apoyará en un comité técnico que representará a Andalucía-Biotec Salud en las comisiones técnicas estatales de las distintas LAs del BAS que se celebran periódicamente cada 3-4 meses. Estas líneas de actuación no coinciden totalmente con las establecidas en la CA de Andalucía, y algunas de las estatales no se han definido como tales en el Plan Complementario de Andalucía, pero para las que Andalucía-Biotec Salud tendrá representantes a nivel estatal. Este comité técnico estará formado en parte por varios de los IPs de los proyectos que abordan las LA1-3 de Andalucía-Biotec Salud y en parte por destacados expertos en los campos temáticos que cubren las LAs estatales: el Dr. Francisco O'Valle (UGR) como representante en la LA1 estatal (Plataforma BREM de criome aplicada a la medicina personalizada), los Drs. Joaquim Dopazo y Javier López-Ríos como representantes en la LA2 estatal (Implementación y análisis de bases de datos en medicina de precisión), la Dra. Olga Genilloud (Fundación MEDINA) como representante de la LA3 estatal (Plataforma de cribado de fármacos y análisis interacciones farmacodiana), el Dr. Miguel A. Moreno como representante en la LA4 estatal (Desarrollo de modelos biológicos para cribado y estudio de la actividad de moléculas terapéuticas), la Dra. Amelia Aránega (Universidad de Jaén) como representante de la LA5 estatal (Desarrollo de nanofármacos, biodistribución, toxicidad y acciones terapéuticas en modelos de patología), y el Dr. Mario Delgado como representante de la LA6 estatal (Técnicas y procesos para terapias avanzadas y dirigidas, formación quirúrgica y robótica médica). Del mismo modo a lo propuesto para la reunión del comité de coordinación estatal, la unidad de coordinación de Andalucía-Biotec Salud se reunirá con los representantes del comité técnico antes y después de la reunión del comité técnico estatal para actualizar la información que se tratará o se derive de dichas reuniones.

- Centralizar la coordinación de las acciones de protección intelectual, tratando de identificar posibles resultados/productos/tecnologías patentables y organizar con las distintas instituciones y sus OTTs cómo llevar a cabo la protección y explotación de los mismos y estudiar la potencialidad de generar empresas de carácter tecnológico derivadas de las estructuras científico/tecnológicas que los generen. Las propuestas de las tres líneas de actuación son en principio susceptibles de tales acciones, tanto en la implementación de herramientas de armonización y análisis masivo de bases de datos, como en la generación de nuevos biomodelos y la obtención de resultados que se deriven de la biotecnología basada en transcriptómica, epigenómica y proteómica de célula única tanto de modelos como de muestras de pacientes de diversas patologías. Para ello, desde la unidad de coordinación se realizará con cada uno de los responsables científicos de los proyectos una actualización periódica trimestral del estado de las distintas plataformas biotecnológicas en materia de novedad y actividad inventiva de todos los avances y resultados que se produzcan.

- Coordinar la comunicación de Andalucía-Biotec Salud a varios niveles: 1) interno entre los participantes en las distintas líneas de actuación de Andalucía-Biotec Salud, 2) supracomunitario con los comités de coordinación, científicos y técnicos estatales del Plan Complementario BAS, y 3) externo con los distintos agentes e instituciones interesadas en las plataformas biotecnológicas que se estén implementando, tanto en entornos profesionales, sanitarios y académicos (centros de investigación, universidades, hospitales e industria farmacéutica/biotecnológica) como a nivel de sociedad y público en general. El propósito es centralizar toda la información de interés relativa a las actividades y resultados tanto del Plan Complementario BAS a nivel estatal como a nivel autonómico en los canales de comunicación previstos para los agentes interesados, fomentando la participación y la mejor gestión. Además, se pretende fomentar la participación ciudadana, mediante la organización de actividades en el ciclo de eventos de divulgación científica andaluz: Noche Europea de los Investigadores, Semana de la Ciencia, efemérides científicas... Se realizarán trabajos de divulgación del conocimiento científico producido a través de las líneas de actuación, mediante notas de prensa y marketing de contenidos sobre las actividades y resultados de las investigaciones en medios especializados y generalistas. Para centralizar las actividades que se lleven a cabo y que la respuesta de comunicación y difusión de Andalucía-Biotec Salud sea uniforme y homologada se van a definir protocolos de comunicación y divulgación de la producción científica que se realice, mediante la elaboración al inicio del proyecto de una guía de Divulgación Científica y de Comunicación Científica en el Plan Complementario de Andalucía-Biotec Salud. Para garantizar el cumplimiento de publicación de las fuentes financiadoras y organismos participantes se distribuirá una guía de obligado seguimiento en materia de comunicación, que incluya entre otras cosas instrucciones de como publicitar actividades, iniciativas, convocatorias, recursos humanos, uso de logotipos, y comunicación de resultados en conferencias, publicaciones,..., Será crítico garantizar la difusión de información sobre las convocatorias y expresiones de interés para atraer a solicitantes, de forma que los investigadores tengan la información pertinente, creando unas plantillas de difusión de los requisitos del proceso y unas estrategias de difusión dinámicas dirigidas a los públicos correctos. Se establecerá un punto de centralización y coordinación de la información científica producida para vehicular en los respectivos canales de comunicación autonómicos y estatales del Plan Complementario. La unidad de coordinación centralizará la gestión y ejecución de los canales de comunicación definidos en el Plan de Comunicación. Se proponen los siguientes canales de comunicación: 1) la elaboración de una página web Andalucía-Biotec Salud (elaborada a través del gestor de contenidos WordPress al inicio del proyecto) que constará de secciones que incluyan información sobre el proyecto (descripción de los objetivos, líneas de actuación, equipo de investigación y centros participantes,...), noticias y divulgación científica, resultados (presentación de los logros del proyecto, con la información más relevante en forma de publicaciones y materiales visuales), convocatorias y empleo; 2) el uso de los canales estatales establecidos en redes sociales y página web (<https://planescomplementariosalud.es/>, <https://twitter.com/PpccSalud>, <https://www.linkedin.com/company/ppccsalud/>) y de canales en redes sociales específicamente creadas y gestionadas por Andalucía-Biotec Salud; 3) canales a nivel de centro participante (redes sociales y webs de los centros participantes); 4) eventos públicos (realización de seminarios en centros con interés en el proyecto, así como en congresos, para aumentar el conocimiento entre la comunidad científica); 5) medios de comunicación tradicionales, para llegar a una audiencia amplia y diversa; 6) *newsletters* (boletines electrónicos periódicos con la información más relevante para los interesados). El responsable de la comunicación y difusión de Andalucía-Biotec Salud será el representante en el comité de comunicación estatal, que se reúne mensualmente para la actualización periódica de los progresos del Plan de Actuación y el

anuncio de hitos, convocatorias, noticias científicas..., y además mantendrá contactos periódicos (semanales o bisemanales) con los responsables científicos de los proyectos para mantener una actualización de la información relevante que se desee o pueda (reservándose aquello que puede ser protegido intelectualmente) comunicar.

Estos objetivos se abordarán de forma continuada a lo largo de todo el periodo de ejecución del proyecto, que será de 30 meses, con los plazos marcados más arriba en cuanto a reuniones periódicas y realización de informes intermedios/finales y guías.

Presupuesto de la Unidad de Coordinación (LA4)

Personal (100.000€): El mayor gasto en la creación de la Unidad de Coordinación va a corresponder al capítulo de contratación de personal de apoyo. Se prevé la contratación de una persona con titulación superior y con experiencia contrastada en comunicación y divulgación científica para coordinar todas las tareas de comunicación y difusión de Andalucía-Biotec Salud a todos los niveles (24 meses, a tiempo completo, 70.000€), así como de una persona con titulación superior en ciencias de la vida y con experiencia en gestión, análisis de datos y transferencia de tecnología (18 meses, a tiempo parcial, 30.000€). Ambas serán imprescindibles para dar apoyo al coordinador de la propuesta en la Unidad de Coordinación en sus dos grandes objetivos: coordinación y comunicación a nivel interno y supracomunitario.

Equipamiento y fungible (3.000€): Se prevé la adquisición de dos ordenadores portátiles para el trabajo que realizarán los responsables de apoyo a la coordinación y de comunicación (1.800€). Gastos en marketing físico (folletos, enaras publicitarias y carteles, 400€) y materiales para las obligaciones de comunicación (500€). Gastos de creación y mantenimiento de página web y redes sociales (700€).

Otros gastos: 3.000€. Gastos de participación en ferias y eventos de ciencia (transporte, materiales y dietas, 700€), de auditoría (1.200€), y de viajes de para la coordinación intergrupos y con el comité de coordinación estatal, al menos una vez al año (2.100€).

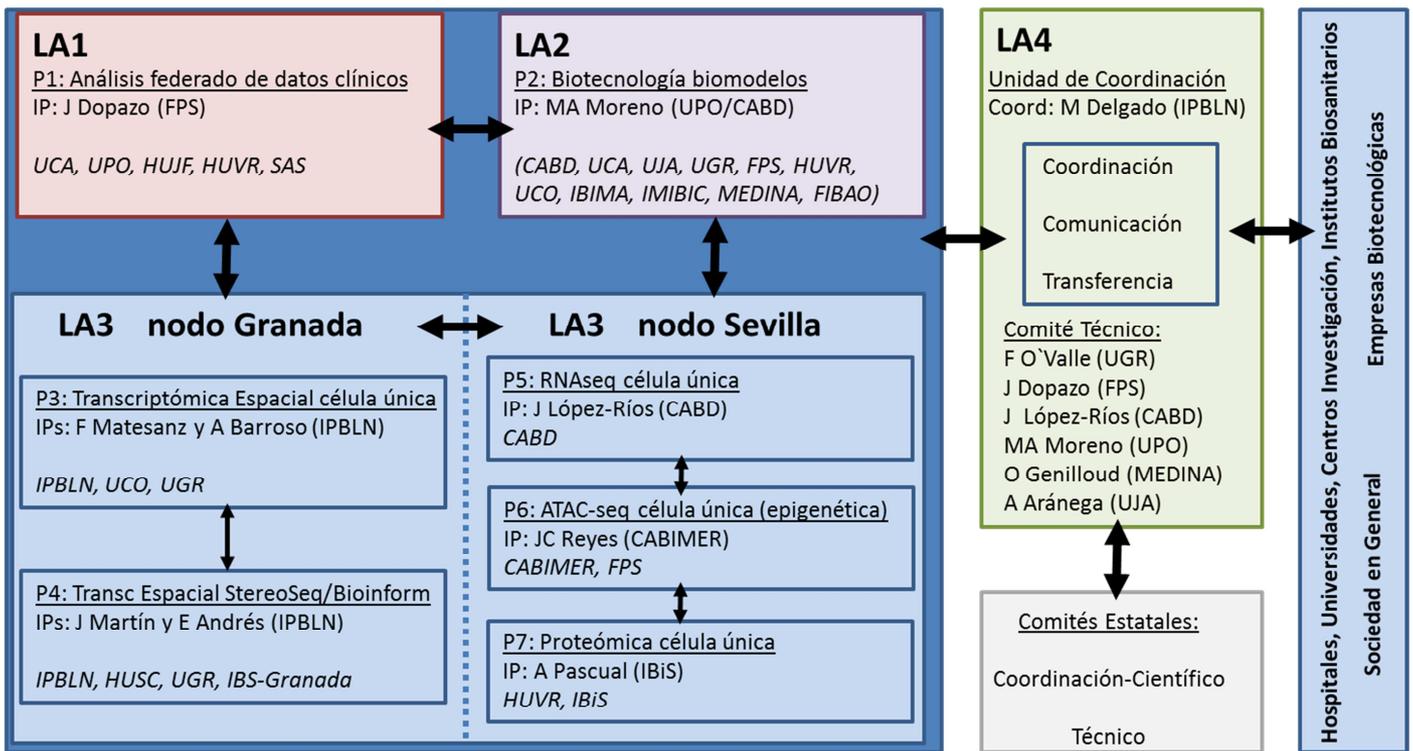
Costes indirectos (15% costes directos): 15.900€. **PRESUPUESTO TOTAL: 121.900€.**

Presupuesto de la Propuesta Integrada: El presupuesto global para la ejecución de Andalucía-Biotec Salud se desglosa en la siguiente Tabla en subproyectos para cada LA y en conceptos. El 50% del presupuesto se dedicará a la contratación de 10 Titulados superiores o Ingenieros y 3 Técnicos de Apoyo. El 47% del presupuesto se dedicará a la adquisición de nuevo equipamiento (equipo informático para instalar aplicaciones de uso federado de datos, secuenciador SteroSeq-MGI, equipo Chromium iX para procesamiento de célula única fijada) que resultará crítico para el establecimiento de las plataformas biotecnológicas propuestas en las distintas LAs, así como en el fungible para la puesta a punto de las nuevas plataformas y la financiación de proyectos para pruebas del concepto. El 3% del presupuesto se dedicará a otros gastos no contemplados en los conceptos anteriores (difusión y comunicación de resultados, viajes/reuniones de coordinación, envío de muestras, auditorías). El 15% de gastos indirectos viene marcado por la convocatoria. Este reparto entre los distintos conceptos se ajusta en gran medida al sugerido en la convocatoria, con algunas desviaciones en casos específicos debido a la naturaleza propia del subproyecto, y que han sido debidamente justificados en la memoria en el apartado correspondiente de cada uno de ellos.

Línea actuación/ Proyecto	Personal	Equipamiento Fungible	Otros gastos	15% Costes indirectos	Total
LA 1/P1	200.000€	30.000€	3.000€	34.950€	267.950€
LA 2/P2	100.000€	76.000€	14.130€	28.520€	218.649€
LA 3/P3	61.250€	100.000€	1.000€	24.337€	186.587€
LA 3/P4	52.500€	193.620€	1.500€	37.143€	284.763€
LA 3/P5	82.000€	160.000€	15.000€	38.550€	295.550€
LA 3/P6	35.000€	36.750€	5.250€	11.550€	88.550€
LA 3/P7	35.000€	41.000€	1.000€	11.550€	88.550€
LA 4/Coor	100.000€	3.000€	3.000€	15.900€	121.900€
TOTAL	665.750€	640.370€	43.750€	202.629€	1.552.499€

A continuación se incluye un **Esquema de organización de Andalucía-Biotec Salud**, en el que se describen los proyectos que se ejecutarán en cada línea de actuación (LA1-4), con los correspondientes investigadores principales y las entidades participantes en el equipo de investigación (siglas en cursiva, y en el caso del P2 como potenciales participantes en los proyectos que se financien). Las flechas indican interacciones y coordinaciones bidireccionales entre proyectos, nodos y LAs, así como los flujos a la hora de generar protocolos unificados, actividades de formación y comunicación. Se incluye la coordinación con los comités científicos y técnicos a nivel estatal, así como los canales de comunicación y difusión con los potenciales actores interesados en el proyecto tanto a nivel profesional como académico y social.

Andalucía-Biotec Salud



Proyecto 1: IMPLEMENTACIÓN DE TECNOLOGÍAS PARA EL DESCUBRIMIENTO Y ANÁLISIS FEDERADO DE DATOS CLÍNICOS Y ÓMICOS

IP: Joaquim Dopazo Blázquez; Institución: Fundación Andaluza Progreso y Salud

- **Resumen de la propuesta:** El objetivo principal de esta propuesta es posibilitar el descubrimiento, interoperabilidad y reusabilidad de los datos disponibles en grandes bases de datos del SSPA y otras bases de datos generadas en proyectos de investigación para acelerar y aumentar las capacidades de generación de conocimiento biomédico. Para ello se desarrollará, reutilizará y desplegará software open que permita consultar catálogos de distintos tipos de datos, seleccionarlos y poner en marcha una infraestructura que permita analizarlos de forma segura.

- **Antecedentes del proyecto**

En el momento actual, en el que la digitalización y el desarrollo vertiginoso de las tecnologías de generación de datos han abaratado su producción, el verdadero reto está en la integración, manejo e interpretación de los datos para generar nuevo conocimiento. La armonización e integración de datos (historias médicas, datos biométricos, datos ómicos, imagen, modelos animales, etc.) es esencial para poder utilizar todo el poder del *Big data* y de la Inteligencia Artificial (IA), no solo dentro de Andalucía sino en colaboración con otras CCAA y proveedores sanitarios privados. De hecho, en los últimos años se han puesto en marcha diversas iniciativas que van en ese sentido, como el proyecto IMPaCT-data¹ del ISCIII, el Data Lake Sanitario español², y a nivel europeo el *European Health Data Space*³ y el proyecto EHDEN⁴, la infraestructura EUCAIM de imagen de cáncer federada o el *Beyond 1 million genomes*⁵ y su proyecto de implementación GDI (Genomic Data Infrastructure). Todas estas iniciativas están dentro de un marco de Ciencia Abierta que se articula en la estrategia EOSC (European Open Science Cloud) que marcará directrices, políticas y requisitos a todos los proyectos de datos de salud europeos previamente indicados, en lo que respecta al uso en investigación, y en el que los principios FAIR⁶ (por sus siglas en inglés: *Findable, Accessible, Interoperable and Reusable*) proporcionan directrices para mejorar los datos de modo que sean Localizables, Accesibles, Interoperables y Reutilizables con el objetivo de garantizar que los datos puedan localizarse, accederse, interoperar y reutilizarse por parte de personas y máquinas.

Solo en imagen médica, en Andalucía se generan cada año 11 millones de pruebas radiológicas y más de 2 millones de preparaciones histológicas. La Base Poblacional de Salud⁷ (BPS) de Andalucía es una base de datos estructurada con las historias clínicas de más de 13 millones de pacientes a la que se está empezando a añadir datos genómicos tanto de pacientes como de patógenos (incluyendo unos 30.000 genomas de SARS-CoV-2 y más de 1000 genomas bacterianos procedentes del circuito de vigilancia epidemiológica⁸). Anualmente, se incluyen un millón de diagnósticos codificados en SNOMED CT en los informes de anatomía patológica. Además, se dispone de un innovador entorno de investigación de confianza (*Trusted Research Environment*⁹), la plataforma iRWD¹⁰, para el análisis seguro de datos protegidos por la RGPD dentro de la red corporativa del SSPA.

- **Objetivos del proyecto**

El objetivo principal de este proyecto es la armonización, integración y FAIRificación de las bases de biodatos (incluyendo datos ómicos de distintos tipos, clínicos, de imagen, de modelos animales, etc.), que es el requisito esencial para poder explotar los datos contenidos en estas y optimizar la capacidad de generar nuevo conocimiento biomédico. Este objetivo aplica tanto a las bases de datos generadas para proyectos de investigación específicos como, especialmente, a las grandes bases de datos existentes en nuestro sistema de salud, en particular BPS, las bases de fármacos tanto prescritos por receta como los prescritos en el ámbito hospitalario, el PACS de imagen o la futura base de datos genómica de Andalucía, para permitir su uso secundario. Aunque BPS usa algunos estándares y ontologías, no está en un formato estándar que haga inmediata su interoperabilidad. Para ello se plantea desarrollar una pasarela que permita su armonización, mapeando su contenido a un modelo estándar de datos comunes. Entre los estándares se propone OMOP¹¹ por su adopción masiva, extensión y uso en otros proyectos.

El segundo objetivo consiste en el desarrollo de un sistema de descubrimiento similar a los estándares Beacon¹² usados por la Global Alliance for Genomics and Health (GA4GH), tanto para datos genómicos, que es su objetivo inicial, como para datos de imagen o clínicos.

¹ <https://impact-data.bsc.es/>

² <https://espanadigital.gob.es/lineas-de-actuacion/data-lake-sanitario>

³ https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_es

⁴ <https://www.ehden.eu/>

⁵ <https://b1mg-project.eu/>

⁶ <https://pubmed.ncbi.nlm.nih.gov/26978244/>

⁷ <https://www.sspa.iuntadeandalucia.es/servicioandaluzdesalud/profesionales/sistemas-de-informacion/base-poblacional-de-salud>

⁸ https://www.clinbioinfosspa.es/COVID_circuit/

⁹ <https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england>

¹⁰ <https://www.clinbioinfosspa.es/projects/iRWD/>

¹¹ <https://www.ohdsi.org/data-standardization/the-common-data-model/>

¹² <https://beacon-project.io/>

El tercer objetivo consiste en el desarrollo de las herramientas que posibiliten el uso de estos datos. Dado que se dispone de la plataforma iRWD para el análisis seguro de datos protegidos por la RGPD, se propone construir sobre esta un sistema que permita el uso de metodologías de IA federadas (*federated machine learning*) para el desarrollo de proyectos concretos, permitiendo así alinear esta estructura con la filosofía de análisis federado de datos promovida por los proyectos IMPaCT-data, FAIR4Health y el Espacio Español de datos de Salud, así como el *European Health Data Space* y el proyecto EHDEN. Finalmente se propone construir un repositorio de procedimientos de análisis, que vayan desde la extracción y la codificación de los datos hasta los pipelines de análisis, que sea consultable en el contexto de proyectos debidamente aprobados, para poder replicar experimentos realizar experimentos similares re-aprovechando el código y pipelines que ya han demostrado su utilidad en otros experimentos, así como la propia experiencia del equipo investigador, como por ejemplo software de ELIXIR, de FAIR4Health, etc.

Todas las aplicaciones desarrolladas para los objetivos de este proyecto se harán en código abierto y siguiendo estándares y guías de buenas prácticas de programación¹³.

- **Contribución del proyecto a la línea de actuación a la que se adscribe**

Este proyecto enmarca en la LÍNEA DE ACTUACIÓN 1: “Armonización de bases de datos del SSPA para su integración y análisis conjunto en la iniciativa nacional”, proporcionando una herramienta para hacer interoperables los datos de las BBDD del SSPA (BPS, las base de fármacos tanto prescritos por receta como los prescritos en el ámbito hospitalario, PACS de imagen y genética), y cualquier otra base de datos de proyectos de investigación, permitiendo su transformación a estándares fácilmente federables (LA1.1). En particular se harán “descubribles” usando estándares construidos alrededor de Beacon (LA1.2 y LA1.3), y se trabajará en un entorno de investigación de confianza en el que se puedan hacer estudios federados (LA1.4).

Este subproyecto tiene un compromiso expreso de colaboración con las actuaciones conjuntas que emanen del Proyecto de Dirección Científica, tanto en Andalucía como a las actuaciones conjuntas con otras CCAA.

- **Metodología y plan de trabajo. Cronograma.**

Se trata de una propuesta metodológica cuyos objetivos implican el desarrollo de software open para facilitar el descubrimiento, interoperabilidad y reusabilidad de los datos disponibles en grandes bases de datos del SSPA y otras bases de datos generadas en proyectos de investigación. El plan de trabajo será llevado a cabo por los participantes y se compone de las siguientes tareas:

Tarea 1. Desarrollo de una pasarela que transforme los datos clínicos en modelo OMOP. Los modelos estándar permiten la interoperabilidad y facilitan el uso de los datos en proyectos federados. La pasarela permite una transformación de los datos sin necesidad de cambiar la estructura inicial de la base de datos, pero además proporciona una información clave para futura transformación de la estructura de la base de datos en caso de plantearse. Duración 4T.

Tarea 2. Desarrollo de un catálogo consultable siguiendo los estándares de IMPaCT-data para el descubrimiento de datos clínicos sin revelar datos individuales. El catálogo permitirá conocer la existencia de datos susceptibles de ser incluidos en estudios y permitirá a los usuarios crear sinergias dentro del sistema. Esto implica una primera fase de evaluación de impacto en protección de datos para el diseño de las posibles consultas y una segunda fase de metadatos para la FAIRificación y desarrollo de la aplicación para realizar las consultas. Duración: 4T.

Tarea 3. Desarrollo de un sistema Beacon según estándares GA4GH para el descubrimiento de datos ómicos (genómicos, proteómicos, etc.). El sistema Beacon permite el descubrimiento sin revelar datos individuales y permite evaluar la posibilidad de incluirlos en estudios cooperativos en función de los consentimientos para su uso secundario. Este sistema incluirá funcionalidades específicas para la detección de pacientes potencialmente utilizables en ensayos clínicos. Duración: 2T

Tarea 4. Desarrollo de un sistema de almacenamiento estándar de datos ómicos (genómicos, proteómicos, etc.) basado en los estándares de EGA, alineado con las recomendaciones de IMPaCT-data. Este sistema de almacenamiento estará conectado con el Beacon de la Tarea 2 de forma automática. Duración: 2T

Tarea 5. Puesta a punto de un sistema computacional en iRWD que permita el análisis de datos clínicos federado, incluyendo técnicas de inteligencia artificial federada. El sistema permitirá descargar datos armonizados de BPS y de EGA en el entorno seguro de iRWD e instalar e implementar algoritmos que permitan análisis federados. Duración: 6T

Tarea 6. Desarrollo de un repositorio de algoritmos y métodos de análisis usados, que sea consultable. Permitirá poder replicar experimentos y reaprovechar código y pipelines que ya han demostrado su utilidad en otros experimentos según la filosofía FAIR y los estándares de reproducibilidad en ciencia. Duración: 2T

Cronograma (distribución de Tareas en Trimestres):

¹³ <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1009481>

Tareas	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
1. Desarrollo de pasarela de conversión a OMOP										
2. Desarrollo de catálogo de datos clínicos consultable										
3. Desarrollo de sistema Beacon para descubrimiento de datos genómicos										
4. Desarrollo de sistema de almacenamiento estándar de datos genómicos										
5. Desarrollo de sistema computacional en iRWD que permita análisis de datos federado										
6. Desarrollo de repositorio consultable de algoritmos y métodos de análisis										

- **Grado de innovación de la propuesta y justificación del potencial impacto científico y/o tecnológico del proyecto.**

Se trata de una propuesta altamente innovadora en cuanto a su implementación ya que sería la primera implementación tecnológica real que permita articular a nivel de sistema de salud la posibilidad de usar sus datos en proyectos de investigación debidamente aprobados, tanto internos como en colaboración con otros proyectos de investigación, o con sistemas de salud de otras CCAA de forma federada. Constituirá un pilar fundamental para el descubrimiento de nueva información en enfermedades no caracterizadas genéticamente, así como en la evaluación de nuevos fármacos en pacientes con un genotipo heterogéneo, apoyando la futura medicina personalizada.

- **Resultados esperados a los que contribuye dentro de la línea de actuación. Plan de difusión y explotación.**

Proporcionar la posibilidad de usar datos clínicos y ómicos (genómicos, proteómicos, etc.) armonizados y FAIRificados provenientes de la Base Poblacional de Salud andaluza en proyectos cooperativos con otras CCAA. Posibilidad de participar de forma federada en el Proyecto Espacio de datos del Ministerio de Sanidad (Data Lake Sanitario con estructura federada). Puesta a punto de un sistema que permita localizar datos clínicos y genómicos de interés en las bases de datos, identificar la posibilidad de uso secundario de estos y cuando sea posible, descargar en un sistema de análisis seguro. Además, se dotará a la base de datos con la capacidad de descubrir de forma eficiente la existencia de pacientes con un perfil genómico compatible con ensayos clínicos.

El plan de difusión incluye la presentación en congresos y publicación en revistas de ciencia de datos. En paralelo se usará la oficina de prensa de la Fundación progreso y Salud y de otras entidades participantes para su difusión en prensa y en medios digitales.

En principio no hay plan de explotación de resultados. En cualquier caso, si se detectase que los resultados son susceptibles de ser patentados y explotados, la Oficina de Transferencia de tecnología (OTT) de la FPS asesoraría en los pasos que hay que dar.

- **Implicaciones éticas y/o de seguridad.**

No existen implicaciones éticas ya que, aunque todos los desarrollos se hacen para su uso final con datos sensibles, en ningún momento del desarrollo se requiere el uso de datos sensibles y los entregables son los códigos en formato abierto de las aplicaciones desarrolladas.

- **Presupuesto detallado del proyecto y justificación del mismo**

Personal (200.000€): 85% del presupuesto, distorsión producida por la necesidad única de contratar a personal implicado en tareas de programación. Se requiere de la contratación de dos ingenieros informáticos o de otras ramas con experiencia en informática y ciencia de datos para participar en las partes de desarrollo de aplicaciones de todas las tareas durante toda la duración del proyecto (30 meses).

Equipamiento/infraestructura (30.000€): 13% del presupuesto, desviación necesaria dado el coste del personal a contratar. Este equipamiento contempla la adquisición de un equipo informático en el que se instalarán las aplicaciones de uso federado de datos (26.000€) y dos equipos de sobremesa como estaciones de trabajo de los dos ingenieros contratados (4.000€).

Otros: 3.000€ (movilidad, presentación de resultados en publicaciones en abierto, informe del auditor). 1% del presupuesto, no hay desviación.

Costes indirectos (15% costes directos): 34.950€. **PRESUPUESTO TOTAL: 267.950€.**

Proyecto 2: GENERACIÓN Y CONSOLIDACIÓN DE HERRAMIENTAS BIOTECNOLÓGICAS Y DE EDICIÓN GENÓMICA PARA EL DESARROLLO Y USO DE BIOMODELOS

IP: Miguel Angel Moreno; Institución: Universidad Pablo Olavide (CABD-UPO).

- **Resumen:** El empleo de biomodelos en biología y medicina es fundamental para la investigación en enfermedades humanas. Agrupar esfuerzos en torno a este objetivo puede posibilitar la generación y uso de estos modelos de un modo más efectivo y sólido. En este proyecto se propone la formación de una red de trabajo para la creación de un catálogo de aproximaciones biotecnológicas mediante la recopilación, consolidación y obtención de herramientas y protocolos en edición genómica de última generación que sirvan de base para el desarrollo de modelos en biomedicina desde levaduras a mamíferos pasando por modelos *ex vivo* (organoides, cultivos primarios, etc). Además, se propone financiar, como prueba de principio, la generación de un número limitado de biomodelos de enfermedades humanas resultantes de este consorcio de trabajo que ayude a consolidar la creación de esta plataforma de biomodelos en la comunidad andaluza. La plasmación de esta plataforma de biomodelos i) fortalecerá las interacciones entre Institutos de Investigación y Universidades de la comunidad andaluza con intereses comunes y ii) posibilitará la creación de un grupo de trabajo con el fin de satisfacer demandas científicas a cuestiones donde la generación o uso de biomodelos respondan a preguntas fundamentales en biología fundamental y medicina con un impacto tanto a nivel regional, nacional como internacional.

- **Antecedentes del proyecto**

El uso de animales y modelos celulares es una práctica esencial en investigación biológica y médica. Cada sistema presenta una serie de características que deben ser tenidas en cuenta a la hora de la generación de los biomodelos requeridos y que han permitido y siguen permitiendo que la ciencia en biomedicina avance y se desarrolle. En este sentido el uso de sistemas unicelulares como las levaduras ha contribuido a dilucidar la función de genes conservados en eucariotas que ha resultado muy útil para empezar a entender la base biológica de determinadas enfermedades humanas. Por otro lado, el empleo de animales invertebrados (como el nematodo *Caenorhabditis elegans* o la mosca de la fruta *Drosophila melanogaster*) y vertebrados (por ejemplo, el ratón y el pez cebra) así como el uso de modelos celulares de mamífero en cultivo o de sistemas 3D *ex vivo* u organoides han permitido un mayor desarrollo experimental y de conocimiento fundamental para el avance científico en biomedicina. Más allá de estos y otros modelos más usados tradicionalmente en biología y biomedicina como anfibios, cerdos o incluso primates no humanos, los avances en edición génica de los últimos años han permitido el uso y la implementación de nuevos sistemas como salamandras o ajolotes, que son sistemas empleados para entender la biología regenerativa. En este sentido la generación de mutantes para modelar enfermedades humanas en diversos sistemas animales y modelos en experimentación se ha catapultado durante la última década dado el desarrollo de técnicas de edición genómica mediante aproximaciones basadas en las endonucleasas de "dedos de zinc", TALEN o más recientemente los sistemas procarióticos CRISPR-Cas que han extendido y democratizado notablemente la edición génica dado su flexibilidad y coste reducido [PMID: 32572269]. El sistema CRISPR-Cas más utilizado proviene originalmente de *Streptococcus pyogenes* [PMID: 22745249] que consiste en dos componentes en los que un ARN guía (gRNA) dirige la endonucleasa Cas9 (SpCas9) a su sitio de destino. El reconocimiento de la diana se produce primero mediante la detección de una secuencia fija denominada PAM (del inglés *protospacer adjacent motive*) y luego mediante la complementariedad entre el gRNA y el ADN diana en el genoma. SpCas9 genera cortes de ADN de cadena doble (CCD) que pueden acabar induciendo mutaciones (esencialmente inserción o deleción '-*indels*'-) [PMID: 32572269].

Por otro lado, La mayoría de los más de 50.000 cambios genéticos actualmente asociados con enfermedades humanas son mutaciones puntuales (PMID: 29234789). A pesar de que la tecnología CRISPR-Cas se ha convertido en una revolución en biotecnología y biomedicina, existen importantes limitaciones que condicionan su uso de manera eficiente y precisa. De hecho, para recapitular mutaciones concretas y generar modelos genéticos específicos (o avatares) se necesitan sistemas de edición precisos ya que las mutaciones *indels* no suelen ser útiles para generar una modificación específica. En los últimos 5 años se han desarrollado tecnologías basadas en la tecnología CRISPR-Cas que permiten de facto la introducción de mutaciones concretas de una manera controlada y relativamente precisa. De este modo, aproximaciones como la edición de base (*base editing*) o la edición de calidad (*prime editing*) [PMID: 32572269] así como el uso de Sistemas CRISPR-Cas con secuencias PAM variadas y más simples (PMID: 33483498) pueden permitir un mayor acceso al genoma y la modificación precisa del mismo de un modo eficiente.

Aunque en Andalucía existen laboratorios e institutos que han sido capaces de usar e implementar herramientas y aproximaciones de última generación para desarrollar biomodelos en enfermedades humanas, una coordinación de los mismos para establecer y consolidar una plataforma de trabajo eficiente y robusta posibilitará la unificación de esfuerzos entorno a preguntas científicas relacionadas, así como la accesibilidad de estas aproximaciones a la comunidad científica regional, nacional e internacional. Finalmente, desde la financiación incluida en este proyecto se impulsarán un número limitado de biomodelos que serán la prueba de principio del funcionamiento y productividad de la plataforma de trabajo. Esta propuesta será por tanto la base de una red de trabajo que, idealmente, deberá seguir creciendo y consolidándose a través de convocatorias futuras en esta misma línea de trabajo.

- **Objetivos del proyecto**

En este proyecto se plantean **dos objetivos principales**:

- 1) Formación de una red de trabajo en biomodelos a nivel andaluz:
Generación, consolidación y optimización de un catálogo de recursos, herramientas y protocolos que abarquen posibilidades heterogéneas para hacer edición genómica en diferentes biomodelos con distinto nivel de resolución y precisión.

- 2) Pruebas de principio de la obtención y desarrollo de biomodelos de enfermedades humanas generadas a través de la interacción y uso de la plataforma de los grupos de investigación de la propuesta.

- **Contribución del proyecto a la línea de actuación a la que se adscribe el proyecto.**

Esta propuesta está enfocada a la creación de una plataforma de trabajo y colaboración a nivel andaluz que más allá de reforzar la potencialidad en investigación, diseño y desarrollo de biomodelos por parte de los Centros de Investigación y Universidades de la Comunidad Autónoma Andaluza pretende construir una base sólida para dar una continuidad a esta plataforma. De este modo se proyecta una red de trabajo que pueda también proporcionar soluciones y aportaciones a problemas científicos a nivel nacional e internacional dentro de la temática de generación y uso de modelos en biomedicina lo cual encaja en la temática general LA2. De este modo la propuesta de manera directa abarca lo recogido en la línea 2.2 y 2.4 dado que el primer objetivo tiene como fin generar, consolidar y optimizar las herramientas y protocolos necesarios para el diseño, la generación y la catalogación de biomodelos desde levaduras hasta modelos murinos pasando por modelos *ex vivo*, incluyendo la potencial interacción con plataformas de cribado que formarán parte del consorcio y que proporcionarían la posibilidad de llevar a cabo validaciones experimentales de fármacos y tratamientos sobre los biomodelos. Entre las herramientas a consolidar y catalogar se cuentan vectores virales y no virales con un impacto en la línea de actuación 2.5. Esto en su conjunto está enfocado de un modo final a satisfacer lo recogido en la línea de actuación 2.1 que será reforzado con los objetivos principales de esta propuesta con la generación de subgrupos de biomodelos priorizados dentro del consorcio y que contribuirán a la generación de un catálogo de biomodelos con un potencial crecimiento ilimitado dentro de la plataforma.

La línea de actuación 2 se compromete a colaborar con los otros proyectos gestionados por la Unidad de Coordinación además de participar en las actuaciones conjuntas que surjan de otros proyectos en otras CCAA. Esta propuesta fomenta, por otro lado, la participación de forma transversal a nivel de la CA Andaluza gracias a la participación de agentes OPIs/Universidades/SSPA proporcionando recursos, conocimiento y generando biomodelos de distinta índole dentro del propio proyecto. Finalmente, el IP actuará como representante andaluz en el comité técnico nacional correspondiente. En la metodología se detallan posibles interacciones concretas con las otras propuestas.

- **Metodología y plan de trabajo.**

Objetivo 1. Formación de una red de trabajo en biomodelos a nivel andaluz.

Esta parte del proyecto se centrará en la generación y consolidación de un catálogo de recursos, herramientas y protocolos que abarquen posibilidades heterogéneas para realizar edición genómica y optimizar aproximaciones moleculares para la obtención de diferentes biomodelos a distinto nivel de resolución y precisión y que mediante este catálogo sean accesibles al consorcio de esta propuesta y a la comunidad científica en general. Además, se llevará a cabo una reunión presencial de inicio de trabajo de la plataforma invitando a los laboratorios implicados en esta propuesta para que se compartan conocimientos, aproximaciones e ideas y así poder avanzar de un modo coordinado y coherente para cubrir los objetivos del proyecto. Se llevarán a cabo varias sesiones para exponer las aproximaciones y recursos de cada grupo y también para coordinar los esfuerzos en la generación del catálogo de herramientas y posibles colaboraciones para desarrollar proyectos científicos ya sean financiados o no por esta propuesta, así como para establecer los criterios para la convocatoria interna planeada en el objetivo 2.

De este modo primero se propone generar un catálogo y una base de datos accesible de:

a) Sistemas CRISPR-Cas para edición genómica en distinto formato y con diferentes características y flexibilidad para el uso en biomodelos.

Se planea generar una colección de recursos que estará a disposición del consorcio de esta propuesta internamente en un principio y después abierto a la comunidad científica mediante una página web (ver más adelante). Dado que la variedad de PAM en los distintos sistemas CRISPR-Cas resulta crucial a la hora de poder seleccionar la diana que se requiera para llevar a cabo la modificación genómica de interés se plantea tener una colección de recursos y una base de datos de distintas proteínas Cas, SpCas9, Nme2Cas9, SaCas9, LbCas12a, AsCas12a, miniCas (Cas12f/phi) y las versiones con PAM flexibles SpG y SpRY. De estos sistemas se generará una colección de plásmidos (que se mantendrán a -20 °C) y bacterias (*E. coli* mantenidas a -80 °C) para expresar las proteínas endonucleasas Cas y sus guías de RNA correspondientes en formatos que permitan i) la expresión desde promotores de RNA polimerasa II y III, ii) su obtención como mRNA por transcripción *in vitro* y iii) la purificación de la proteína Cas para su uso como complejo ribonucleoproteico (RNP). Además, se consolidará una colección de vectores para generar lentivirus (LV) y virus adenoasociados (AAV) que expresen las proteínas Cas más comúnmente usadas: SpCas9 y Cas12a (Lb o As). Aquí se incluirá al menos una colección para llevar a cabo escrutinios genéticos con CRISPR-SpCas9.

Como se ha apuntado en la introducción, para llevar a cabo edición génica de precisión se necesitan herramientas que de manera concreta modifiquen cambios concretos en el genoma. En este objetivo, además de la colección descrita se generará un catálogo similar de editores de base y de calidad cuyos efectores (proteína Cas y gRNA) sean expresados en los distintos formatos definidos anteriormente para los sistemas CRISPR-Cas. De este modo se recopilarán vectores que lleven a cabo las ediciones de base hasta la fecha descritas (C-T, sistema CBE y A-G, sistema ABE, esencialmente) basados en SpCas9 y en sus versiones de PAM reducidas SpG y SpRY. Se procederá igualmente para la generación de vectores de edición de calidad. Esta colección no estará limitada a estos recursos y se ampliará con el tiempo, la consolidación de la red y las propuestas de la misma, estando abierta a la incorporación de nuevos recursos a sugerencia de los laboratorios participantes en el consorcio y que puedan ser útiles en la generación de biomodelos. Estos recursos vendrán acompañados de protocolos establecidos o en marcha por los distintos laboratorios implicados en la plataforma de trabajo o basados en publicaciones científicas que facilitarán el uso de estas tecnologías por la comunidad del consorcio de esta propuesta además de otros usuarios externos a ella.

Uno de los objetivos de la reunión de trabajo inicial, además de conocer las líneas científicas involucradas en la red, es la puesta en común de aproximaciones e ideas que hagan este catálogo lo más útil posible para el desarrollo y uso de distintas aproximaciones en la obtención de biomodelos y determinar las estrategias para completarlo mediante la generación de nuevas herramientas o la adquisición de las mismas. Así, los recursos se recopilarán a través del consorcio de trabajo desde los distintos investigadores asociados y los que no estén disponibles se

adquirirán través de servicios de obtención de herramientas como *addgene*, entre otros. Aunque no estén directamente relacionados con la generación de biomodelos, se plantea la posibilidad de incluir en el catálogo a vectores de expresión de Cas13 y Cas7-11 cuya diana es el ARN pero que se pueden usar como sistemas de eliminación o edición del ARN y ser también útil en el modelaje de enfermedades en sistemas. Además, en esta colección se recopilarán otro tipo de vectores y herramientas que sean útiles para la generación de biomodelos como plásmidos para la inserción de etiquetas fluorescentes, la generación de moldes para la obtención de *knock-in*, etc.

b) Modelos eucariotas en biomedicina

Para la formación de este grupo de trabajo se ha abierto a la participación a los distintos centros de investigación y Universidades de Andalucía y se ha organizado una red en la que se incluyen laboratorios que poseen una alta experiencia en la generación de biomodelos de distinta índole y/o en el uso de ellos para estudios enfocados en biomedicina. Este grupo de laboratorios además de participar en la reunión inicial que permitirá coordinar esfuerzos en generar los catálogos de recursos servirán también como agente consultor a la hora de necesitar de determinados biomodelos bien en proyectos dentro del consorcio o fuera de él. Esta será una posibilidad futura que se ofertará de manera oficial a través de la creación de la web anteriormente mencionada y en coordinación con la web de Andalucía-Biotec Salud y del Plan Complementario estatal. En este apartado se describen los subgrupos de trabajo en biomodelos que se establecerán, sus participantes y algunas de las posibilidades que se catalogarán en la colección de recursos y que se explotarán durante la ejecución de esta propuesta:

Levaduras: Modelo unicelular eucariota para el que se dispone de una amplia diversidad metodológica en bioquímica, genética y biología molecular y celular para responder cuestiones básicas asociadas a enfermedades. Laboratorios: Rafael R. Daga (CABD-UPO, Sevilla): Optimización de una plataforma robotizada y biosensor de levaduras de fisión que simula envejecimiento celular por agregados proteicos. Jesús de la Cruz (IBIS-US, Sevilla): *Saccharomyces cerevisiae* para entender patologías asociadas a procesos conservados evolutivamente en eucariotas.

Modelos ex vivo (cultivos celulares humanos, cultivos primarios de pacientes, iPSC): Los modelos *ex vivo* posibilitan la generación de escrutinios genéticos de un modo sencillo. Los modelos iPSC y los cultivos primarios de pacientes brindan una serie de ventajas relacionadas con la medicina personalizada como primer paso en terapia génica o farmacológica. Rosario Fernández y Olga Genilloud (Fundación Medina, Granada): biomodelos de enfermedades humanas con líneas celulares o iPSC knock-out y knock-in mediante CRISPR-Cas9. Escrutinios CRISPR en líneas celulares. Elena Reales y Álvaro Juárez (INIBICA, Cádiz): Generación de cultivos de fibroblastos primarios de pacientes y modelos celulares renales que expresan proteínas etiquetadas por knock-in. Elena González (BIONAND-IBIMA, Málaga): Generación de iPSC humanas para modelos. José Antonio Sánchez Alcázar (CABD-UPO, Sevilla): Generación de cultivos de fibroblastos primarios derivados de pacientes con enfermedades genéticas y reprogramación directa a músculo o neurona. Francisco Martín y Pilar Muñoz (Genyo, UGR, Granada) y F. Javier Molina Estévez (FIBAO, Granada): generación de modelos KO-específicos para inmunoterapia, electroporación de RNPs en células primarias y en líneas inmortalizadas tejido específicas para el estudio de enfermedades genéticas. Modelos humanizados mediante edición *ex vivo* de células madre hematopoyéticas humanas y su posterior injerto en ratones inmunodeficientes. Juan Antonio Marchal (UGR-BioFabi3D lab e ibs.GRANADA): cultivos primarios tanto de tumores sólidos (células tumorales, CSCs, CAFs) como de tejidos normales (MSCs, fibroblastos dérmicos, queratinocitos, EPCs y condrocitos).

Modelos 3D y organoides: Estos modelos 3D permiten trabajar sobre el fondo genético de pacientes o de modelos animales que se desee en un contexto fisiológico más adecuado como alternativa o complementación a los modelos animales. Isabel González y Carlos López (IBIMA-Plataforma BIONAND, Málaga): cultivos 3D derivados de pacientes o de modelos pluricelulares y organoides derivados de biopsias de pacientes. Maribel Torres (Univ. Jaén): Optimización de tejidos reconstruidos en 3D mediante usos de biorreactor fluido promotor de microambientes fisiológicos. María Almuedo (CABD-CSIC, Sevilla): Generación de organoides desde embriones de pez cebra.

C. elegans: Modelo conveniente para estudiar la biología del desarrollo, facilidad para la manipulación genética, desarrollo estereotípico y una buena proporción de genes conservados con humanos. Plataforma sencilla para cribados de distinta índole. Laboratorios: Peter Askjaer y Manuel Muñoz (CABD-CSIC, Sevilla): técnicas para el control preciso de la edición del genoma mediada por CRISPR-Cas9 con diversas aplicaciones. Generación de modelos de enfermedades humanas.

D. melanogaster: Sistema manipulable genéticamente, económico, ciclo vital corto con un alto porcentaje de genes ortólogos con humanos que permite el modelaje de enfermedades humanas neurodegenerativas, metabólicas o cáncer, entre otras. Fernando Casares y Sol Sotillos (CABD-CSIC, Sevilla): CRISPR-Cas y recombinación homóloga para avatares y biomodelos en *D. melanogaster*.

D. rerio (pez cebra): Modelo de vertebrado que debido a su facilidad de manejo, recursos genéticos y versatilidad permite mimetizar enfermedades humanas y además lo convierten en un sistema conveniente para la optimización de herramientas de edición genómica *in vivo*. Plataforma de vertebrado económica y simple para cribados de distinta índole. Miguel A. Moreno-Mateos y Gloria Brea (CABD-UPO, Sevilla): Optimización de sistemas CRISPR-Cas y edición genómica de precisión *in vivo* en modelos de teleósteos. Modelos y avatares de enfermedades mitocondriales humanas en el pez cebra. Sara R. Heras y Ana Ariza (GENYO, Granada): Modelos y aproximaciones CRISPR-Cas en el pez cebra para entender el papel de RNAs no codificantes en enfermedades congénitas.

M. musculus: Modelo más usado y establecido en biomedicina con recursos y posibilidades de toda índole. Manuel Tena (IMIBIC-UCO, Córdoba): Experiencia y disponibilidad de cepas murinas SpCas9 (LSL-Cas9^{eGFP}) que permiten la expresión de SpCas9 con control temporal y espacial (vía cruce con líneas adecuadas, disponibles para expresión en hígado, o diversas poblaciones neuronales). Oscar Ocaña (Univ. Jaén): Desarrollo de protocolos de electroporación para la generación de modelos de ratón modificados genéticamente y como plataforma tecnológica común.

Plataformas de Cribado: Rosario Fernández y Olga Genilloud (Fundación Medina, Granada): plataforma de cribado y escrutinios *in vitro* de alto rendimiento: toxicidad, análisis de fármacos etc. Juan de Dios Alché (EEZ-CSIC, Granada): Cribado y caracterización de muestras vegetales y agroalimentarias de diversos tipos para el análisis en biomodelos animales e *in vitro*. Manuel Muñoz (CABD-CSIC, Sevilla): plataforma de escrutinio de fármacos en el modelo *C. elegans*.

En este grupo de trabajo los investigadores e investigadoras responsables representan los distintos niveles académicos y científicos. Así la red la componen personal Ramón y Cajal, Miguel Servet, Nicolás Monardes, Contratados doctores, Profesores Titulares y Catedráticos de Universidad, así como Científicos Titulares, Investigadores Científicos y Profesores de Investigación del CSIC todos ellos provenientes de Institutos de Investigación que representan gran parte (6 de las 8 provincias) de la geografía andaluza.

Además de la colección de recursos, la reunión inicial y la organización para aunar esfuerzos que potencien nuestra capacidad de generar y optimizar biomodelos en Andalucía, se planea poner estos recursos disponibles a la comunidad científica mediante el diseño de una página web donde se pueda y acceder a la información aquí recogida y catalogada, tanto de recursos como de laboratorios, y que sirva de escaparate para la promoción de potenciales colaboraciones directas así como propuestas en las que los distintos grupos de trabajo servirán como una plataforma de consulta para contribuir con el diseño de biomodelo más apropiado dado una problemática concreta. Este recurso se realizará en coordinación con las LA1 y LA4 de esta propuesta.

Objetivo 2. Prueba de principio de la generación y/o uso de modelos experimentales para responder cuestiones en biomedicina

Para plasmar a través de proyectos los resultados de la formación del consorcio de trabajo y de la consolidación y catálogo de recursos y dado el presupuesto del que se dispone, se propone realizar una convocatoria interna dentro de los grupos que participan en la red de trabajo y definidos en el objetivo 1 para la ejecución de tres propuestas. En esta convocatoria interna se financiarán tres subproyectos donde el objetivo será la generación, consolidación y/o caracterización de determinados modelos que respondan cuestiones relacionadas con la biomedicina. Las características de financiación de cada subproyecto serán definidos en el apartado de presupuesto. Para valorar estas propuestas se contará con un comité externo de 4 personas expertos en la materia: dos miembros del comité serán de otras líneas de actuación dentro de esta propuesta otros dos serán asesores o asesoras independientes. Una vez acontecida la reunión inicial y a partir de lo que se decida en la misma se abrirá la convocatoria interna donde para la concesión de los tres proyectos, de entre las propuestas que se reciban, se valorará la interacción entre distintos grupos asociados, que los proyectos tengan un claro impacto en la generación de biomodelos con interés en biomedicina y que se alineen bien con las líneas de actuación de esta convocatoria, que sean propuestas viables en el tiempo de ejecución de este proyecto (dos años en total y los subproyectos 15 meses, ver cronograma), entre otras variables que serán consensuadas en la reunión inicial. De manera notable, la caracterización de los biomodelos podrá beneficiarse de la interacción con la línea de actuación 3 de este proyecto coordinado, lo que completa la conexión de esta red de trabajo en biomodelos con las otras propuestas. Los resultados de este objetivo serán una demostración del posible desarrollo de este consorcio que servirá como prueba de concepto de lo idóneo de esta plataforma y de la potencialidad de la misma.

Cronograma:

Tarea	Año 1		Año 2	
1.a. Reunión inicial del consorcio de biomodelos	■			
1.b. Catálogo de recursos y protocolos		■	■	■
1.c. Página web del consorcio				■
2.a. Convocatoria interna y resolución de proyectos		■		
2.b. Ejecución de los proyectos concedidos			■	■

Hitos año 1: Se habrá llevado a cabo la reunión presencial de inicio de proyecto con la coordinación de los subgrupos de trabajo, así como se ejecutará la convocatoria interna de los proyectos y se iniciarán los mismos. Además, se iniciará el catálogo de herramientas y recursos para la generación de biomodelos.

Hitos año 2: Se culminarán los proyectos donde se espera que al menos alguno de ellos sea parte o incluso el peso principal de alguna publicación científica. Se completará el catálogo de recursos que en conjunción con la creación de la página web consolidará la formación de la plataforma y la disponibilidad de los recursos recopilados en el primer año y medio de desarrollo.

- **Grado de innovación de la propuesta y justificación del potencial impacto científico y/o tecnológico del proyecto.**

La catalogación y consolidación de recursos y métodos para la generación de biomodelos y la creación de una red coordinada de trabajo tendrá un impacto notable a medio y largo plazo en la capacidad científico-técnica de la Comunidad andaluza. Aunque con una financiación considerablemente mayor este tipo de consorcios se han convertido recientemente en una tendencia a nivel internacional o nacional. Así redes como la *Somatic Cell Genome Editing* del NIH en EEUU o consorcios a nivel europeo como *COST Action (GenE-Humdl)*, a la que el IP de esta propuesta pertenece, van en la línea de lo aquí proyectado. Esta propuesta planea iniciar una innovadora red de trabajo centrada en la generación y uso de biomodelos a nivel regional, nacional e internacional. Aunque el tiempo de ejecución es corto, de los proyectos de la propuesta interna puede haber resultados que sean susceptibles de protección intelectual. Esto será también valorado a la hora de la concesión de las ayudas durante la ejecución de este proyecto.

- **Resultados esperados a los que contribuye dentro de la línea de actuación. Plan de difusión y explotación.**

Aunque no se dispone de mucho tiempo para tener un impacto a corto plazo, la creación de un sitio web donde esté recopilada toda la información y sirva de consulta y enlace con la comunidad científica interesada servirá para establecer colaboraciones y para generar proyectos que involucren la generación de biomodelos de un modo más sólido con la ayuda de los grupos de trabajo de la propuesta. Esto fomentará de manera notable la difusión de los resultados del objetivo 1 que además serán publicitados por canales diversos como notas de prensa desde las instituciones involucradas en el grupo de trabajo o redes sociales como Twitter. Más allá de esto, los resultados de los

proyectos financiados en la convocatoria interna serán susceptibles de ser publicados como artículos o comunicaciones a congresos donde se reflejará la contribución en estos trabajos del proyecto que aquí se presenta.

- **Implicaciones éticas y/o de seguridad.**

Cada subgrupo de trabajo mencionado en el objetivo 1 dispone de los permisos de experimentación animal necesarios para la generación y manipulación de biomodelos (en el caso de los trabajos de vertebrados) así como las autorizaciones para el acceso a muestras de pacientes.

- **Presupuesto detallado del proyecto y justificación del mismo.**

Personal (100.000€): Se prevé la contratación de una persona a cargo de la coordinación del objetivo 1 (personal de apoyo técnico y auxiliar de gestión, 18 meses, 44.000€), encargada de la organización de la reunión inicial, la ejecución del catálogo de recursos y protocolos y el desarrollo del sitio web en colaboración con las LA1 y LA4. Además, dos de las propuestas que se financien dentro del objetivo 2 dispondrán de financiación para la contratación de personal de apoyo técnico (12 meses, 28.000€/contrato). Este personal llevará a cabo tareas específicas dentro de las propuestas concretas que sean financiadas en la convocatoria interna.

Equipamiento y fungible (76.000€): Se distribuirá en los distintos proyectos de la convocatoria interna de la siguiente manera para los gastos de ejecución/fungible: 36.000€ para uno de los proyectos y 20.000€/proyecto para los otros dos proyectos dotados con contratos de apoyo.

Otros gastos: 14.130€. Gastos para organizar la reunión inicial que incluirá una noche de alojamiento (posiblemente en un hotel en Antequera al ser un lugar equidistante en la geografía andaluza) y pensión completa por un día para aproximadamente 30 personas, más el alquiler de la sala de reuniones y los medios audiovisuales (aproximadamente 6.000€). El resto irá destinado a la consolidación de recursos que sean necesarios para el catálogo de herramientas y protocolos, así como a la creación, mantenimiento de la página web y auditoría del proyecto.

Costes indirectos (15% costes directos): 28.520€. PRESUPUESTO TOTAL: 218.649€.

Proyecto 3: PLATAFORMA DE TRANSCRIPTOMICA ESPACIAL PARA APLICACIONES BIOMÉDICAS

IP1: Fuencisla Matesanz del Barrio; IP2: Alicia Barroso del Jesus. Instituciones: IPBLN-CSIC/Universidad de Granada

- **Resumen de la propuesta:** La propuesta contempla la implementación y puesta en marcha de una plataforma de transcriptómica espacial (*spatial transcriptomics*, ST) basada en la tecnología GeoMx®/nCounter®, para su utilización en investigación biomédica y clínica. Se trata de una solución integral que abarca el tratamiento experimental de las muestras clínicas y el procesamiento bioinformático de los datos e interpretación de resultados. La plataforma estará ubicada en el PTS de Granada como resultado del esfuerzo conjunto del Consejo Superior de Investigaciones Científicas (CSIC) y la Universidad de Granada (UGR) a través del Instituto de Parasitología y Biomedicina "López-Neyra" (IPBLN) y el Centro de Investigaciones Biomédicas (IIBM), respectivamente. No obstante, se concibe como una iniciativa abierta que pretende la integración de diferentes agentes del conocimiento en el ámbito científico y clínico dentro y fuera de la Comunidad Autónoma de Andalucía. Esta plataforma permitirá obtener información de relevancia biomédica a partir de muestras clínicas tanto frescas como de archivo y pretende servir como una potente e innovadora herramienta para el desarrollo de procedimientos diagnósticos y terapias más personalizadas. El proyecto constará de una primera fase de implementación y puesta a punto de la tecnología en un modelo experimental murino, una segunda fase de validación de la tecnología en muestras clínicas humanas de glioblastoma, y una última fase de prueba de concepto en el que la plataforma se abrirá a la comunidad científica desarrollando proyectos externos de interés.
- **Antecedentes del proyecto**

Los procesos patológicos ocurren en contextos tisulares complejos en los que las interacciones celulares, la influencia del microambiente que rodea las células, y la infiltración y distribución de células inmunes, entre otros factores, juegan papeles protagonistas en la aparición de las alteraciones, la progresión de la enfermedad y la respuesta a los tratamientos. Esto es especialmente trascendente en patologías como el cáncer, enfermedades neurodegenerativas o enfermedades complejas autoinmunes.

La ST proporciona información del perfil transcripcional de las células asociada a su localización bidimensional en el tejido. Es por ello que la ST aporta un nuevo estrato de información molecular fundamental en la investigación biomédica y en su traslación a la clínica.

Actualmente, existen tres tecnologías principales que permiten abordar este tipo de estudios: GeoMx®/nCounter® (Nanostring), Visium® (10x Genomics) y Stereo-seq® (MGI) cada una con sus fortalezas y debilidades. De entre ellas, GeoMx®/nCounter® presenta tres características exclusivas que la convierten en la tecnología de elección para este proyecto dado su marcado carácter clínico y traslacional:

1. La tecnología GeoMx®, a diferencia de Visium® y Stereo-seq®, es independiente del nivel de integridad del RNA de la muestra estudiada, lo que le permite trabajar sobre cualquier muestra parafinada. Esta característica tiene una repercusión enorme en la aplicabilidad de la tecnología, puesto que los cortes de tejidos parafinados son el objeto de estudio habitual en la patología clínica. Esto permite asociar la nueva información molecular a la información histológica convencional en los procesos diagnósticos y, además, pone a disposición de la ST las muestras de archivo acumuladas durante años, lo que potencia enormemente la capacidad de descubrimiento.
2. La segunda característica diferenciadora que justifica la utilización de esta aproximación metodológica es que, en contraposición con las otras tecnologías que, inexcusablemente obtienen información de la totalidad del tejido procesado, esta permite al investigador o profesional de la salud seleccionar zonas específicas del tejido (ROIs, *regions of interest*), basándose en las tinciones o marcadores de enfermedad, lo que descarta un análisis indiscriminado y permite reducir de manera drástica los costes de secuenciación asociados, un parámetro clave para la traslación efectiva de la tecnología al ámbito de la salud.
3. Por último, esta tecnología posibilita el análisis de un grupo reducido de genes de las ROIs (frente al transcriptoma completo), práctica habitual en el ambiente de diagnóstico clínico, más focalizado en genes relacionados con la patología analizada. Esta aproximación permite una lectura directa de resultados mediante análisis de fluorescencia en el scanner nCounter®, prescindiendo de la utilización de tecnología de NGS para generar resultados. Esto reduce el coste del ensayo y el tiempo de obtención de resultados.

En definitiva, actualmente, la combinación de la SCS y la ST constituye una poderosa herramienta para relacionar los fenómenos patológicos de los tejidos humanos y las alteraciones moleculares constituyendo una de las bases presentes y futuras en las que ha de sustentarse la medicina de precisión.

- **Objetivos del proyecto**

1. Puesta en marcha de la plataforma de ST GeoMx®/nCounter® (Nanostring) sobre un modelo experimental animal con organización tisular compleja (cerebro murino). Establecimiento de protocolos normalizados de trabajo y validación de la reproducibilidad de los procedimientos experimentales.
2. Validación de la plataforma ST y de los procedimientos establecidos con un modelo de enfermedad humana del SNC: glioblastoma
3. Análisis paralelo mediante SCS (plataforma Chromium®, 10x Genomics). Comparativa directa entre ambas aproximaciones tecnológicas (SCS vs ST)
4. Definición e implementación de los procedimientos bioinformáticos de análisis para cada tecnología y estudio comparado de resultados.

5. Establecimiento de un decálogo para la selección de la tecnología de elección dependiendo de la muestra y del objetivo del estudio. Formación de un comité asesor para dar soporte a los grupos interesados.
6. Apertura de la plataforma y fase de prueba de concepto mediante la realización de proyectos externos

- **Contribución del proyecto a la línea de actuación.**

La línea de actuación 3.3 de la presente convocatoria, contempla la caracterización ómica de muestras y biomodelos a escala de célula individual, incluyendo expresamente técnicas de SCS y ST, por lo que la propuesta se encuadra perfectamente en los objetivos propuestos. El equipo de investigación tiene una naturaleza altamente multidisciplinar y aporta los modelos biológicos, la infraestructura, el conocimiento y la experiencia para el desarrollo del proyecto gracias a un esfuerzo sinérgico entre diferentes instituciones andaluzas. Los grupos del Dr. Pablo Galindo y del Dr. Francisco O'Valle, pertenecientes a la Universidad de Granada contribuyen con el equipamiento de ST (GeoMx®/nCounter®) valorado en más de 1 millón de euros, personal con formación para su manejo y los conocimientos de anatomía patológica para la preparación de tejidos. El grupo de OncoObeidad y Metabolismo del Dr. Raúl Luque del Instituto Maimónides de Investigación Biomédica de Córdoba (IBIMIC; Universidad de Córdoba) aporta las muestras y datos clínicos para la validación de la plataforma y el conocimiento del modelo biológico de glioblastoma. Por su parte el grupo de la Dra. F Matesanz (IPBLN-CSIC), que actuará como IP1 de este proyecto, posee una amplia experiencia en estudios ómicos en enfermedades neurodegenerativas y las Unidades Científico-Técnicas de Genómica y de Bioinformática del IPBLN disponen de las plataformas de NGS y SCS (valoradas en más de medio millón de euros) junto con el conocimiento tecnológico y la experiencia necesaria para el desarrollo experimental del proyecto (Unidad de genómica dirigida por la Dra. Alicia Barroso-del Jesús con más de 200 proyectos realizados con éxito desde 2018, y que actuará como IP2 de este proyecto) y un equipo de profesionales en el ámbito de las ciencias de la computación, altamente cualificados y con experiencia en el análisis de datos de sc-omics, además de los recursos de almacenamiento de información y computación necesarios (Dra. Laura Terrón perteneciente a la Unidad de Bioinformática).

La posibilidad de ofertar un flujo de trabajo completo, incluyendo parte experimental y análisis de datos, es una característica singular de enorme valor y, a nuestro juicio, una de las principales aportaciones de esta propuesta a la línea de actuación en la que se encuadra y por extensión a la iniciativa global.

La última fase del proyecto implica la apertura de la tecnología a otros modelos de interés sanitario, para lo cual ya hemos recabado expresiones de interés y contamos con los siguientes candidatos para su participación:

- Grupo del Dr. Diego Franco Universidad de Jaén, estudio de daño cardiaco.
- Grupo de la Dra. Patricia Páez, Universidad de Málaga/IBIMA Interés en hidrocefalia, caracterización de biomodelos y desarrollo terapias celulares.
- Grupo del Dr. Jesús Gonzalo Ponce. Universidad de Cádiz con interés en el empleo de tecnologías ómicas para el estudio de correlación entre ejercicio físico y salud.

La presentación de esta propuesta supone la aceptación de un compromiso expreso de colaboración en el marco de las actuaciones conjuntas que se propongan por parte del Proyecto de la Dirección Científica tanto en Andalucía como con otras Comunidades Autónomas.

- **Metodología y plan de trabajo**

El plan de desarrollo se organiza en hitos consecutivos e interdependientes que permitan un avance sólido y progresivo hacia la consecución de los objetivos:

1. **Puesta en marcha de la plataforma utilizando el modelo de cerebro murino:**
 - a) Optimización de las técnicas de obtención y preparación de muestras: Inclusión en bloque de parafina y criopreservación de tejido. Las muestras de tejido serán proporcionadas por la Unidad de Experimentación Animal del IPBLN.
 - b) Selección de marcadores para los diferentes linajes celulares del SNC y desarrollo de técnicas de marcaje óptimas.
 - c) Utilización de las plataformas GeoMx®/nCounter® y GeoMx®/Illumina para la generación de datos de TS a partir de las muestras experimentales. Análisis de paneles génicos (GeoMx® Custom Panels) y detección directa por fluorescencia en escáner nCounter® frente a análisis de transcriptoma completo (GeoMx® Mouse Whole transcriptome Atlas) y obtención de datos mediante técnicas de NGS y secuenciación Illumina.
 - d) Desarrollo de las herramientas bioinformáticas necesarias para el análisis de datos.
 - e) Análisis comparativo de datos, interpretación y obtención de información biológica relevante.
2. **Validación de la plataforma TS con un modelo de enfermedad humana del SNC (glioblastoma):**
 - a) El material será recogido tras la cirugía e inmediatamente trasladado a Granada por personal de la Unidad de genómica IPBLN (UGIPBLN). La muestra fresca se destinará al análisis de SCS y el resto del tejido se preservará, bien por inclusión en parafina o bien mediante congelación.
 - b) Para los estudios de ST se aplicarán los marcadores y protocolos normalizados establecidos en la fase previa del proyecto con las adaptaciones necesarias por la migración del modelo murino al humano y del modelo fisiológico al modelo tumoral.
 - c) Los estudios de SCS se realizarán siguiendo los protocolos establecidos previamente en la UGIPBLN
 - d) Se realizarán análisis comparativos SCS vs ST y ST GeoMx®
3. **Prueba de concepto de la plataforma TS:**
 - a) Se realizará un llamamiento a través de las herramientas de difusión establecidas para la participación de agentes del conocimiento externos al proyecto.
 - b) Se creará un comité de expertos para evaluar la idoneidad y relevancia de las propuestas externas recibidas y su selección.

- c) Se aplicarán las tecnologías de SCS, ST o SCS+ST, en el contexto de los proyectos incluidos en la iniciativa Andalucía-Biotec Salud, atendiendo a los diseños experimentales particulares, enfermedades estudiadas y objetivos planteados en cada proyecto.
- d) Se abordará el análisis personalizado para cada proyecto, obtención de datos biológica o clínicamente relevantes y presentación de los datos de manera fácilmente interpretable por los investigadores (*ready for publish results*)

- **Cronograma**

Objetivo por trimestre:	Trimestres										
	1	2	3	4	5	6	7	8	9	10	
Formación especializada en tecnología GeoMx®/nCounter®, adquisición de material y reactivos necesarios	*										
Incorporación del personal de refuerzo a la UGIPBLN y formación continua		*	*	*							
Puesta a punto de preparación de muestras para ST (modelo murino)	*	*									
Realización de técnicas de ST (modelo murino) paneles génicos/nCounter		*	*								
Realización de técnicas de ST (modelo murino) transcriptómica total/NGS		*	*								
Desarrollo de pipelines de análisis de datos de ST y estudios comparativos			*	*							
Coordinación con grupos clínicos y establecimiento de protocolos para la toma, recogida y transporte de muestras clínicas). Tramitación de CI			*								
Creación y mantenimiento una de base de datos del proyecto con información clínica de pacientes			*	*	*	*	*				
Recepción, registro y preparación de muestras clínicas humanas				*	*	*	*	*	*	*	*
Realización de técnicas de ST (modelo glioblastoma humano) paneles/génicos /nCounter				*	*	*					
Realización de técnicas de ST (modelo glioblastoma humano) transcriptómica total/NGS				*	*	*					
Realización de técnicas de SCS (modelo glioblastoma humano)				*	*	*					
Análisis comparativo de datos, interpretación y obtención de información biológica relevante					*	*	*	*	*	*	*
Creación de un comité de valoración y selección de propuestas externas						*					
Realización de proyectos externos de TS							*	*	*	*	*
Análisis personalizado de proyectos externos y presentación de resultados							*	*	*	*	*
Difusión y visualización de progresos y resultados			*	*	*	*	*	*	*	*	*

- **Grado de innovación de la propuesta y justificación del potencial impacto científico y/o tecnológico del proyecto.**

La técnica de transcriptómica espacial es altamente innovadora en biomedicina ya que permite analizar la expresión génica asociada a la localización celular. Esto ofrece una mayor comprensión de la organización y función de los tejidos y células en su entorno tridimensional, así como de las intrincadas relaciones intercelulares y su impacto en los procesos fisiológicos y patológicos, conocimiento que era inalcanzable con las técnicas de transcriptómica convencionales (*bulk o single cell*). La plataforma será la única existente en Andalucía y una de las dos disponibles a nivel nacional. Además, el modelo de validación seleccionado tiene a su vez una enorme relevancia clínica ya que los glioblastomas son tumores relativamente frecuentes, sin tratamiento eficaz y con un pronóstico devastador. Se trata de un claro ejemplo en el que la ST puede suponer una herramienta de valor incalculable para la identificación/caracterización/localización de subpoblaciones celulares relevantes que conduzcan al hallazgo de nuevas dianas terapéuticas. Por tanto, se trata de una propuesta innovadora en el sentido de que promueve el acceso a una tecnología de vanguardia inexistente hasta el momento en nuestro entorno autonómico y además presenta un enorme impacto con una doble vertiente, tecnológica por un lado con la implementación de la plataforma abierta y científica por otro con los avances esperables en el conocimiento del glioblastoma.

- **Resultados esperados a los que contribuye dentro de la línea de actuación. Plan de difusión y explotación.**

Uno de los hitos más importantes que se espera alcanzar es la capacidad para ofertar flujos de trabajo completos, en ómicas de célula única y TS, con aplicabilidad a un amplio abanico de problemas biomédicos, desde el área de la investigación básica al área de la clínica. Lo que redundará en una mayor accesibilidad de todos los centros biomédicos/hospitalarios a esta tecnología promoviendo el carácter traslacional y su implementación con fines de medicina personalizada. La oferta de una respuesta integral, aunando la experimentación con el análisis e interpretación de datos, resultará en una herramienta única e inestimable para la comunidad biomédica.

El resultado esperado más relevante, por su carácter innovador y por su potencial impacto en la investigación clínica, es, sin duda, la posibilidad de implementar la tecnología de transcriptómica espacial y de los sistemas de análisis correspondientes y adquirir la capacidad de, en base a los protocolos establecidos en la prueba de concepto, adecuar dichos procedimientos a las particularidades de la demanda de los centros y grupos interesados. Se trata de la herramienta más vanguardista en la actualidad en el campo de las ómicas y su disponibilidad en el marco de la iniciativa Andalucía-Biotec Salud/BAS debe promover un avance cualitativo en la calidad de la investigación biomédica.

Los resultados generados a lo largo del desarrollo del proyecto se difundirán a través de la página web oficial de la iniciativa. Todos los datos científicamente relevantes serán recogidos y enviados para su publicación en revistas científicas revisadas por pares. También se planea la difusión en sociedad de los hitos más importantes a través de las cuentas oficiales en redes de los centros y las unidades participantes.

- **Implicaciones éticas**

La utilización del modelo experimental murino garantizará un acceso menos restrictivo a muestras biológicas durante la primera fase de puesta a punto tecnológica, limitando el uso de muestras humanas a la segunda fase de validación. Para ello se utilizarán tejidos primarios de cerebro

obtenidos tras el sacrificio de ratones adultos, siguiendo las directrices del Real Decreto 53/2013, de 1 de febrero e informando de ello al Comité de Ética de Experimentación Animal del IPBLN para que lleve a cabo el seguimiento de estos animales. En este proyecto se utilizarán también muestras biológicas de pacientes de glioblastoma y datos clínicos. La recogida de muestra se canalizará a través del Biobanco de Andalucía, que a su vez se nutrirá de un circuito ya establecido y puesto a punto de recogida y procesamiento de muestras en fresco en la gran mayoría de los hospitales de Andalucía [ej. Hospital Universitario Reina Sofía (Córdoba), Hospital Universitario Virgen de las Nieves (Granada), Hospital Universitario Virgen del Rocío (Sevilla), Hospital Regional Universitario (Málaga) y Hospital Neuro-Traumatológico (Jaén)]. El facultativo informará a los pacientes sobre el proyecto y se encargará de recabar los consentimientos informados (CI) (Ley 41/2002, de 14 de noviembre). El Biobanco garantizará que los donantes tendrán conocimiento en todo momento de los proyectos en los que sus muestras están siendo utilizadas, así como de la publicación de los resultados de la investigación, en el contexto global del estudio. El tratamiento de los datos y su libre circulación, en el contexto del Proyecto, seguirá el marco legal vigente en materia de protección de datos (Ley Orgánica 7/2021, de 26 de mayo, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales). El proyecto se someterá al comité de ética del hospital donde se obtendrán las muestras, así como al comité de ética del CSIC.

- **Presupuesto detallado**

Personal (61.250€): 37% del presupuesto. A pesar de la participación de un excelente plantel de investigadores, la propuesta supone un reto debido i) a la complejidad y novedad de las técnicas y análisis planteados y ii) a la dificultad inherente a las muestras clínicas implicadas en las etapas de validación, prueba de concepto y desarrollo de proyectos externos. Además, los participantes implicados desarrollan una intensa actividad científico-técnica e investigadora y poseen ya unas obligaciones adquiridas ineludibles. Por tanto, aunque el desarrollo del proyecto se sustente en la cualificación y experiencia de los integrantes, la complejidad y magnitud de las tareas a realizar requiere ineludiblemente la incorporación de personal con dedicación exclusiva al proyecto. Será imprescindible la contratación durante 21 meses de un titulado superior adscrito a la UGIPBLN (en la que recaerá la mayor parte de la labor experimental) para reforzar el equipo de trabajo. El candidato/a debe tener un nivel formativo de grado en ciencias de la vida, preferentemente en el ámbito de la genética, y debe tener experiencia profesional en las técnicas de NGS y SCS, además de disponibilidad para desplazarse dentro de la CCAA andaluza para la recogida y transporte de las muestras clínicas.

Equipamiento/infraestructura (100.000€): 62% del presupuesto. Este contempla la adquisición de los fungibles necesarios para su puesta a punto de una tecnología singular recientemente adquirida por el PTS (50.000€/año): preparación y tinción de cortes de tejido (5.000€), reactivos de ST (40.000€), reactivos de SCS (14.000€), reactivos de secuenciación (35.000€), fungible general de laboratorio (6.000€). Este alto porcentaje se debe principalmente a la incorporación de una nueva tecnología que es crucial para el desarrollo de la línea de actuación y para el avance futuro de la medicina de precisión. Además, el coste de los reactivos de una técnica recién lanzada al mercado es siempre muy elevado. Este equipamiento es crítico y necesario para la consecución de los objetivos marcados en la convocatoria.

Otros: 1.000€ (gastos de movilidad, envío/recogida de muestras, informe del auditor, otros costes, etc.). 1% del presupuesto.

Costes indirectos (15% costes directos): 24.337€. **PRESUPUESTO TOTAL: 186.587€.**

Andalucía-Biotec Salud LA3 (Nodo Granada)

Proyecto 4: IMPLEMENTACIÓN DE TECNOLOGÍAS DE TRANSCRIPTÓMICA ESPACIAL DE RESOLUCIÓN CELULAR CON APLICACIÓN BIOSANITARIA Y DESARROLLO DE MÉTODOS COORDINADOS DE ANÁLISIS BIOINFORMÁTICOS ESTANDARIZADOS

IP1: Javier Martín Ibáñez; IP2: Eduardo Andrés León. Institución: IPBLN-CSIC

- **Resumen de la propuesta:** Esta propuesta se enmarca dentro del ámbito del desarrollo de tecnologías innovadoras para el diagnóstico, con enfoque en la medicina personalizada. El objetivo es la implementación de una plataforma de transcriptómica espacial (ST) de célula única de alta resolución. Este equipamiento proporcionará un mejor y más preciso diagnóstico clínico, así como el avance en la comprensión de una amplia gama de enfermedades. Nuestra propuesta aborda un enfoque integral que va desde el procesamiento experimental de muestras, al desarrollo de protocolos de análisis optimizados, posibilitando una rápida interpretación de los resultados obtenidos. Con el objeto de alcanzar la adecuada potencia de análisis para la implementación de esta plataforma, proponemos la adquisición de un sistema de secuenciación “ómica” de resolución espacial mejorada de MGI® que estará ubicada en el Instituto de Parasitología y Biomedicina “López-Neyra” (IPBLN) y, la contratación de personal que proporcionará el apoyo necesario para la creación de protocolos “sc-ómicos” de análisis bioinformáticos enfocados a dar soluciones a necesidades clínicas. Este proyecto tiene un diseño transversal que integra líneas de actuación de diferentes agentes del conocimiento en el ámbito científico y clínico de la Comunidad Autónoma de Andalucía, y se alinea con el resto de propuestas de medicina personalizada de otras comunidades autónomas.

- **Antecedentes del proyecto**

En el campo biomédico, la transcriptómica espacial está proporcionando importantes avances dada su capacidad de obtener perfiles transcriptómicos de miles de dianas con una resolución espacial. Esto permite determinar los procesos biológicos que hay detrás de una

amplia gama de enfermedades y trastornos y por ello va a revolucionar la medicina personalizada en los próximos años. En la actualidad, la tecnología más novedosa y con mayor resolución es la secuenciación “ómica” de resolución espacial mejorada, Stereo-seq® de MGI (del inglés: *spatial enhanced resolution omics-sequencing*). Esta se basa en la identificación de todos los ARNm expresados en cada una de las células del tejido de estudio, sin selección previa. De esta manera, no sólo permite una identificación de célula única “sc-omic” más precisa, sino que también facilita la identificación de procesos moleculares alterados en los que interviene un número muy reducido de genes. Es importante destacar que Stereo-seq® no sólo es una técnica de menor coste, sino que también tiene la mayor definición del mercado y, por tanto, proporciona una resolución a nano-escala, que posibilita generar nuevos conocimientos sobre los complejos procesos moleculares que subyacen a la función de los tejidos y las enfermedades.

El análisis computacional de los datos de ST de alta resolución, propicia la transición de un campo exclusivamente experimental a otro computacional, con el desarrollo paralelo de la bioinformática y de nuevos métodos de análisis. El estudio integrado de muestras del transcriptoma de miles de ARNm a nivel celular (sc-RNASeq) junto con su localización bidimensional en un tejido, permite añadir una nueva capa de información molecular fundamental en la investigación biomédica y su traslación a la clínica. La aparición de estas tecnologías acompaña el desarrollo de nuevos algoritmos específicos que son necesarios para el análisis de los datos producidos.

La creación, manejo y modificación de estos algoritmos en la investigación biomédica para adecuarla a la naturaleza de los proyectos, requiere de la creación de protocolos de análisis bioinformáticos. Estos protocolos se caracterizan por ser:

- Protocolos dinámicos y reproducibles. Dado el carácter transversal de este proyecto con las otras líneas de actuación y otros proyectos de otras CCAA, estos protocolos deben dar respuesta a distintos tipos de muestras de diferentes patologías.
- Capaces de armonizar y anonimizar los datos con el objetivo de poder analizar cualquier tipo de datos “sc-ómico” generado en estos proyectos y realizar informes de resultados integrables en bases de datos de pacientes, de forma segura.
- Idóneos para generar resultados de forma rápida y así, facilitar la toma de decisiones clínicas en tiempo real. Por tanto, es importante que estos protocolos estén optimizados computacionalmente para un uso grandes clústeres de computación científica.

• **Objetivos del proyecto**

El objetivo principal es el desarrollo de protocolos “sc-ómicos”, tanto experimentales como de análisis bioinformáticos, y su aplicación a problemas de salud priorizados por el consorcio. Para ello los objetivos específicos son:

1. Puesta en marcha de una plataforma de transcriptómica espacial de célula única en base a la tecnología Stereo-seq® (MGI), usando en su inicio el modelo experimental de piel en pacientes de enfermedades autoinmunes como la esclerodermia.
2. Desarrollo de protocolos experimentales que aseguren la reproducibilidad de los resultados y minimizar la variabilidad técnica.
3. Creación de una unidad de bioinformática centralizada de referencia, para favorecer la comunicación y organización con los grupos implicados en los análisis “sc-ómicos” en el proyecto global. Entre los objetivos de esta unidad, podemos destacar:
 - a. Desarrollo de protocolos de análisis optimizados de ST de la LA3 que permitan el análisis e integración rápido de resultados.
 - b. Elaboración de métodos de análisis bioinformático “sc-ómicos” estandarizados sobre modelos descritos en la LA2, que permitan el almacenamiento de resultados en las bases de datos objeto de la LA1.
 - c. Producción de protocolos que contemplen las mejoras prácticas para el almacenamiento e intercambio seguro de los datos experimentales producidos por las distintas tecnologías ómicas de las diferentes líneas de actuación.
 - d. Generación de recursos formativos de aplicación autonómica y/o nacional, en el uso de los flujos de análisis desarrollados que faciliten su aplicación en los diferentes entes del ámbito biomédico y clínico.

• **Contribución del proyecto a la línea de actuación.**

La línea de actuación 3 se centra en la caracterización de células individuales y biomodelos, mediante técnicas de secuenciación de célula única y transcriptómica espacial, así como su análisis bioinformático. Este proyecto se alinea con los objetivos propuestos y es el resultado de un esfuerzo de colaboración entre varias instituciones de Andalucía, como el Hospital Universitario Clínico San Cecilio (HUSC), la Universidad de Granada (UGR), el Instituto de investigación Biosanitaria de Granada (IBS) y el Instituto de Parasitología y Biomedicina “López-Neyra” (IPBLN). Estas instituciones generarán un circuito para dar acceso a muestras biológicas de piel que permitirá la puesta a punto de las técnicas de transcriptómica espacial y el desarrollo de los protocolos de análisis. El IPBLN, además, aportará el espacio físico para la instalación de la plataforma de Stereo-seq® así como la experiencia y conocimiento del personal altamente cualificado de la Unidad de Genómica. A su vez, la Unidad de Bioinformática ofrece un equipo de trabajo formado por profesionales con enorme experiencia en el desarrollo de protocolos de análisis de datos “sc-ómicos”. Asimismo, proporcionará los recursos computacionales necesarios para el desarrollo de los pipelines de análisis de datos así, como el almacenamiento seguro de datos.

Desde el inicio de la elaboración de este proyecto, se marcó como requisito indispensable el compromiso expreso de colaboración con el resto de líneas de actuación de esta convocatoria, así como el resto de actuaciones nacionales. Por ello es importante destacar que la plataforma de ST de célula única será de aplicación a aquellos biomodelos desarrollados por la LA2 priorizados por la LA4 de coordinación, de acuerdo con las entidades de las CCAA participantes.

Por otro lado, los protocolos de análisis creados por la unidad de bioinformática, permitirán el estudio de todos los datos “sc-ómicos” generados. En primer lugar, colaborando en la puesta a punto de la técnica de ST de los proyectos de la LA3, tanto el basado en la tecnología de Nanostring como el de Stereo-seq. En segundo lugar, ayudando a la correlación de resultados de los biomodelos de la LA2 y, en tercer lugar, en la creación de resultados en formatos compatibles con su almacenamiento en las bases de datos de uso de la LA1.

• **Metodología y plan de trabajo. Cronograma.**

La metodología empleada en la puesta a punto de los protocolos experimentales de ST de célula única, utilizará muestras de piel de individuos sanos proporcionadas por el HUSC y el IBS, mientras que el IPBLN, las necesarias para el estudio de la esclerodermia. Posteriormente, para su continuo desarrollo, se aplicará a aquellos biomodelos desarrollados en la LA2, así como a proyectos de investigadores que ya han mostrado su interés. El plan de trabajo detallado es:

- Optimización de las técnicas de obtención, manipulación y preparación de muestras: Protocolos específicos para muestras de tejido fresco de piel.
- Utilización de la tecnología de ST de célula única. Puesta a punto de los protocolos de tinción de ácidos nucleicos específicos para la localización tisular de las células a través de análisis de imagen.
- Obtención de los datos de transcriptómica celular mediante técnicas de secuenciación masiva con MGI.
- Aplicación de las tecnologías de ST o scST en el contexto de los proyectos incluidos en la convocatoria atendiendo a los diseños experimentales particulares, enfermedades estudiadas y objetivos planteados en cada proyecto.

En el caso de los protocolos de análisis de datos, la metodología se basará en *Nextflow* como marco de trabajo de computación, dado que está diseñado específicamente para la automatización de pipelines de análisis de grandes cantidades de datos biomédicos. Para favorecer la reproducibilidad y replicación de los resultados, así como minimizar las dependencias de versiones entre programas, estos flujos de trabajo operarán sobre *Singularity*. Esta es una herramienta de contenedores de programas, diseñada específicamente para el uso en sistemas de computación científica de alto rendimiento. El plan de trabajo diseñado es:

- Desarrollo de los primeros flujos de trabajo de análisis haciendo énfasis en los datos de ST de la tecnología de Nanosting, así como de Stereo-seq. Validación de resultados con datos públicos previamente publicados.
- Contacto con los responsables de cada proyecto para el establecimiento de una unidad de bioinformática de referencia que englobe al personal involucrado en los análisis de datos de las líneas de actuación. Desarrollo de otros pipelines de análisis sc-ómicos priorizados por la unidad de coordinación según las necesidades del proyecto global y del resto de entidades nacionales.
- Optimización de los flujos de trabajo para el análisis de datos de enfermedades de interés en el marco de la convocatoria, para facilitar la toma de decisiones clínicas en tiempo real. Presentación de los resultados de manera clara y comprensible para los investigadores y médicos, así como su posible almacenamiento en bases de datos clínicas.
- Creación de un gabinete asesor y de formación a nivel autonómico principalmente, pero también nacional, en el uso de los pipelines de análisis creados para facilitar su uso en entidades del ámbito biomédico y clínico.

Cronograma	Trimestres									
Objetivo por trimestre:	1	2	3	4	5	6	7	8	9	10
Adquisición del secuenciador MGI con capacidad para ST Stereo-Seq e instalación.	*	*								
Puesta a punto del secuenciador y los protocolos necesarios para su funcionamiento.		*	*	*						
Desarrollo del protocolo de recogida y tratamiento de muestras de piel para ST	*	*								
Realización de técnicas de ST de célula única en el modelo de piel de esclerodermia				*	*					
Realización de técnicas de ST de célula única de biomodelos de la LA2					*	*	*			
Creación de la Unidad de Bioinformática de coordinación. Recopilación común de pipelines de análisis sc-ómicos necesarios. Protocolo de almacenamiento y desarrollo técnico.	*	*								
Incorporación de personal de refuerzo en la Unidad de Bioinformática del IPBLN y formación			*	*						
Desarrollo de protocolos armonizados de datos "sc-ómicos"				*	*	*	*			
Optimización de flujos de trabajo para la toma de decisiones clínicas en tiempo real.				*	*	*				
Análisis comparativo de datos, interpretación y obtención de información biológica relevante entre líneas de actuación.					*	*	*	*	*	*
Creación de un gabinete asesor y de formación para extender el uso de las técnicas analíticas desarrolladas en el ámbito clínico y biomédico.						*	*	*	*	*
Difusión y visualización de progresos y resultados			*	*	*	*	*	*	*	*

● **Grado de innovación de la propuesta y justificación del potencial impacto científico y/o tecnológico del proyecto.**

Este proyecto se caracteriza principalmente por la incorporación de nuevas tecnologías que por su diseño y capacidades van a revolucionar la medicina personalizada en los próximos años. Con este proyecto planteamos que Andalucía y los andaluces sean los primeros en disponer de las mejores herramientas para el diagnóstico clínico y la investigación biomédica. La secuenciación "ómica" de resolución espacial mejorada (Stereo-seq) no está aún disponible en ninguna institución pública española y por eso creemos importante apostar por una plataforma biotecnológica única, de proyección futura. En el campo biomédico, la ST tiene un enorme impacto científico dado que posibilita avanzar nuestra comprensión de una amplia gama de enfermedades y trastornos, como el cáncer, las enfermedades neurodegenerativas, cardiovasculares y autoinmunes, entre otros. En cuanto a su impacto en la esclerodermia, la ST tiene un importante papel en la identificación de nuevos genes y procesos celulares involucrados en el trastorno del tejido conectivo que afecta a órganos y sistemas en el cuerpo, lo que lo hace extrapolable a otras enfermedades inmunológicas, así como patologías de la piel como el melanoma.

El análisis computacional de los datos generados por esta técnica, proporciona una imagen más completa de los patrones de expresión génica, lo que ayuda a los investigadores a identificar nuevas dianas terapéuticas y desarrollar tratamientos más eficaces para diversas enfermedades ayudando a la traslación de los resultados. En consecuencia, los protocolos de análisis son complejos y costosos computacionalmente, precisando de personal cualificado para el desarrollo y optimización de los flujos de trabajo que permitan obtener diagnósticos a "tiempo real".

Asimismo, cabe destacar que esta propuesta cuenta con un equipo de investigación formado por científicos de primer nivel con experiencia en la traslación de resultados a la clínica. Las muestras de tejido de los pacientes de esclerodermia proceden del grupo del profesor de investigación Javier Martín, reconocido investigador a nivel mundial en el estudio de las bases genéticas de las enfermedades autoinmunes, coordinador del Servicio de Bioinformática del IPBLN, y actúa como IP1 de este proyecto. Las muestras de tejido sano serán proporcionadas por diferentes colaboradores, como el Dr. José Luis Calleja, facultativo de la Unidad de Enfermedades Sistémicas del Servicio de Medicina Interna del HUSC, el Dr. Ricardo Ruiz Villaverde, jefe del servicio de Dermatología del HUSC e investigador responsable del grupo de Enfermedades Inmunomediadas de la Piel del IBS, así como el Dr. Norberto Ortego, Catedrático de Medicina de la UGR, jefe de la Unidad de

Enfermedades Autoinmunes Sistémicas del HUSC e investigador responsable del grupo de bases de fisiopatología y terapéutica médica del área de medicina de precisión del IBS. Además, el proyecto cuenta con la Dra. Luz Canet Antequera de la unidad de genómica del IPBLN, con una amplia experiencia en el desarrollo de protocolos experimentales, así como del Dr. Eduardo Andrés-León, responsable de la unidad de Bioinformática, que cuenta con más de 20 años de experiencia en el diseño de flujos de trabajo para el análisis automatizado y optimizado, de datos de secuenciación masiva en entornos de computación científica de alto rendimiento, y actúa como IP2 de este proyecto.

- **Resultados esperados a los que contribuye dentro de la línea de actuación. Plan de difusión y explotación.**

La LA3 se enfocará en el desarrollo de protocolos experimentales y bioinformáticos estandarizados, que se utilizarán para abordar problemas identificados de manera coordinada dentro del consorcio BAS a través de Proyectos específicos de Prueba de Concepto. Esto proporcionará nuevas capacidades tecnológicas a los centros de investigación en Andalucía para el estudio de patologías. En particular, se creará una unidad coordinada de análisis "sc-omics" para impulsar el creciente uso de protocolos de análisis en centros biomédicos y hospitalarios que implementen o requieran técnicas "sc-omics". Los objetivos finales incluyen la caracterización y el análisis rápido de muestras de pacientes para ayudar en la toma de decisiones clínicas en tiempo real y monitorizar la respuesta a dichas decisiones mediante una iteración clínica/análisis, así como integrar el conocimiento para descubrir y desarrollar nuevas soluciones terapéuticas y mejorar su aplicación en la clínica.

La difusión y explotación de los resultados y conclusiones del proyecto tendrán un impacto directo en la salud de los andaluces, por lo que se planea llegar a una amplia gama de públicos relevantes, incluyendo la comunidad académica, la industria, las organizaciones gubernamentales y no gubernamentales, asociaciones de enfermos y la sociedad en general. Esto se realizará a través de publicaciones científicas punteras de carácter científico técnico y hacia la sociedad desde de una página web, en la que cada subproyecto pueda informar de sus logros y avances, en pos de la transparencia.

Por último, las tecnologías innovadoras que se crearán en este proyecto ya han despertado el interés de numerosos investigadores, por lo que existirá una explotación futura de las técnicas desarrolladas. Entre esos investigadores, podemos destacar: Dra. Marta Alarcón Riquelme (FPS-GENyO), Dr. Jesús Gustavo Ponce González (Universidad de Cádiz) y Dr. David M. Landeira Frías (UGR-GENyO)

- **Implicaciones éticas y/o de seguridad.**

Las muestras para la puesta a punto proceden del esfuerzo colaborativo entre el IPBLN, el HUSC y el IBS por parte de los investigadores Dr. Javier Martín, Dr. José Luis Calleja, Dr. Ricardo Ruiz y el Dr. Norberto Ortego. Todas las muestras cuentan con la aprobación de los pertinentes comités éticos de la Junta de Andalucía y, parten de un consentimiento informado a los individuos que proporcionarán las muestras. En caso necesario, el tratamiento de los datos y su intercambio, en el contexto del proyecto, seguirá el marco legal vigente en materia de protección de datos (Ley Orgánica 7/2021, de 26 de mayo, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales). El proyecto, además, se someterá al comité de ética de la institución de origen de las muestras, así como al del CSIC.

- **Presupuesto detallado del proyecto y justificación del mismo.**

Personal (52.500€): 22% del presupuesto, distorsión producida por la compra de equipamiento. El IPBLN contribuye con personal altamente cualificado de las unidades de genómica y bioinformática. Sin embargo, la tecnología a implantar es única en el ámbito científico público lo que representa un desafío que exigirá reforzar la plantilla de las unidades de genómica y bioinformática del IPBLN. Por lo tanto, será necesario contratar durante 18 meses una persona con un título de grado en ciencias de la vida y experiencia en el análisis bioinformático de grandes datos de secuenciación, así como conocimiento en la programación en entornos de computación científica, para fortalecer el equipo de trabajo de la unidad de Bioinformática en las tareas de desarrollo necesarias para estos análisis.

Equipamiento/infraestructura (193.620€): 77% del presupuesto, desviación necesaria dado el coste del equipo a instalar. Este contempla la adquisición de un equipamiento singular (Secuenciador MGI para Stereo-seq, 155.000€) y los fungibles necesarios para su puesta a punto (19.310€/año): preparación y tinción de cortes de tejido (2.500€), reactivos de ST (15.120€), reactivos de SCS (7.000€), reactivos de secuenciación (10.000€), fungible general de laboratorio (2.000€), fungible informático (2.000€). Este alto porcentaje se debe principalmente a la incorporación de una nueva tecnología que es crucial para el desarrollo de la línea de actuación y para el avance futuro de la medicina de precisión. Además, el coste de los reactivos de una técnica recién lanzada al mercado es siempre muy elevado. Este equipamiento es crítico y necesario para la consecución de los objetivos marcados en la convocatoria.

Otros: 1.500€ (costes de movilidad, envío de muestras, informe del auditor, otros costes, etc.). 1% del presupuesto, no hay desviación.

Costes indirectos (15% costes directos): 37.143€. **PRESUPUESTO TOTAL: 284.763€.**

Proyecto 5: IMPLEMENTACIÓN DE TECNOLOGÍAS DE TRANSCRIPTÓMICA A NIVEL DE CÉLULA ÚNICA A PARTIR DE MUESTRAS CLÍNICAS

IP: Javier López-Ríos Moreno; Institución: Centro Andaluz de Biología del Desarrollo (CABD-CSIC)

- **Resumen de la propuesta:** La tecnología de transcriptómica a nivel de célula única (scRNA-seq) está revolucionando el estudio de los procesos biológicos al proporcionar un nivel de resolución sin igual a la hora de definir la complejidad celular de tejidos. Estas tecnologías, de rápida implantación en el ámbito académico, están permitiendo un entendimiento fundamental de los procesos del desarrollo, la homeostasis tisular, el envejecimiento de los organismos y las bases moleculares y celulares de múltiples procesos patológicos. El objetivo de la presente propuesta es capitalizar la extensa experiencia en la generación y análisis de datos de scRNAseq por parte de los Servicios de Genómica y Bioinformática del Centro Andaluz de Biología del Desarrollo (CABD, Sevilla) para la implementación de protocolos robustos que permitan acceder a estudios de tipo scRNAseq a partir de muestras humanas de diferente origen (biopsias, piezas quirúrgicas, xenoinjertos derivados de pacientes), incluyendo muestras fijadas y parafinadas, el tipo de muestra más compleja para procesar, pero también el más relevante para la transferencia al uso clínico. A la finalización del proyecto, la plataforma tecnológica resultante dispondrá de las capacidades científico-técnicas para extender sus servicios a centros de investigación biomédica y hospitalarios que precisen de este tipo de tecnología en estrategias de medicina de precisión (MP).

- **Antecedentes del proyecto**

El estudio anatomopatológico de biopsias y resecciones quirúrgicas constituye una herramienta clínica fundamental para el diagnóstico y pronóstico de las enfermedades humanas. Los Servicios de Anatomía Patológica han ido complementando el análisis histopatológico de la morfología y composición tisular y celular con técnicas inmunohistoquímicas y de patología molecular, incluyendo el uso de secuenciación masiva de paneles de genes implicados en patologías concretas con el fin de identificar a nivel genómico mutaciones específicas o variantes de riesgo. La significancia clínica de estas variantes o sus combinaciones permite obtener información sobre el diagnóstico del paciente, la prognosis de la enfermedad o la respuesta de la misma a tratamientos terapéuticos específicos. Por ejemplo, estas aproximaciones moleculares han revolucionado en las últimas décadas la capacidad de definir nuevos subtipos patológicos de tumores, impactando directamente en la elección de tratamiento y en la información que recibe el paciente para hacer frente a su enfermedad. Más recientemente, se ha introducido el análisis transcriptómico global a partir de muestras homogenizadas (RNA-seq; [1]), con el fin de identificar genes diferencialmente expresados entre el tejido patológico y el sano. Este tipo de estudios, todavía no rutinario en la clínica, presentan la desventaja conceptual de que, al realizarse a partir de tejidos complejos, carecen de resolución celular. Este inconveniente es particularmente difícil de soslayar en el caso de tumores, al ser estos tejidos altamente heterogéneos y en los que ciertos tipos celulares de potencial relevancia como dianas terapéuticas, como las células *stem* cancerosas, son muy poco frecuentes [2].

Las técnicas de transcriptómica global con resolución celular basados en microfluídica (scRNA-seq; [3]), de muy reciente desarrollo e implementación en centros de investigación, están permitiendo caracterizar con alta resolución la complejidad celular de tejidos en múltiples contextos biológicos, tanto normales como patológicos [4, 5]. La capacidad técnica de marcar todo el transcriptoma celular con “códigos de barras” específicos de célula (*barcodes*) y de molécula (UMI: *unique molecule identifier*) permite, tras la secuenciación masiva de las genotecas resultantes, asignar cada transcrito a una célula concreta y la cuantificación lineal de las moléculas de RNA en las miles de células recuperadas en el proceso. El análisis bioinformático subsiguiente, basado en la similitud de los perfiles transcriptómicos individuales, permite agrupar las células en grupos o *clusters*, los cuales recogen desde la identidad celular (tipos celulares, tanto de progenitores hasta células terminalmente diferenciadas), el estado metabólico o la fase de ciclo celular. Cada cluster se anota en base a su perfil de expresión diferencial respecto al resto y la información disponible en las bases de datos. La resolución de este tipo de tecnologías es tal que *clusters* de pocas decenas de células pueden ser claramente identificados y sus perfiles moleculares extraídos. El desarrollo continuo de nuevas herramientas bioinformáticas permite además inferir información sobre las relaciones ontogénicas de las distintas subpoblaciones o la identificación de redes de regulación activas en cada uno de los *clusters*. El Centro Andaluz de Biología del Desarrollo (CABD, Unidad de Excelencia María de Maeztu 2022-2025) dispone de experiencia extensa en la generación y análisis de datos de scRNA-seq a través de su Servicio de Genómica Funcional (SGF), cuyo director científico es el solicitante, y su Servicio de Bioinformática. Desde septiembre de 2020, el SGF del CABD ha procesado más de 70 genotecas de scRNA-seq, a partir de muestras de ocho especies diferentes (seis vertebrados, incluyendo humano y ratón, y dos invertebrados), lo que coloca al CABD en una posición de liderazgo para el desarrollo e implementación de tecnologías de scRNA-seq en estrategias de medicina de precisión (MP).

- **Objetivos del proyecto**

Las tecnologías de scRNA-seq se están extendiendo con rapidez en el ámbito académico, pero no se han incorporado todavía a la práctica clínica. Entre las principales razones de ello están: (i) la falta de desarrollo de protocolos estandarizados y robustos para el procesamiento de muestras de origen clínico, incluyendo muestras fijadas y parafinadas; (ii) el alto coste por muestra; (iii) la ausencia de experiencia en análisis bioinformático que permita la transferencia eficiente e informativa de los resultados de scRNA-seq al ámbito clínico. El presente proyecto capitalizará la dilatada experiencia del CABD en la generación y análisis de datos de scRNA-seq para dar soluciones a estos problemas principales, contribuyendo a la generación de una estructura de conocimiento y capacidades técnicas que permita la implementación de estas tecnologías en el ámbito clínico, tanto a nivel autonómico como nacional. Los objetivos específicos son:

- a) Desarrollo y estandarización de protocolos robustos de aislamiento de suspensiones de células o núcleos de alta calidad para su procesamiento para scRNA-seq a partir de muestras fijadas, incluyendo secciones parafinadas de tejidos humanos.
- b) Instalación y puesta en marcha en el SGF del CABD de un equipo para aplicaciones de célula única con capacidad de procesamiento de muestra fijada y optimización de los procedimientos de combinación de muestras durante la generación y secuenciación de genotecas.
- c) Desarrollo de *pipelines* bioinformáticos de fácil implementación que permitan el procesamiento de datos de secuenciación masiva de genotecas de scRNA-seq y otras ómicas y la generación de informes para la práctica clínica sobre el perfil molecular de las muestras de manera que puedan resultar informativos para la diagnosis, prognosis y la elección de alternativa terapéutica o la respuesta de la enfermedad a la misma.
- d) Integración con el resto de subproyectos en la línea de actuación 3 y con el Proyecto de Dirección Científica para el análisis mediante scRNA-seq y otras ómicas de xenoinjertos derivados de pacientes (PDX) de sarcoma de Ewing (suministrados por investigadores del IBIIS, Sevilla), estableciendo una prueba de principio para el uso de estas tecnologías en medicina de precisión y personalizada.
- e) Compilación de un catálogo vivo de las capacidades de procesamiento y análisis de muestras de scRNA-seq y otras tecnológicas ómicas con resolución celular en la Comunidad Autónoma Andaluza, estableciendo una colaboración transversal entre Servicios de Genómica Funcional y de Bioinformática para el intercambio de protocolos, herramientas y capacidades formativas.

- **Contribución del proyecto a la línea de actuación.**

El presente proyecto implementará la tecnología de scRNA-seq, siendo los perfiles de expresión génica los más inmediatamente interpretables en el contexto clínico (**Actuación 3.3; objetivos a-d**). Además, desarrollaremos protocolos estandarizados y robustos para el procesamiento y análisis de muestras clínicas humanas que puedan ser directamente importados por otros Servicios tanto académicos como sanitarios (**Actuación 3.2; objetivos a-c y e**). El solicitante se compromete a que las capacidades generadas, unidas a las ya presentes, se compartirán y utilizarán para priorizar proyectos transversales en los que se aplicará la tecnología de scRNA-seq y otras ómicas, en coordinación con el resto de proyectos del consorcio Andalucía-Biotec Salud y el de Dirección Científica, y en colaboración con otros agentes andaluces (SSPA/OPIs/Universidades) y de otras CCAA (**Actuación 3.1; objetivo d y e**). El solicitante participará como representante andaluz en el comité nacional correspondiente.

- **Metodología y plan de trabajo. Cronograma.**

Preparación de muestras y genotecas para scRNA-seq (objetivos a-b, d). La puesta a punto de la preparación de suspensiones celulares/de núcleos (objetivo a) se realizará a partir de muestras de 3-5 órganos (tejidos blandos o fibrosos, tejidos con matriz extracelular abundante, etc.) y/o distinto origen clínico (biopsias/piezas quirúrgicas/PDXs). Los tejidos se obtendrán de biobancos y nuestra red de colaboradores (ver listado más abajo) o se aislarán a partir de ratón, lo que nos permitirá optimizar protocolos para diferentes órganos y condiciones de procesamiento sin recurrir a valiosas muestras humanas. Dada la variedad de tejidos y métodos de preservación, se primará la optimización de extracción de células/núcleos a partir de tejidos fijados, incluyendo aquellos incluidos en parafina o congelados, tomando como punto de partida protocolos indicados por el fabricante (10x Genomics) o recientemente descritos [6], explorando la capacidad de indexación de muestras para optimizar costes. Las muestras se disociarán utilizando mezclas enzimáticas *ad hoc* (eg. liberasa TL, liberasa TL+colagenasa, etc.) y se optimizará la digestión manual versus automatizada (gentleMACS Octo Dissociator). La calidad de la extracción de muestras se determinará mediante conteo automatizado de células, la evaluación de su integridad, rendimiento, presencia de restos celulares, etc. El equipo presupuestado (Chromium iX; objetivo b) es actualmente el único equipo en el mercado con capacidades técnicas de procesar muestras clínicas fijadas (ie. secciones de parafina) para scRNA-seq, al recurrir a paneles de sondas con cobertura global del transcriptoma (Chromium Fixed RNA Kit, Human Transcriptome). Una vez que dicho equipo esté instalado en el SGF del CABD, se explorarán las condiciones de multiplexado de muestras y genotecas para su secuenciación (25.000 pares de lecturas por célula; NovaSeq 6000; CABIMER) para optimizar costes, considerando que se cargarán 8.000-10.000 células por muestra, lo que permitirá detectar poblaciones poco frecuentes.

Análisis bioinformático de datos de scRNA-seq (objetivos c, d). El alineamiento, normalización y QC de los datos de secuenciación de genotecas de scRNA-seq se llevarán a cabo mediante herramientas estándar (Cell Ranger; paquete Seurat R; [7], que permiten además el curado de datos (ej. eliminación de células muertas y dobletes). Seurat incorpora también herramientas para la reducción de la dimensionalidad, generación de *clusters* basados en la similitud de los perfiles de expresión génica, determinación de estado de ciclo celular e integración de datos multiómicos. La anotación automatizada de *clusters* [4, 8] será curada manualmente. Los estudios de trayectoria se integrarán mediante herramientas como STREAM [9], mientras que para los análisis de ontología génica/redes de regulación se usarán algoritmos como SCENIC [10].

Además del IP del proyecto, el Dr. J López-Ríos, y del equipo de investigación formado por el Dr. I Cases (responsable del Servicio de Bioinformática del CABD) y la Dra. Ana Fernández-Miñán (responsable del Servicio de Genómica Funcional del CABD), colaborarán en la ejecución del proyecto tanto con el aporte de muestras como en el análisis bioinformático los siguientes investigadores: Enrique de Álava (Jefe de Servicio de Anatomía Patológica Hosp. Virgen del Rocío e IP IBIIS, Sevilla). Diego Franco (Catedrático de Biología Celular; Univ. de Jaén); M^a José Sánchez (Científico Titular, CABD-CSIC); Patricia Ybot (Científico Titular, IBIIS, Sevilla); Jesús G. Ponce (Profesor Titular; Univ. de Cádiz); Ana B. Muñoz-Manchado (Investigadora Ramón y Cajal; Univ. de Cádiz); Juan A. G. Ranea (Catedrático de Biología Molecular y Bioquímica; Univ. de Málaga).

CRONOGRAMA	AÑO 1		AÑO 2	
	Meses 1-6	Meses 6-12	Meses 13-18	Meses 19-24
Objetivos específicos				
Desarrollo y estandarización de protocolos de preparación de muestras para scRNA-seq (objetivo a). Personal: NN1, AFM, JLR				
Instalación y puesta en marcha del equipo de scRNA-seq con capacidades para muestra fijada (objetivo b). Personal: NN1, AFM, JLR				
Desarrollo de pipelines bioinformáticos para el análisis de muestras scRNA-seq (objetivo c). Personal: NN2, IC, JLR				
Integración subproyectos de técnicas a nivel de célula única: prueba de concepto (objetivo d). Personal: AFM, NN2, IC, JLR				
Generación de herramientas colaborativas: catálogo de capacidades andaluzas a nivel de célula única (objetivo e). Personal: JLR				
Acciones adicionales				
Reuniones presenciales/online para articular acciones colaborativas entre proyectos de la línea de actuación 3 (2-3/año)				
Participación del IP como representante andaluz en el comité nacional de Biotecnología aplicada a la Salud correspondiente (2-3/año)				
Depósito de datos en repositorios públicos				
<i>NN1: titulado superior contratado con cargo al proyecto (tareas experimentales); NN2: titulado superior contratado con cargo al proyecto (análisis bioinformático); AFM: Dra. Ana Fernández-Miñán (responsable Servicio de Genómica Funcional CABD); IC: Dr. Ildefonso Cases (responsable Servicio de Bioinformática CABD); JLR: Javier López-Ríos</i>				

Bibliografía: 1. Byron, S.A., et al. *Nat Rev Genet*, 2016. 17(5):257-71. 2. Battle, E. and H. Clevers. *Nat Med*, 2017. 23(10):1124-1134. 3. Zheng, G.X., et al. *Nat Commun*, 2017. 8:14049. 4. Tabula Sapiens, C., et al. *Science*, 2022. 376(6594): eab4896. 5. Neftel, C., et al. *Cell*, 2019. 178(4):835-849 e21. 6. Vallejo, A.F., et al. *bioRxiv*, 2022. doi.org/ 10.1101/2022.08.23.505054. 7. Satija, R., et al. *Nat Biotechnol*, 2015. 33(5):495-502. 8. Pasquini, G., et al. *Comput Struct Biotechnol J*, 2021. 19:961-969. 9. Chen, H., et al. *Nat Commun*, 2019. 10(1):1903. 10. Van de Sande, B., et al. *Nat Protoc*, 2020. 15(7):2247-2276.

- **Grado de innovación de la propuesta y justificación del potencial impacto científico y/o tecnológico del proyecto.**

Las tecnologías de célula única suponen una revolución en MP, puesto que revelan la heterogeneidad del tejido a escala genómica, lo que permitirá la definición de nuevos subtipos patológicos con diagnósticos y pronósticos diferenciales, así como su respuesta a tratamientos concretos. Este tipo de métodos no son directamente implementables en Servicios Clínicos, sino que requieren de desarrollo tecnológico previo en centros de investigación punteros, que sean expertos en los mismos y desarrollen protocolos sólidos para la rutina clínica que además optimicen costes de manera que la tecnología pueda ser verdaderamente accesible. Además, la complejidad de los datos requiere de la elaboración de informes concretos e informativos, con datos relevantes para la clínica (ej. en tumores, perfiles moleculares de poblaciones proliferativas, invasivas, del estroma o de células *stem* cancerosas, así como de las relaciones ontogénicas entre subpoblaciones).

- **Resultados esperados a los que contribuye dentro de la línea de actuación. Plan de difusión y explotación.**

Como parte de las acciones descritas, se generarán: (i) protocolos experimentales robustos para el procesamiento de muestras para scRNA-seq a partir de muestras clínicas, con especial atención al coste/célula, los cuales serán compartidos mediante publicación en revistas académicas/repositorios técnicos de acceso abierto (<https://www.protocols.io>; DIGITAL.CSIC, etc.), siguiendo los principios FAIR (Findable, Accesible, Interoperable and Reusable); (ii) *pipelines* bioinformáticos igualmente compartidos mediante repositorios dedicados (ej. GitHub) o institucionales; (iii) datos de scRNA-seq a partir de muestras clínicas depositados en bases de datos abiertas como GEO. En el caso de que, además de desarrollo tecnológico, se deriven publicaciones académicas adicionales, estas serán igualmente de acceso abierto; (iv) catálogo publicado en la web del consorcio Andalucía-Biotec Salud que compile las capacidades andaluzas en tecnologías de célula única. Dicho catálogo será también la base para canales de colaboración basados en listas de correo que permitan a los usuarios compartir protocolos, realizar consultas técnicas, anunciar convocatorias/ofertas de trabajo/cursos de formación, etc.

- **Implicaciones éticas y/o de seguridad.**

En el presente proyecto se utilizarán muestras humanas proporcionadas por la Red de Biobancos de Andalucía y/o Servicios de Anatomía Patológica del SSPA, según la Ley 14/2007, de 3 de julio, de Investigación Biomédica y precedido en cada caso por la evaluación favorable del Comité de Ética de Investigación correspondiente, y su tramitación vía el Portal de Ética de la Investigación Biomédica de Andalucía. Por otro lado, también se utilizarán tejidos de ratón de laboratorio para el establecimiento de protocolos de extracción de células o núcleos para scRNAseq, si bien no se realizará ningún procedimiento de experimentación animal, ni se utilizarán animales modificados genéticamente. El animalario del CABD está registrado con número REGA ES410910008004 y todo el personal implicado en la manipulación y eutanasia de animales ostenta la capacitación exigida (funciones a, b y c según Orden ECC/566/2015, de 20 de marzo).

- **Presupuesto detallado del proyecto y justificación del mismo**

Personal (82.000€): Se solicita un titulado universitario (7 meses, al inicio del proyecto, 22.000€), encargado de la optimización de los protocolos de preparación de muestras para scRNA-seq bajo la supervisión de JLR/AFM. La preparación rutinaria de muestras ya optimizadas durante el resto del proyecto será tarea del SGF. Se solicita además un titulado universitario (18 meses, 60.000€), con formación en R y supervisado por JLR/IC, encargado de implementar los *pipelines* bioinformáticos para análisis de datos de scRNAseq y la elaboración de informes semi-automatizados de tipo R Markdown para la comunicación con Servicios Clínicos. *28% del presupuesto, distorsión producida por la compra de equipamiento.*

Equipamiento/infraestructura (160.000€): 95.000€ (adquisición equipo célula única fijada para el SGF del CABD); 65.000€ (fungible: compra de reactivos para la preparación de muestras y genotecas para scRNA-seq). *54% del presupuesto, desviación necesaria dado el coste del equipo a instalar.*

Otros: 15.000€ (gastos de servicios centrales y externos, incluidos secuenciación, costes de movilidad, informe del auditor, otros costes, etc.). *5% del presupuesto, no hay desviación.*

Costes indirectos (15% costes directos): 38.550€. **PRESUPUESTO TOTAL: 295.550€.**

Proyecto 6: IMPLEMENTACIÓN DE TECNOLOGÍAS DE ACCESIBILIDAD DE LA CROMATINA A NIVEL DE CÉLULA ÚNICA EN ESTRATEGIAS DE MEDICINA PERSONALIZADA

IP: José Carlos Reyes; Institución: Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER)

- **Resumen de la propuesta:** Las tecnologías OMICAS de célula única constituyen un cambio de paradigma para aproximarse a la comprensión del funcionamiento de las diferentes escalas organizativas que constituyen un ser vivo, desde la escala celular a la de órgano. La aplicación de estas tecnologías a la clínica está en sus comienzos pero tiene un incuestionable futuro especialmente en algunas áreas clínicas donde comprender la heterogeneidad de la población celular es esencial. Un ejemplo paradigmático es la oncología donde la heterogeneidad genética y epigenética de las células tumorales origina desde comportamientos metastáticos a resistencias a tratamientos. La tecnología denominada ATAC-seq de célula única (scATAC-seq) permite el conocimiento de las regiones accesibles del genoma, que en su mayoría son las regiones reguladoras: promotores, potenciadores y aisladores. Estos elementos son específicos del tipo celular y cambian según el grado de diferenciación celular, así como por efecto de rutas de señalización. Conocer el repertorio de elementos reguladores accesibles de una célula es tanto como conocer sus potencialidades y características, y es complementario a conocer su transcriptoma y su proteoma. El objetivo general de este subproyecto es implementar metodologías de scATAC-seq en muestras de origen clínico o modelos derivados de ellas, por lo que se alinea completamente con la LÍNEA DE ACTUACIÓN 3 (LA3). Como prueba de concepto usaremos xenoinjertos derivados de pacientes (PDX), un modelo de gran interés para dilucidar nuevos tratamientos y biomarcadores en oncología. Los resultados se pondrán en común y se compararán con los de otros subproyectos que usen diferentes tecnologías de célula única.

- **Antecedentes del Proyecto**

Tecnologías OMICAS de célula única (sc-OMICS): Las técnicas clásicas de anatomía patológica usadas rutinariamente en la clínica se basan en las diferentes morfologías y características de tinción de los diferentes tipos celulares y tejidos, y en el uso de anticuerpos contra determinados marcadores. Sin embargo, las recientes técnicas de análisis del transcriptoma (scRNA-seq) y del epigenoma a nivel de célula única sugieren que la cantidad de tipos celulares existente es muy superior a la previamente identificada por técnicas clásicas (1). Estas técnicas están revelando una gran heterogeneidad dentro de cada tipo celular cuyas consecuencias en la clínica están solo comenzando a ser evaluadas. Un área donde la heterogeneidad celular tiene una importancia innegable es en oncología. Tras el descubrimiento de que un mismo tumor a menudo contiene diferentes clones genéticamente diferentes (2), publicaciones recientes demuestran también una alta heterogeneidad epigenética superpuesta a la heterogeneidad genética. Así, diferentes grados de diferenciación, o diferentes estados en el gradiente epitelio-mesénquima se superponen provocando diferentes comportamientos celulares que influyen el grado de metastización (3, 4) y de resistencia ante terapias (5, 6). Esta información es especialmente importante para la ayuda en la toma de decisiones clínicas cuando las características de una población mínima de células pueda tener enormes consecuencias mediante su expansión clonal, como ocurre con poblaciones de células troncales cancerosas (CSC), o de célula que hayan llevado a cabo la transición epitelio-mesénquima (EMT) y puedan delaminarse de un tumor primario. Por otra parte, las sc-OMICS ofrecen una información sin precedentes acerca del microambiente tumoral y el grado de infiltración (7).

Análisis de la accesibilidad de la cromatina en células únicas: Los distintos tipos celulares se caracterizan por diferentes patrones de expresión génica que, a su vez, vienen en gran parte determinados por la regulación de la transcripción. Dicha regulación se lleva a cabo mediante un conjunto de regiones genómicas reguladoras de la transcripción, siendo los promotores, potenciadores (*enhancers*) y aisladores los mejor conocidos. El conjunto de promotores y *enhancers* activos o disponibles para ser activados (*primed*) son característicos de cada tipo celular, se establecen durante el proceso de diferenciación celular y se caracterizan por ser regiones accesibles de la cromatina a las que se unen factores transcripcionales específicos de secuencia (TF) y factores generales de la transcripción. Mutaciones en estos elementos o su desregulación epigenética está en la base de multitud de enfermedades congénitas y el cáncer (8). El conocimiento de la fracción accesible del genoma de una célula nos informa, por lo tanto, de su programa de regulación génica, de qué tipo celular es, en qué estado de diferenciación e incluso en qué estado transcripcional se encuentra. La técnica denominada ATAC-seq determina las regiones accesibles del genoma al salto de la transposasa Tn5 y revela básicamente *enhancers*, promotores y elementos aisladores del genoma (9). Las tecnologías de ATAC-seq adaptadas para célula única (scATAC-seq) complementan las técnicas de scRNA-seq y ofrecen en ocasiones una información menos sujeta a ruido transcripcional pero más difícil de anotar.

- **Objetivos del Proyecto**

Las sc-OMICS están poco implantadas por el momento en Andalucía debido entre otros motivos al elevado coste, los complejos métodos de análisis bioinformático requeridos y el limitado conocimiento de la tecnología en las Unidades de Genómica de los centros de Investigación. La implantación en los servicios hospitalarios es aún menor. El **objetivo general** de este subproyecto es la implementación de las tecnologías de scATAC-seq, su aplicación a muestras clínicas, la estandarización de protocolos que se pondrán a disposición de la comunidad científica y la comparación con otras tecnologías sc-OMICS (scRNA-seq, sc-proteomics, spatial transcriptomics) desarrolladas en otros subproyectos de este proyecto coordinado.

Los objetivos específicos son los siguientes:

- Objetivo 1: Implementar la tecnología de scATAC-seq en la Unidad de Genómica de CABIMER.

- Objetivo 2: Aplicación de scATAC-seq a modelos de xenoinjertos derivados de pacientes (PDX) de sarcoma de Ewing y gliomas.
- Objetivo 3: Generación de pipelines de análisis informático para análisis de datos de scATAC-seq e integración con otros datos de sc-OMICS.

- **Contribución del proyecto a la línea de actuación.**

El proyecto se adecua perfectamente a la LA3: "Caracterización ómica de muestras y biomodelos a escala de célula única". Además, en este subproyecto nos coordinaremos con otros subproyectos del Nodo Occidental de la LA3, mediante el uso del mismo modelo experimental (PDX de sarcoma de Ewing). Esto nos permitirá investigar hasta qué punto las tecnologías de scRNA-seq, scATAC-seq y scProteomics son solapantes o complementarias, investigar software bioinformáticos de integración de datos e intercambiar protocolos para los pasos primeros de la experimentación que en todos los casos requieren la disgregación del tejido fresco hasta célula única viable. Nuestros datos y nuestros protocolos experimentales serán puestos en común con los otros subproyectos tan pronto como sean obtenidos.

Manifestamos nuestro compromiso expreso de la colaboración del proyecto a las actuaciones conjuntas que emanen del proyecto de dirección científica, tanto en Andalucía como a las actuaciones conjuntas con otras comunidades autónomas. Nuestros datos y nuestros protocolos experimentales serán puestos en común con los otros subproyectos tan pronto como sean obtenidos.

- **Metodología y Plan de Trabajo. Cronograma.**

Objetivo 1.- Implementar la tecnología de scATAC-seq en la Unidad de Genómica de CABIMER. A) Preparación de suspensiones de célula única de varios tejidos tumorales diferentes. Dado que las muestras de tumores humanos son enormemente preciadas, las primeras pruebas se harán con modelos murinos. B) Realización de librerías de scATAC-seq mediante la tecnología de 10X-Genomics en la Unidad de Genómica de CABIMER que dispone de un equipo Chromium. C) Secuenciación de las librerías en la Unidad de Genómica de CABIMER mediante el equipo Novaseq de Illumina. *Entregable 1: Estandarización de protocolos para scATAC-seq.*

Objetivo 2.- Aplicación de scATAC-seq a modelos de xenoinjertos derivados de pacientes (PDX). Las tecnologías de PDX están avanzando como excelentes modelos que recapitulan parcialmente las características de un tumor concreto y permiten testar terapias de forma personalizada (10). Una importante ventaja de los PDX es la disponibilidad de una importante cantidad de tejido fresco sin necesidad de biopsias de gran tamaño o de repetidas biopsias. Utilizaremos tejido de PDX para la realización de scATAC-seq y la comparación con scRNA-seq, sc-proteomics llevados a cabo en otros subproyectos, usando la misma muestra de tejido para las tres tecnologías. Disponemos de PDX de sarcoma de Ewing procedentes del laboratorio del Dr. Enrique de Álava (IBIS, Sevilla) y de gliomas humanos procedentes del laboratorio de Vivian Capilla González (FPS, CABIMER, Sevilla). *Entregable 2: Generación de datos de scATAC-seq de PDX.*

Objetivo 3.- Generación de pipelines de análisis informático para análisis de datos de scATAC-seq e integración con otros datos de sc-OMICS Análisis bioinformático de las muestras, *clustering* para identificar tipos celulares, estudio de trayectorias. Análisis de heterogeneidad intra y intertumoral. Estos análisis los haremos con las capacidades bioinformáticas internas. Además en colaboración con otros subproyectos usaremos métodos de integración de datos ÓMICOS de varios orígenes. *Entregable 3: Comparación e Integración de datos de diferentes sc-OMICS.*

Cronograma

Objetivos	Semestre 1	Semestre 2	Semestre 3	Semestre 4
Objetivo 1.- Implementar la tecnología de scATAC-seq	*			
Objetivo 2.- Aplicación de scATAC-seq a PDX			**	
Objetivo 3.- Generación de pipelines de análisis informático e Integración sc-OMICS				***

*Entregable 1. Estandarización de protocolos para scATAC-seq; ** Entregable 2. Generación de datos de scATAC-seq de PDX; *** Entregable 3 Comparación e Integración de datos de diferentes sc-OMICS

Además del IP del proyecto, el Dr JC Reyes y el equipo de investigación formado por el Dr. Andrés Aguilera y la Dra. Vivian Capilla, todos ellos investigadores del CABIMER, colaborarán en el proyecto los siguientes investigadores: E. de Álava del Hosp. Virgen del Rocío-IBIS, Sevilla. Benoît Gauthier, I. Pineda Torra, y F. Díaz-Corrales de FPS-CABIMER, Sevilla. A.J. López-Contreras, M. Álvarez-Dolado del CABIMER-CSIC, Sevilla. D. Landeira de Universidad de Granada-GENYO. D. Pozo de Universidad de Sevilla-CABIMER, A. Rojas Univ. Pablo de Olavide-CABIMER. Elosia Andujar Pulido y Mónica Pérez Alegre de la Unidad de genómica de CABIMER.

- **Grado de innovación de la propuesta y justificación del potencial impacto científico y/o tecnológico del proyecto.**

La propuesta pretende implementar tecnologías absolutamente novedosas, aún muy poco desarrolladas, en nuestro ámbito regional. El futuro de estas tecnologías en el campo de la biomedicina en estudios básicos se está poniendo de manifiesto con una enorme cantidad de publicaciones de altísimo nivel en los últimos años. En el campo clínico y sobre todo en oncología la identificación de pequeñísimas poblaciones con características metastáticas o de resistencia a determinados quimioterapias podría condicionar completamente los tratamientos. Por lo tanto el futuro de estas tecnologías es innegable.

- **Resultados esperados a los que contribuye dentro de la línea de actuación. Plan de difusión y explotación.**

i) El resultado más importante que esperamos de este proyecto es metodológico y consiste en la implementación de tecnologías de scATAC-seq en la Unidad de Genómica de CABIMER y la puesta a la disposición de la comunidad científica y sanitaria andaluza y nacional de este conocimiento y de los datos obtenidos mediante su depósito en bases de datos. ii) La comparación de varias sc-OMICS de las mismas muestras ofrecerá información de su redundancia o complementariedad para futuras decisiones acerca de qué tecnologías usar sobre todo en la clínica. iii) Finalmente, los PDX son importantes modelos para la medicina personalizada (10). Comprender a nivel de célula única la heterogeneidad que adquieren estos modelos, y las interacciones entre las células humanas tumorales y las células murinas son resultados científicos de relevancia que esperamos obtener. Los resultados se presentarán en congresos y se publicarán en revistas especializadas del campo.

REFERENCIAS: 1. C. Tabula Sapiens *et al. Science* **376**, eabl4896 (2022). 2. M. Gerlinger *et al. N Engl J Med* **366**, 883-892 (2012). 3. S. F. Roerink *et al. Nature* **556**, 457-462 (2018). 4. M. Latil *et al. Cell stem cell* **20**, 191-204 e195 (2017). 5. A. N. Hata *et al. Nature medicine* **22**, 262-269 (2016). 6. S. M. Shaffer *et al. Nature* **546**, 431-435 (2017). 7. F. Wu *et al. Nature communications* **12**, 2540 (2021). 8. J. Nasser *et al. Nature* **593**, 238-243 (2021). 9. J. D. Buenrostro *et al. Nat Methods* **10**, 1213-1218 (2013). 10. E. R. Zanella, E. Grassi, L. Trusolino. *Nat Rev Clin Oncol* **19**, 719-732 (2022)

- **Implicaciones éticas y/o de seguridad.**

Las muestras utilizadas dispondrán de los permisos legales preceptivos informes del comité ético de experimentación animal de CABIMER, y consentimientos informados de las muestras humanas que procederán de Biobancos oficiales andaluces.

- **Presupuesto detallado del proyecto y justificación del mismo**

Personal (35.000€): Se solicita un titulado universitario durante un año de contrato, encargado de la optimización de los protocolos de preparación de muestras, disgregación de tejidos, generación de librerías de scATAC-seq para la secuenciación. *45% del presupuesto.*

Equipamiento/infraestructura (36.750€): adquisición de productos para experimentación, incluyendo Kits scATAC-seq 10X genomics, reactivos para disgregación de tejidos, reactivos de secuenciación y secuenciación de muestras. *48% del presupuesto.*

Otros: 5.250€ (gastos de difusión de resultados, costes de movilidad, informe del auditor, otros costes, etc.). *7% del presupuesto.*

Costes indirectos (15% costes directos): 11.550€. **PRESUPUESTO TOTAL: 88.550€.**

Andalucía-Biotec Salud LA3 (Nodo Sevilla)

Proyecto 7: IMPLEMENTACIÓN DE TECNOLOGÍAS DE PROTEÓMICA A NIVEL DE CÉLULA ÚNICA CON POTENCIAL APLICACIÓN EN MEDICINA PERSONALIZADA

IP: Alberto Pascual; Institución: Instituto de Biomedicina de Sevilla (IBiS)

- **Resumen de la propuesta:** La tecnología de análisis a nivel de célula única (sc) está revolucionando el estudio de los procesos biológicos al proporcionar un nivel de resolución sin igual a la hora de definir la complejidad celular de tejidos. Estas tecnologías, de rápida implantación en el ámbito académico, están permitiendo un entendimiento fundamental de los procesos del desarrollo, la homeostasis tisular, el envejecimiento de los organismos y las bases moleculares y celulares de múltiples procesos patológicos. El objetivo de la presente propuesta es desarrollar la potencialidad de la plataforma de “sc-proteomics” del Instituto de Biomedicina de Sevilla (IBiS) a partir de muestras humanas de diferente origen (biopsias, piezas quirúrgicas, xenoinjertos y avatares derivados de pacientes). A la finalización del proyecto, la plataforma tecnológica resultante dispondrá de las capacidades científico-técnicas para extender sus servicios a centros de investigación biomédica y hospitalarios que precisen de este tipo de tecnología en estrategias de medicina de precisión (MP).

- **Antecedentes del proyecto**

El proyecto de Biotecnología Aplicada a la Salud (BAS) tiene como foco el desarrollo de tecnologías biomédicas y bioinformáticas que contribuyan a la implantación efectiva de la Medicina de Precisión (MP), una nueva medicina centrada en el individuo y que hará que la diagnosis, tratamiento y prognosis de la enfermedad sea más eficiente en los sistemas de salud españoles. Este es un objetivo muy ambicioso que necesita de múltiples capacidades tecnológicas y de análisis, así como de su implantación dentro de los sistemas de salud de las CCAA.

La línea de actuación 3 está relacionada con el uso de tecnologías de sc. El origen y el desarrollo de la enfermedad suele estar asociado a tipos celulares específicos, con frecuencia poco abundantes, y a respuestas celulares heterogéneas (por ejemplo, en el contexto de tumores, a infección por patógenos o a sensibilidad a un fármaco). Es por ello que la MP necesita de una caracterización ómica (genómica, transcriptómica, epigenómica, proteómica y metabolómica) de las muestras biológicas derivadas de pacientes/controles y biomodelos con resolución de sc. Estas técnicas son capaces de identificar mutaciones somáticas asociadas a enfermedad, tipos celulares críticos para la enfermedad pero poco abundantes (e.g. cáncer stem cells, células tumorales circulantes), caracterizar poblaciones con diferente sensibilidad a

patógenos, composición de microbiota o drogas, marcadores específicos (de enfermedad, de resistencia, etc.), así como monitorizar la evolución de una patología u optimizar tratamientos. En particular, las tecnologías de sc proteomics están actualmente en desarrollo y constituyen una frontera de conocimiento técnico, debido a la complejidad de detectar los proteomas de célula única con la adecuada sensibilidad. El IBiS ha desarrollado una plataforma de sc-omics para la medicina de precisión del ISCIII, dotado con la mejor tecnología disponible en este momento para realizar estas técnicas. Sin embargo, el uso de estas técnicas y, en particular, de la sc proteomics, está actualmente lejos de la realidad clínica, a pesar de su gran potencial para definir nuevos biomarcadores a nivel proteico. En este contexto, este subproyecto busca aunar las fortalezas de BAS para el desarrollo técnico de esta tecnología con muestras clínicas.

- **Objetivos del proyecto**

Como prueba de concepto, este subproyecto pretende la implantación de tecnologías de sc-proteomics en un entorno de muestras clínicas/investigación clínica. En concreto y dado el estado inicial de estas tecnologías, se pretende conseguir los siguientes objetivos:

1. Desarrollo de un protocolo reproducible de extracción e identificación del proteoma de sc en muestras derivadas del entorno clínico.
2. Coordinación entre los diferentes subproyectos de la LA3 y dentro de la BAS.
3. Identificación de potenciales clientes de la tecnología a desarrollar en el entorno clínico español y establecimiento de alianzas.

Como tipo de muestra inicial de esta prueba de concepto se han escogido los cultivos ortotópicos de tumores humanos de Sarcoma de Ewing en ratones (PDX) por diferentes motivos:

1. Existe una colección de estos modelos en el IBiS que nos va a permitir disponer de muestra fresca necesaria para esta aproximación y otras de la LA3 (ver debajo).
2. Estas muestras provienen de estudios realizados por el equipo de investigación del Dr. Enrique de Álava, integrante del equipo investigador de este subproyecto, cuentan con una excelente trazabilidad clínica sobre el estado de la patología, desarrollo y evolución del tumor y respuesta a fármacos, haciendo de esta una colección ideal para responder a preguntas clínicas sobre biomarcadores y heterogeneidad tumoral en respuesta a fármacos.
3. Los PDX alinean perfectamente con el desarrollo en otras LA de este proyecto para la generación de biomodelos (LA2) y se dispone de una clínica bien documentada en Andalucía para poder desarrollar estudios bioinformáticos a posteriori de los resultados de sc y transcriptómica espacial (LA3) en la LA1.

Para el desarrollo de estos objetivos, el proyecto cuenta con el siguiente personal involucrado de forma activa: como IP del proyecto, el Dr. Alberto Pascual, responsable de sus infraestructuras técnicas y responsable científico del servicio de proteómica del IBiS, y como componentes del equipo de investigación, el Dr. Enrique de Álava, Jefe de Servicio de Anatomía Patológica del HUVR, investigador principal del IBiS y coordinador del nodo de Biobanco Andaluz del HUVR, del programa de Oncohematología y Genética del IBiS y de la estrategia de MP de la Junta de Andalucía, y la Dra. Marisa Mateos Martín, responsable técnico del servicio de proteómica del IBiS.

Además contará con la colaboración inicial de los siguientes grupos de investigación que han mostrado un interés inicial por este proyecto, aunque el mismo está abierto a nuevas propuestas: Grupo de Investigación en Neuropsicofarmacología y Psicobiología, Dras. Esther Berrocoso, María Hidalgo Figueroa e Irene Suárez, Universidad de Cádiz. Contribuirá con muestras de avatares de ratón conteniendo células humanas de pacientes con trastornos psiquiátricos. Grupo Terapia Regenerativa Cardiovascular y Proteómica Aplicada, Dra. M. Carmen Durán Ruiz, Universidad de Cádiz. Especialista en proteómica que apoyará el trabajo y análisis desde el IBiS. Grupo Exercise Physiology, Nutrition, and Health Research, Dr. Jesús Gustavo Ponce González, Universidad de Cádiz. Especialista en tejidos humanos con gran colección de muestras que podrán ser analizadas mediante sc-proteomics.

- **Contribución del proyecto a la línea de actuación.**

El proyecto desarrolla la vertiente sc-proteomics descrita en la línea de actuación inicial y se complementa y coordina con otros proyectos de LA3, donde se realizarán estudios de sc-RNAseq y sc-ATACseq con las mismas muestras, para obtener resultados conjuntos.

El proyecto se compromete al desarrollo de las acciones conjuntas que emanen del Proyecto de Dirección Científica, tanto en Andalucía como a las actuaciones conjuntas con otras CCAA.

- **Metodología y plan de trabajo. Cronograma.**

1. Desarrollo de un protocolo reproducible de extracción e identificación del proteoma de célula única en muestras derivadas del entorno clínico.
 - 1.a. Se abordará la disgregación y separación de células individuales por técnicas mecánicas de tejido fresco proveniente de PDX de Sarcoma de Ewing, disponibles en el IBiS.
 - 1.b. Separación mediante el equipo de separación celular de Cellenion y deposición mediante técnicas acústicas en pocillos individuales.
 - 1.c. Dispensación mediante robot con capacidad de manejo de nanolitros de soluciones de lisis y tripsinización de proteínas a los pocillos de células individuales.
 - 1.d. Acople de los pocillos individuales al equipo de cromatografía líquida y espectrometría de masas de alta resolución (SCP tims TOF, Bruker).
 - 1.e. Obtención de espectros de célula única.
 - 1.f. Análisis bioinformático en el IBiS y el CABD (subproyecto coordinado por el Dr. Javier López Ríos) para la identificación del proteoma de célula única, la clusterización en células de ratón o humanas, su reparto en tipos celulares.
2. Coordinación entre los diferentes subproyectos de la LA3 y dentro de la BAS.

2.a. El IBiS coordinará el reparto de muestras para los subproyectos de la LA3 de scRNAseq (Dr. J López-Ríos, CABD) y ATASseq (Dr. JC Reyes, CABIMER).

2.b. Se establecerán circuitos de análisis de los resultados obtenidos centrados en el nodo de bioinformática y la coordinación con la proteómica espacial (Dra. F Matesanz, IPBLN) dentro del LA3.

2.c. Se establecerán circuitos de información con las líneas LA1 y LA2 para el análisis e integración de los resultados.

3. Identificación de potenciales usuarios de la tecnología a desarrollar en el entorno clínico español y establecimiento de alianzas.

3.a. Se cuenta ya con grupos colaboradores que pueden aportar muestras de origen clínico de diferentes patologías y órganos, con los que se podrá desarrollar, en función de la disponibilidad presupuestaria nuevos análisis para la optimización de los protocolos de sc proteomics.

3.b. Se extenderá la búsqueda a otros grupos de Andalucía y de otras CCAA.

Cronograma

Objetivos	Semestre 1	Semestre 2	Semestre 3	Semestre 4
Objetivo 1.- SC proteomics	X	X	X	
Objetivo 2.- Coordinación interna BAS			X	X
Objetivo 3.- Coordinación externa BAS				X

- **Grado de innovación de la propuesta y justificación del potencial impacto científico y/o tecnológico del proyecto.**

La propuesta pretende implementar tecnologías absolutamente novedosas, aún muy poco desarrolladas, en nuestro ámbito regional. El futuro de estas tecnologías en el campo de la biomedicina en estudios básicos se está poniendo de manifiesto con una enorme cantidad de publicaciones de altísimo nivel en los últimos años. En el campo clínico y sobre todo en oncología la identificación de pequeñísimas poblaciones con características metastáticas o de resistencia a determinados quimioterapias podría condicionar completamente los tratamientos. Por lo tanto, el futuro de estas tecnologías es innegable.

- **Resultados esperados a los que contribuye dentro de la línea de actuación. Plan de difusión y explotación.**

LA3 resultará en el desarrollo de protocolos experimental y bioinformáticos estandarizados, que serán aplicados a problemas identificados de forma coordinada dentro del consorcio BAS en forma de Proyectos Prueba de Concepto específicos que validen las diferentes aproximaciones. Desde un punto de vista organizativo, el desarrollo de estas actividades permitirá la coordinación, dentro de estos proyectos específicos, de las diferentes capacidades tecnológicas que se están implantando a gran velocidad en los centros de investigación de Andalucía. En particular, la generación de una unidad coordinada de análisis de "single cell omics" permitirá el desarrollo dinámico de protocolos de análisis y la extensión de su uso a un número creciente de centros de investigación biomédica/centros hospitalarios que implementen o demanden técnicas ómicas sc. Los dos objetivos últimos son, por un lado, ofrecer una caracterización y análisis rápidos de muestras de pacientes que ayuden en la toma de decisiones clínicas en tiempo real y un procedimiento de iteración clínica/análisis que posibilite monitorizar la respuesta a esas decisiones, y por otro integrar el conocimiento generado para el descubrimiento y desarrollo de nuevas soluciones terapéuticas y mejorar su traslación a la clínica.

- **Implicaciones éticas y/o de seguridad.**

Las muestras utilizadas dispondrán de los permisos legales preceptivos informes del comité ético de experimentación animal del HUVR/IBiS, y consentimientos informados de las muestras humanas que procederán de Biobancos oficiales andaluces.

- **Presupuesto detallado del proyecto y justificación del mismo**

Personal (35.000€): Se solicita un titulado universitario durante un año de contrato, encargado de la optimización de los protocolos de preparación de muestras, disgregación de tejidos, y que fortalecerá la unidad de proteómica del IBiS (<https://www.ibis-sevilla.es/servicios/proteomica.aspx#page=page-1>), que actualmente cuenta con el IP de este proyecto como su responsable científico y con la Dra. Marisa Mateos como responsable técnico. La complejidad de las técnicas a emplear y la alta carga de trabajo esperada con la implantación de estas técnicas justifican la contratación de este personal. *45% del presupuesto.*

Equipamiento/infraestructura (41.000€): adquisición de productos y fungibles para experimentación, incluyendo reactivos y Kits para sc-proteomics. *53% del presupuesto.*

Otros: 1.000€ (gastos de difusión de resultados, informe del auditor, etc.). *1% del presupuesto.*

Costes indirectos (15% costes directos): 11.550€. **PRESUPUESTO TOTAL: 88.550€.**

Procedimientos implantados para evitar el daño significativo al medioambiente (DNSH) y para prevenir y corregir el fraude, la corrupción y el conflicto de interés.

El desarrollo del proyecto Andalucía-Biotec Salud en su conjunto, y de cada uno de los subproyectos, no prevé daños significativos al medio ambiente, y en consecuencia, se considera que cumple el principio DNSH por lo que respecta al objetivo en cuestión. Todos los procedimientos y protocolos usados implican la separación selectiva y tratamiento de residuos de investigación de acuerdo a la legislación vigente, así como la destrucción de material biológico sobrante. Los equipos solicitados en las distintas LA serán instalados tras la licitación del contrato administrativo de compra según dicta la legislación vigente y habiéndose ajustado plenamente la ejecución a los Pliegos de Cláusulas Administrativas y Prescripciones Técnicas del mismo, lo que garantiza la libre competencia y el uso correcto de fondos públicos. La contratación de personal se realizará de acuerdo con la legislación vigente y las ofertas serán publicitadas en la plataforma propia de Andalucía-Biotec Salud y del Plan Complementario a nivel nacional, así como en las plataformas de contratación abierta del CSIC y del resto de instituciones participantes, garantizando la concurrencia competitiva y la justa evaluación de méritos. En el contexto de la acción coordinada, se constituirá un panel de expertos que posibilite una revisión independiente de los datos y los resultados del proyecto para detectar cualquier posible fraude o sesgo. Todos los investigadores y personal involucrado en el proyecto deberán declarar cualquier posible conflicto de intereses. La evaluación de la convocatoria interna de proyectos de prueba de concepto en las distintas LA se realizará por un comité que evite el conflicto de interés y la corrupción.

En esta propuesta hemos considerado los siguientes **aspectos vinculados con la inclusión, la integración y la perspectiva de género**, aplicando las políticas en esta materia adoptadas e impulsadas por las instituciones estatales (ej. CSIC) y autonómicas (universidades y FPS) participantes en la propuesta:

- **Inclusión:** Se promoverá la inclusión de personas de diferentes orígenes, incluyendo a personas de diferentes grupos étnicos y socioeconómicos, y con discapacidades. Esto será llevado a cabo a través de la selección de un equipo de investigación diverso, el diseño de políticas de inclusión y el acceso a la tecnología para personas con discapacidad.
- **Integración:** Se fomentará la integración de diferentes disciplinas científicas, fomentando la colaboración entre expertos en biología, medicina, informática, estadística y otras áreas relacionadas. Esto permitirá una aproximación interdisciplinaria y la generación de nuevos conocimientos y soluciones.
- **Perspectiva de género:** Se impulsará la perspectiva de género en la investigación, considerando las diferencias de género en la salud y la enfermedad, y en el acceso a los cuidados de salud. Además, se tendrá en cuenta la participación equitativa de mujeres y hombres en el equipo de investigación, asegurando una representación equitativa de ambos géneros en la toma de decisiones y en la divulgación de resultados.