

Vigilancia epidemiológica del COVID-19 en aguas residuales

Investigadores responsables: Rafael Sanjuán y Pilar Domingo-Calap

Instituto de Biología Integrativa de Sistemas (Universitat de València-CSIC)

Paterna, 23 de noviembre de 2021

1. Resultados científico-técnicos

1.1. Introducción

La vigilancia epidemiológica basada en aguas residuales nos permite hacer un seguimiento de la circulación de patógenos o sustancias nocivas en la población. Las pruebas individualizadas, ya sean por PCR o pruebas rápidas de antígenos, son esenciales para el control de brotes, tal como se ha demostrado en el caso de la COVID-19. Sin embargo, en un escenario de pandemia como la actual, su uso generalizado en la población se ve a menudo limitado desde los puntos de vista técnico y económico. El seguimiento a través de las aguas residuales es un método alternativo sencillo, económico, rápido y no invasivo, que permite monitorizar las poblaciones y llevar a cabo un control de las epidemias desde una aproximación complementaria a las pruebas individualizadas. Estudios previos han demostrado que se puede detectar una persona infectada por SARS-CoV-2 por cada 1000 habitantes a través del análisis de aguas residuales, valor que está cerca de los valores de incidencia considerados como alarmantes. Sin embargo, son necesarios más estudios para determinar mejor los umbrales de detección de esta técnica.

Mediante el análisis de aguas residuales, se espera poder detectar nuevos brotes, así como estudiar la estacionalidad de las epidemias y otros parámetros poblacionales de interés. Además, tanto los pacientes sintomáticos como los individuos asintomáticos pueden contribuir a la presencia de virus en aguas, lo cual es de especial relevancia en el caso de la COVID-19. Es sabido que el SARS-CoV-2 puede ser excretado, tanto por saliva como por heces y orina, no sólo por individuos asintomáticos, sino también antes del inicio de los síntomas y hasta cinco semanas después, siendo por tanto el análisis de aguas residuales un método en principio adecuado para el seguimiento de este virus.

Uno de los objetivos a alcanzar es determinar la validez del análisis de aguas residuales como sistema de alerta temprana, el cual permitiría anticipar de forma sencilla y económica la presencia de niveles preocupantes de SARS-CoV-2 en la población. Para conseguir este objetivo, es necesario monitorizar las estaciones depuradoras de manera periódica (al menos una vez por semana) y siguiendo protocolos consistentes. No obstante, es necesario realizar estudios metodológicos para determinar qué sistemas recogida son los más eficaces.

En concreto, a la hora de hacer el seguimiento en aguas residuales, es importante llevar a cabo un muestreo representativo, teniendo en cuenta factores como la selección del punto de muestreo, el volumen de agua analizado, el horario y tipo de muestra, su almacenamiento hasta

su procesado, o incluso factores ambientales que pueden tener efecto como la temperatura, la radiación solar, la lluvia, o la contribución de aguas de regadío. También es importante implementar una metodología de detección molecular adecuada, mediante protocolos estandarizados, llevando a cabo los controles de recuperación viral e inhibición adecuados. Se deben establecer los límites de detección de las técnicas utilizadas, su validación mediante curvas de estandarización, y la presencia de falsos positivos y falsos negativos. Es importante remarcar que las aguas residuales son una matriz compleja, que varía entre plantas residuales y entre países y que puede dificultar en sí misma la detección del material genético viral. Asimismo, los datos publicados hasta la fecha indican que la señal observada por métodos moleculares en aguas residuales es débil y está cerca del umbral de sensibilidad de las técnicas, aunque es detectable. Por último, para poder utilizar los datos moleculares en aguas residuales como herramienta de vigilancia epidemiológica, es importante poder correlacionarlos con otros datos epidemiológicos con el objetivo de implementar modelos de predicción y seguimiento de casos.

En el I2SysBio, centro mixto de la Universitat de València y el CSIC, llevamos analizando las aguas residuales de las estaciones depuradoras de Pinedo 1, Pinedo 2 y Quart-Benàger desde el inicio de la pandemia. En una primera aproximación, realizamos análisis retrospectivos por PCR de muestras puntuales recogidas desde febrero hasta abril de 2020. Nuestros resultados permitieron determinar que el SARS-CoV-2 era detectable en aguas residuales días antes del primer caso declarado en València, lo que sugiere que el sistema de detección podría alertar de forma temprana (<https://www.sciencedirect.com/science/article/pii/S1438463920305678>).

1.2. Puesta a punto de la metodología

Desde mayo 2020, gracias al convenio entre la Generalitat Valenciana a través de la Conselleria de Agricultura, Desarrollo Rural, Emergencia Climática y Transición Ecológica y la Universitat de València, estamos realizando un seguimiento temporal a razón de tres muestras semanales en las estaciones mencionadas, las cuales recogen aguas de más de un millón de habitantes. Para ello, hemos puesto a punto la metodología de detección por PCR del material genético del coronavirus en aguas residuales. Además de estimar la carga viral en las aguas residuales, desde enero 2021, estamos realizando secuenciación de las muestras de aguas residuales con el fin de conocer las variantes de interés del SARS-CoV-2 en la ciudad de València. Con ello, pretendemos establecer los protocolos que nos permitan obtener una visión poblacional del avance de la pandemia en el área metropolitana de Valencia, mediante el sistema de vigilancia en las aguas residuales, ya que permite monitorizar las poblaciones a lo largo del tiempo siendo un método sencillo, no invasivo, económico.

En una primera aproximación se han puesto a punto las técnicas de concentración, extracción y detección del material genético. El objetivo es determinar la carga viral del SARS-CoV-2 en aguas residuales, una matriz compleja, de la forma más reproducible y eficaz posible. Además, se pretende utilizar un método que permita mantener infectivas las posibles partículas virales. Para ello, realizamos centrifugación a alta velocidad, ya que demostramos que era el mejor método tanto en detección por métodos moleculares como por ensayo en placa (detección de partículas virales infectivas, resultados 2020). La detección por métodos moleculares del SARS-CoV-2 se realiza utilizando la técnica de RT-qPCR con los cebadores recomendados por el CDC (N1 y N2). En el caso de la secuenciación, en una primera instancia se realizó secuenciación masiva de las

muestras mediante la técnica Illumina MiSeq. Sin embargo, las muestras de aguas residuales de baja carga viral son muy difíciles de secuenciar mediante este método. Además, es complicado recuperar genomas completos ya que, en muchos casos, el RNA viral está degradado y no amplifica correctamente. Para solventar este problema, hemos secuenciado las muestras mediante secuenciación Sanger. Con este método, obtenemos secuencias de calidad de regiones de interés del SARS-CoV-2 que son informativas de la variante de interés. En primer lugar, obtenemos secuencias consenso para determinar la variante mayoritaria en las aguas residuales, y procedemos posteriormente con clonación de las bandas de RT-PCR para y secuenciamos un número elevado de clones para poder conocer las variantes minoritarias que circulan por la ciudad de València.

1.3. Resultados

Se han realizado los muestreos a razón de dos o tres muestras por semana desde el mes de mayo 2020 en las tres EDARs principales del área metropolitana de Valencia: Pinedo 1, Pinedo 2 y Quart-Benàger. En primer lugar, se ha comparado la amplificación por RT-qPCR de las dos regiones diana, N1 y N2, observándose una buena correlación entre ambas, y se ha determinado el límite de detección de la técnica implementada (**Figura 1**).

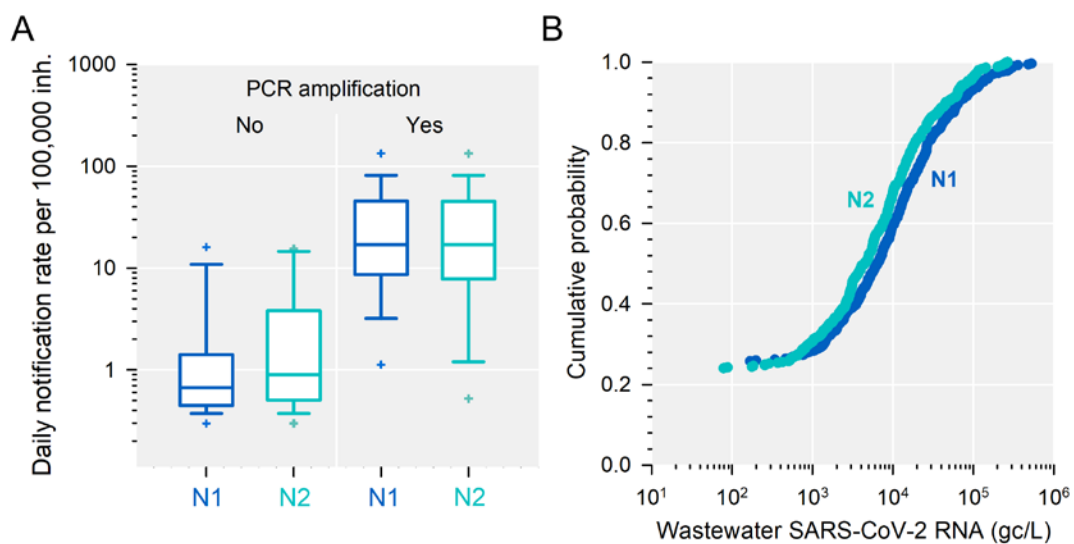


Figura 1. Límite de detección del SARS-CoV-2 en aguas residuales. **A.** Diagrama de cajas de las tasas de notificación diaria por 100.000 habitantes para las muestras de aguas residuales que dieron un producto de RT-qPCR, frente a las que no se amplificaron. Los datos se muestran por separado para las regiones del genoma viral N1 y N2. Los recuadros muestran los percentiles 25, 50 (mediana) y 75. Las barras de error indican los percentiles 10 y 90, y las cruces muestran los percentiles 5 y 95. **B.** Distribución de probabilidad acumulada de la concentración de ARN viral estimada para las regiones genómicas N1 y N2. Según esta distribución, el límite de detección de la RT-qPCR fue de aproximadamente 100 gc/L.

Hemos llevado a cabo análisis comparativos entre la carga viral en aguas compuestas y puntuales, entre las regiones diana N1 y N2, y entre diferentes colectores. Los resultados se muestran en la **Figura 2**. Cabe destacar que las muestras compuestas presentan menor variación que las muestras puntuales y por tanto proporcionan una información más fiable acerca de la carga viral en aguas. Para ello, se ha comparado la toma de muestras puntuales con muestras compuestas de 4h, todas de primera hora de la mañana. Las muestras se analizan el mismo día o dentro de los 3 días siguientes a la toma de las muestras, guardadas a -20°C para su correcta preservación. Los resultados de unidades genómicas por litro obtenidos por RT-qPCR, se comparan con los datos epidemiológicos obtenidos a partir de la Dirección General de Salud Pública. Hemos centrado nuestro análisis estadístico de los datos en las muestras recogidas en Pinedo 1 y 2, puesto que abarcan gran parte de la ciudad de València y nos permiten una correlación con los casos declarados más fiable que en el caso de Quart-Benàger. Además, en esta última EDAR, han tenido problemas en la toma de muestras en períodos largos lo que dificulta su seguimiento temporal.

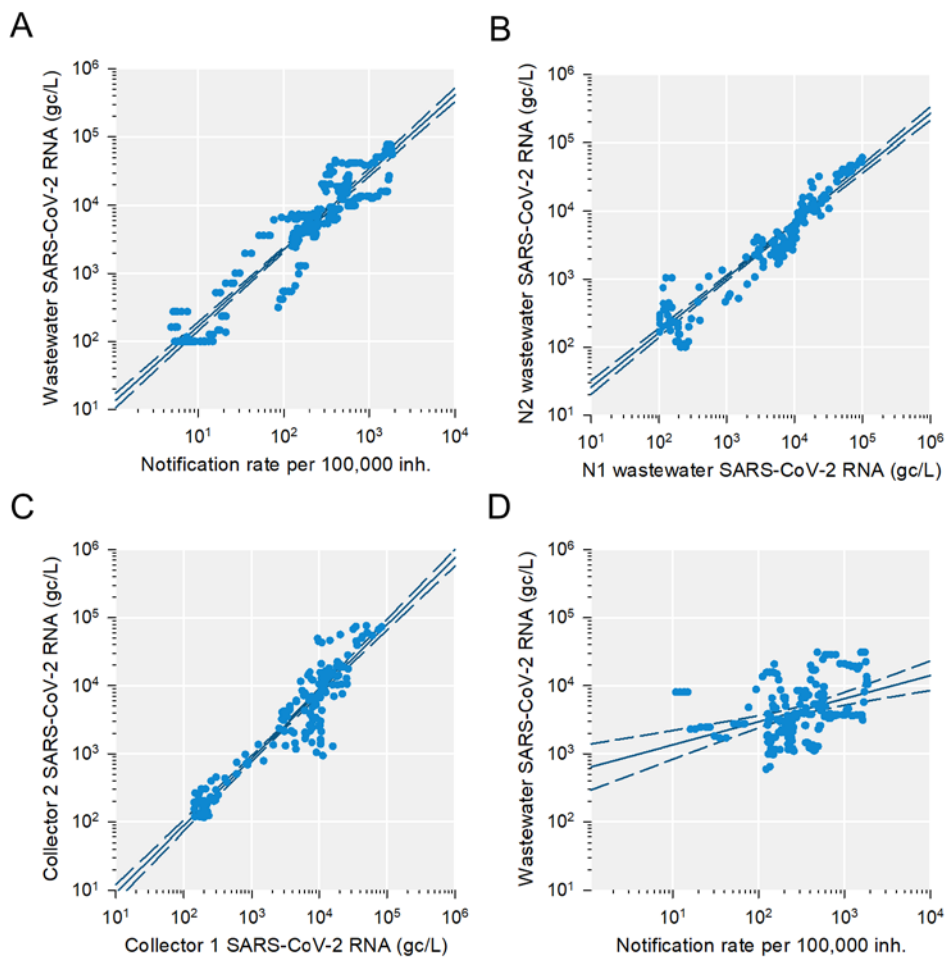


Figura 2. Influencia de la estrategia de muestreo de aguas residuales en la cuantificación del SARS-CoV-2. Se calcularon las concentraciones de ARN viral en las aguas residuales y las tasas acumuladas de

notificación de casos por cada 100.000 habitantes para períodos de 14 días utilizando ventanas deslizantes desplazadas de un día. Se muestran las líneas de regresión lineal log-log por mínimos cuadrados y los intervalos de confianza del 99%. **A.** Tasa de notificación frente a la mediana de la concentración de ARN viral en las aguas residuales compuestas. **B.** Concentraciones medias de ARN viral logarítmico estimadas utilizando la región del genoma del SARS-CoV-2 N1 frente a la N2. **C.** Concentraciones medias de log ARN viral estimadas a partir del colector 1 de la EDAR de Pinedo frente al colector 2. **D.** Tasa de notificación frente a la concentración media de log ARN viral en las aguas residuales utilizando muestras individuales en lugar de muestras compuestas.

Existe correlación entre la señal obtenida por PCR en aguas de Pinedo 1 y 2 y el número de casos COVID declarados en el mismo día. No obstante, existe fuerte variabilidad en los datos. Las razones de esa variabilidad pueden ser varias e incluyen factores como la baja señal obtenida en las PCRs de aguas, oscilaciones en el caudal del agua, la temperatura, etc., así como la variación diaria en el número declarado de casos (por ejemplo, con valores más bajos los fines de semana debido al menor número de pruebas realizadas). Para mejorar dicha correlación, hemos realizado análisis por períodos de una semana y de 14 días, agregando los datos de ese periodo para Pinedo 1 y Pinedo 2. De ese modo, se alcanza una correlación muy buena entre la señal obtenida en aguas y el número de casos COVID activos declarados (**Figuras 3 y 4**). Una vez establecido el método de obtención de la muestra, realización de la qPCR y análisis de datos, es posible establecer un paralelismo temporal entre el número de casos declarados y la señal obtenida en aguas residuales. Dicha relación se muestra en la Figura 4, donde puede observarse que el análisis de las aguas recapitula fielmente la curva epidémica del virus en la ciudad de València.

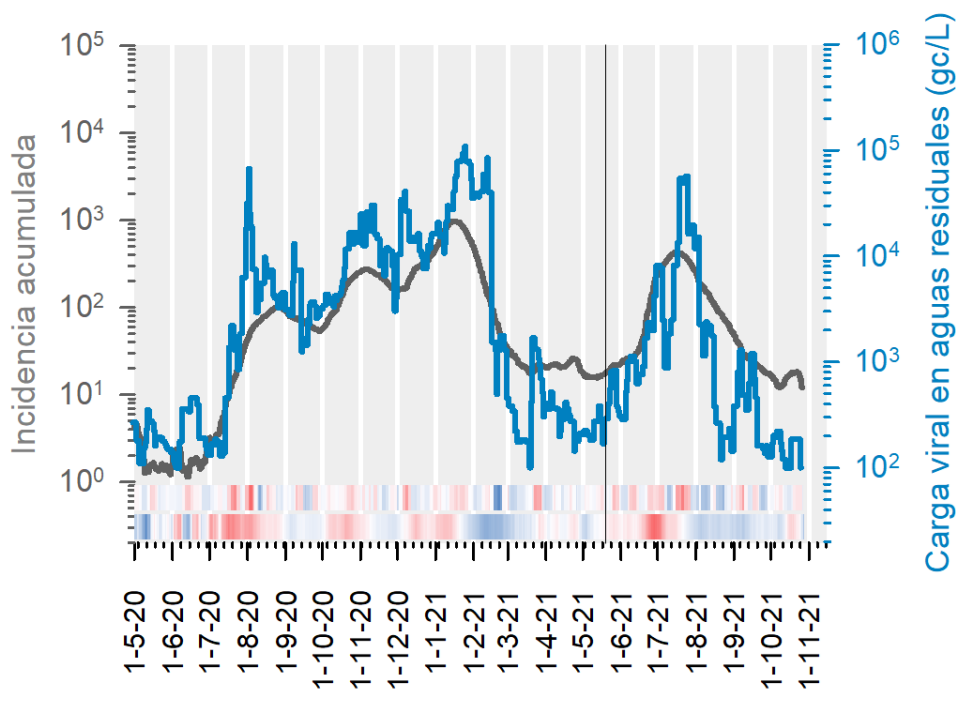


Figura 3. Correlación entre los casos activos declarados en Valencia y la señal obtenida por el análisis de aguas residuales de Pinedo 1 y 2 en períodos ventana de 7 días, desde mayo 2020 hasta noviembre 2021.

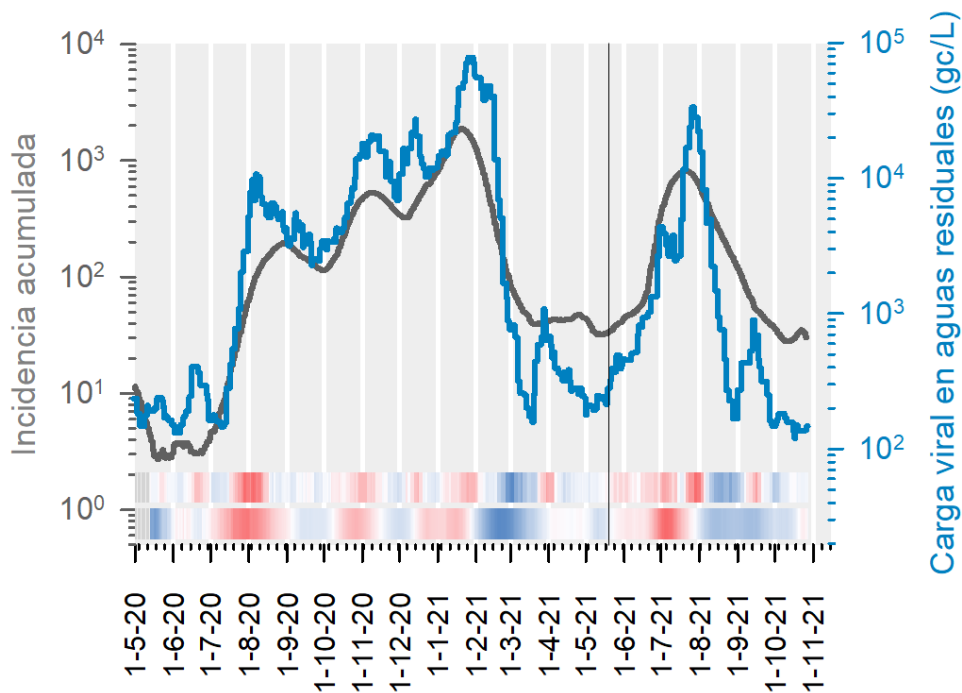


Figura 4. Correlación entre los casos activos declarados en Valencia y la señal obtenida por el análisis de aguas residuales de Pinedo 1 y 2 en períodos ventana de 14 días, desde mayo 2020 hasta noviembre 2021.

Por otra parte, hemos analizado también la correlación entre la carga viral en aguas residuales de una EDAR pequeña, Quart-Benàger, con el número de casos. Los resultados pueden observarse en la **Figura 5**, donde se detecta una mayor variabilidad y dispersión de los datos, así como períodos con ausencia de muestreo.

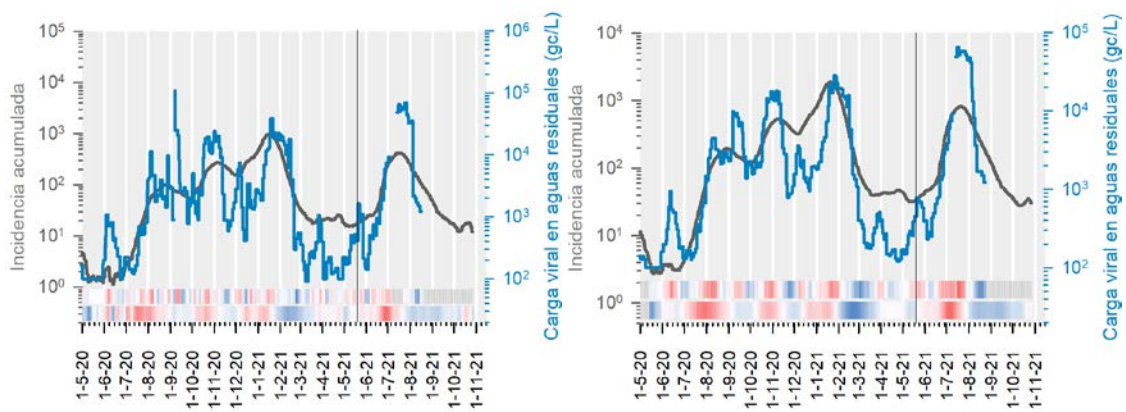


Figura 5. Correlación entre los casos activos declarados y la señal obtenida por el análisis de aguas residuales de Quart-Benàger en períodos ventana de 7 días (izquierda) y 14 días (derecha), desde mayo 2020 hasta noviembre 2021.

1.4. Conclusiones

Pese al previsible control de la epidemia COVID tras la implementación de campañas de vacunación masiva, a fecha de hoy el virus todavía no se ha erradicado y nuevas oleadas epidémicas en la población no inmunizada o por la emergencia de nuevas variantes podrían dar lugar a nuevos picos que sobrecarguen el sistema de salud, al igual que ocurre con la gripe y otros virus. Es también altamente probable que el esfuerzo diagnóstico por PCR u otros métodos decaiga con el tiempo, debido a su elevado coste material y logístico y a la disminución de la alarma social. No obstante, es importante mantener medios de seguimiento epidemiológico de COVID a medio-largo plazo que sean coste-efectivos. Hay que tener en cuenta que la sintomatología de este virus es variable y a menudo muy poco específica, haciendo difícil su diferenciación de otras enfermedades respiratorias comunes. En este contexto, disponer de un método de seguimiento basado en aguas residuales es de gran utilidad, tal y como recomienda la Comisión Europea. Los resultados aquí presentados avalan la posibilidad de implementar esta metodología en la ciudad de València.

Existen diversas razones pues para dar continuidad a estos estudios. En primer lugar, es necesario seguir mejorando la técnica. En esta dirección, sería beneficioso disponer de métodos más eficaces de concentración de las muestras que nos permitan obtener una señal de PCR más fuerte y por tanto más alejada el umbral de detección. En segundo lugar, sería conveniente mantener en el tiempo el estudio de correlación entre los casos declarados y la señal en aguas, especialmente en el momento en el que los casos son muy bajos, pues no hemos podido estudiar qué sucede con la señal medida en aguas en este escenario. En tercer lugar, sería interesante extender la metodología a otros virus respiratorios cuyo seguimiento poblacional pueda resultar de interés. Por último, como se ha señalado, resultaría muy útil mantener en el tiempo una plataforma de seguimiento epidemiológico del COVID (y potencialmente de otros virus) basado en los métodos descritos, una vez éstos hayan quedado establecidos.

2. Justificación económica

Gastos de personal

- Gastos de contratación de María Dolores Arocas Castillo (Técnico Especialista de Laboratorio): 31 392.65 €

Justificación: recolección de muestras en estaciones depuradoras, puesta a punto de la metodología y realización de todas las PCR.

- Dedicación horaria de Rafael Sanjuán Verdeguer (Profesor Titular, Departamento de Genética de la Universidad de València): 2976.70 €

Justificación: investigador principal, supervisión de los ensayos, análisis e interpretación de resultados y actividades de seguimiento del Convenio y difusión. Estas tareas han supuesto una dedicación horaria de 85 h, con un coste/hora calculado en 35.02 €.

- Dedicación horaria de Pilar Domingo Calap (Investigadora del Programa Ramón y Cajal, Instituto de Biología Integrativa de Sistemas de la Universidad de València): 5274.34 €

Justificación: co-investigadora principal, supervisión de los ensayos, análisis e interpretación de resultados y actividades de seguimiento del Convenio y difusión. Estas

tareas han supuesto una dedicación horaria de 181 h, con un coste/hora calculado en 29.14 €.

Gastos de equipamiento:

Material fungible:

- Reactivos de biología molecular: 20 400.68 €

Justificación: material necesario para la realización de las qPCR.

Otros gastos:

- Gastos de desplazamiento y envío de muestras 678.30 €
- Gastos de formación y adquisición de reactivos para la puesta en marcha de un equipo de qPCR: 220.00 €.

Justificación: gastos de combustible fruto de desplazamientos efectuados por María Dolores Arocas Castillo con su vehículo particular para la recogida de aguas residuales en las estaciones de Pinedo y Quart-Benàger, así como otros gastos de envío. Puesta en marcha de un equipo de qPCR dedicado principalmente al presente convenio.

Costes indirectos:

- Gastos asociados al mantenimiento de los laboratorios y las infraestructuras necesarias para la ejecución del Convenio (13% del total): 9100 €.

TOTAL: 70 042.67 €

Rafael Sanjuán Verdeguer

Investigador responsable

Pilar Domingo Calap

Investigadora responsable