Monitoring emergence of SARS-CoV-2 B.1.1.7 Variant through the Spanish National SARS-CoV-2 Wastewater Surveillance System (VATar COVID-19)

Journal:	*Environmental Science & Technology*
Manuscript ID	es-2021-03589a.R2
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Carcereny, Albert; Universitat de Barcelona, Genetics, Microbiology and Statistics; Research Institute of Nutrition and Food Safety (INSA), University of Barcelona, Martínez-Velázquez, Adan; Universitat de Barcelona, Genetics, Microbiology and Statistics; Research Institute of Nutrition and Food Safety (INSA), University of Barcelona Bosch, Albert; Universitat de Barcelona, Department of Genetics, Microbiology and Statistics; Research Institute of Nutrition and Food Safety (INSA), University of Barcelona Allende, Ana; CEBAS CSIC, Department of Food Science & Technology Truchado, Pilar; CEBAS CSIC Cascales, Jenifer; CEBAS CSIC Romalde, Jesús; Universidade de Santiago de Compostela, Departamento de Microbiología Lois, Marta; Universidade de Santiago de Compostela Polo, David; Universidade de Santiago de Compostela Sánchez, Gloria; IATA-CSIC, Biotechnology Pérez-Cataluña, Alba; IATA-CSIC Díaz-Reolid, Azahara; IATA-CSIC Antón, Andrés; Vall d'Hebron Institut de Recerca (VHIR) Gregori, Josep; Vall d'Hebron Institut de Recerca (VHIR); Centro de Investigación Biomédica en Red Enfermedades Hepáticas y Digestivas Garcia-Cehic, Damir; Vall d'Hebron Institut de Recerca (VHIR); Centro de Investigación Biomédica en Red Enfermedades Hepáticas y Digestivas Quer, Josep; Vall d'Hebron Institut de Recerca (VHIR); Centro de Investigación Biomédica en Red Enfermedades Hepáticas y Digestivas Palau, Margarita; Government of Spain Ministry of Health Social Policy and Equality González Ruano, Cristina; MITECO Pintó, Rosa; Universitat de Barcelona, Department of Genetics, Microbiology and Statistics; Research Institute of Nutrition and Food Safety (INSA), University of Barcelona Guix, Susana; Universitat de Barcelona, Genetics, Microbiology and Statistics; Research Institute of Nutrition and Food Safety (INSA), University of Barcelona

Monitoring emergence of SARS-CoV-2 B.1.1.7 Variant through the Spanish National SARS-CoV-2 Wastewater Surveillance System (VATar COVID-19)

Albert Carcereny[1,2†], Adán Martínez-Velázquez[1,2†], Albert Bosch[1,2], Ana Allende[3], Pilar Truchado[3], Jenifer Cascales[3], Jesús L Romalde[4], Marta Lois[4], David Polo[4], Gloria Sánchez[5], Alba Pérez-Cataluña[5], Azahara Díaz-Reolid[5], Andrés Antón[6], Josep Gregori[7,8], Damir Garcia-Cehic[7,8], Josep Quer[7,8], Margarita Palau[9], Cristina González Ruano[10], Rosa M Pintó[1,2#*] and Susana Guix[1,2#*]

[1] Enteric Virus laboratory, Department of Genetics, Microbiology and Statistics, Section of Microbiology, Virology and Biotechnology, School of Biology, University of Barcelona, Barcelona, 08028, Spain.

[2] Research Institute of Nutrition and Food Safety (INSA), University of Barcelona, Santa Coloma de Gramenet, 08921, Spain.

[3] Research Group on Microbiology and Quality of Fruit and Vegetables, CEBAS-CSIC, Murcia, 30100, Spain.

[4] Department of Microbiology and Parasitology, CIBUS-Faculty of Biology & Institute CRETUS, Universidade de Santiago de Compostela, Santiago de Compostela, 15782, Spain.

[5] Department of Preservation and Food Safety Technologies, Institute of Agrochemistry and Food Technology, IATA-CSIC, Paterna, 46980, Spain.

[6] Microbiology Department, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Barcelona Hospital Campus, Barcelona, 08031, Spain.

[7] Liver Unit, Liver Diseases - Viral Hepatitis, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Barcelona Hospital Campus, Barcelona, 08035, Spain.

[8] Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Instituto de Salud Carlos III, Madrid, 28029, Spain.

[9] General Directorate of Public Health, Ministry of Health, Madrid, 28014, Spain

[10] Subdirección General de Protección de las Aguas y Gestión de Riesgos, Ministerio para la Transición Ecológica y el Reto Demográfico, Madrid, 28003, Spain

[†] Equally contributed

Address correspondence to Susana Guix (susanaguix@ub.edu) and Rosa Pintó (rpinto@ub.edu)

* Co-corresponding authors

ABSTRACT

Since its first identification in the United Kingdom in late 2020, the highly transmissible B.1.1.7 variant of SARS-CoV-2, become dominant in several countries raising great concern. We developed a duplex real-time RT-qPCR assay to detect, discriminate and quantitate SARS-CoV-2 variants containing one of its mutation signatures, the ΔHV69/70 deletion, and used it to trace the community circulation of the B.1.1.7 variant in Spain through the Spanish National SARS-CoV-2 Wastewater Surveillance System (VATar COVID-19). B.1.1.7 variant was detected earlier than clinical epidemiological reporting by the local authorities, first in the Southern city of Málaga (Andalucía) in week 20_52 (year_week), and multiple introductions during Christmas holidays were inferred in different parts of the country. Wastewater-based B.1.1.7 tracking showed a good correlation with clinical data and provided information at the local level. Data from WWTPs which reached B.1.1.7 prevalences higher than 90% for ≥ 2 consecutive weeks showed that 8.1±1.8 weeks were required for B.1.1.7 to become dominant. The study highlights the applicability of RT-qPCR-based strategies to track specific mutations of variants of concern (VOCs) as soon as they are identified by clinical sequencing, and its integration into existing wastewater surveillance programs,

as a cost-effective approach to complement clinical testing during the COVID-19 pandemic.

Keywords: SARS-CoV-2, COVID-19, B.1.1.7 Variant, Wastewater-based Epidemiology (WBE), RT-qPCR, NGS

Synopsis: The study includes the development of a RT-qPCR assay to discriminate SARS-CoV-2 B.1.1.7 variant and its implementation on the Spanish sewage surveillance system as a data source to complement clinical pandemic assessment.

INTRODUCTION

Environmental surveillance of specimens contaminated by human faeces is used to monitor enteric virus disease transmission in the population, and several countries have implemented SARS-CoV-2 wastewater monitoring networks to inform decision making during the COVID-19 pandemic [1–3]. In Spain, a nation-wide COVID-19 wastewater surveillance project (VATar COVID-19) was launched in June 2020 (https://www.miteco.gob.es/es/agua/temas/concesiones-y-autorizaciones/vertidos-de-aguas-residuales/alerta-temprana-covid19/default.aspx), and has weekly monitored SARS-CoV-2 levels in untreated wastewater from initially 32 wastewater treatment plants (WWTPs) since then. On March 2021, the European Commission adopted a recommendation on a common approach to establish and make greater use of systematic wastewater surveillance of SARS-CoV-2 as a new source of independent information on the spread of the virus and its variants in the European Union[4]. In situations with low or absent SARS-CoV-2 circulation in the community, wastewater surveillance has proven a useful tool as an early warning system [5–9], and several studies have also tried to infer disease incidence in a community, independent of diagnostic testing availability based on SARS-CoV-2 wastewater concentrations, with considerable uncertainties [10–12].

Despite titanic efforts based on confinement measures and mass-vaccination programs, the emergence of novel variants of concern (VOCs), mainly B.1.1.7, B.1.351 B.1.1.28.1 and recently B1.617.2, so far, suggests that continued surveillance is required to control the COVID-19 pandemic in the long run. Since January 2021, countries within and outside Europe have observed a substantial increase in the number and proportion of SARS-CoV-2 cases of the B.1.1.7 variant, first reported in the United Kingdom [13,14].

Since B.1.1.7 variant has been shown to be more transmissible than the previously predominant circulating variants and infections may be more severe [15], countries where the variant has spread and become dominant are concerned on whether the occurrence of the variant will result in increases in total COVID-19 incidence, hospitalizations, and excess mortality due to overstretched health systems.

The emergence of SARS-CoV-2 variants that may increase transmissibility and/or immune escape, points to an imperative need for the implementation of targeted surveillance methods. While sequencing should be the gold standard for variant characterization, cost-effective molecular assays, which could be rapidly established and scaled up may offer several advantages and provide valuable quantitative information without delay.

This study included the development and validation of a one-tube duplex quantitative real-time RT-PCR (RT-qPCR) assay to detect, discriminate and quantitate SARS-CoV-2 variants containing the ΔHV69/70 deletion from variants lacking it, using allelic discrimination probes. Confirmatory sequencing of a subset of samples was performed to be able to ascertain the validity of these assays to trace the community circulation of the B.1.1.7 variant. The RT-qPCR-based assay improved the current variant tracking capability and could be easily implemented for monitoring the emergence of ΔHV69/70 containing SARS-CoV-2 variants (mainly B.1.1.7) in Spain through the nation-wide wastewater surveillance network.

METHODS

Wastewater sampling. Influent water grab samples were weekly collected from 32 WWTPs (one weekly sample per site) located in 15 different Autonomous Communities in Spain, from middle December 2020 (week 20_51, year_week number) to end of March 2021 (week 21_13) (last week of 2020 was not sampled). All samples were transported on ice to one of the 4 participating laboratories of analysis (A, B, C, D), stored at 4°C and processed within 1-2 days upon arrival. The time between sample collection and arrival to the laboratory ranged between 3-24 hours.

Sample concentration, nucleic acid extraction and process control. 200 ml wastewater samples were concentrated by aluminum hydroxide adsorption-precipitation method, as previously described [7,16]. Briefly, samples were adjusted to pH 6.0, a 1:100 v:v of 0.9 N $AlCl_3$ solution was added, and samples were gently mixed for 15 min at room temperature. Precipitate was collected by centrifugation at 1,700×g for 20 min and pellet was resuspended in 10 mL of 3% beef extract (pH 7.4). After a 10- min shaking 150 rpm, samples were centrifuged at 1,900×g for 30 min, and concentrates were resuspended in 1-2 ml of phosphate buffered saline (PBS). All samples were spiked with a known amount of an animal coronavirus used as a process control virus. Animal coronaviruses differed between participant laboratories and included the attenuated PUR46-MAD strain of Transmissible Gastroenteritis Enteric Virus (TGEV) [17]; Porcine Epidemic Diarrhea Virus (PEDV) strain CV777 (kindly provided by Prof. A. Carvajal from University of Leon) and Murine Hepatitis Virus (MHV) strain ATCC VR-764 [18]. Depending on each laboratory, between 10-100 µl of the animal coronavirus stock were added to 200 ml of sample to obtain final concentrations of $2.5x10^4$-$2.5x10^5$ copies/ml (PEDV and MHV) or $6.9x10^3$ TCID50/ml (TGEV). Nucleic acid extraction from

concentrates was performed from 300 µl using the Maxwell® RSC PureFood GMO and Authentication Kit (Promega Corporation, Madison, US), or 150 µl using the NucleoSpin RNA virus kit (Macherey-Nagel GmbH & Co., Düren, Germany), following the manufacturer's instructions. Each extraction included a negative control and a process virus control used to estimate the virus recovery efficiency. RT-qPCR for process control viruses were performed as previously described [19–21]. Parallel to ISO 15216-1:2017 [22] for the determination of norovirus and hepatitis A virus in the food chain, samples with a virus recovery ≥ 1% were considered acceptable.

SARS-CoV-2 RT-qPCR assays. The N1 assay targeting a fragment of the nucleocapsid gene, as published by US CDC (US-CDC 2020), was used to quantify SARS-CoV-2 RNA in the sewage samples, using PrimeScript™ One Step RT-PCR Kit (Takara Bio, USA) and 2019-nCoV RUO qPCR Probe Assay primer/probe mix (IDT, Integrated DNA Technologies, Leuven, Belgium). Different instruments were used by different participating labs, including CFX96 BioRad, LightCycler 480 (Roche Diagnostics, Germany), Stratagene Mx3005P (Applied Biosystems, USA) and QuantStudio 5 (Applied Biosystems, USA).

The S gene was analyzed by a duplex gene allelic discrimination TaqMan RT-qPCR assay, using 400 nM of the following primers targeting the S gene (For-S21708 5'ATTCAACTCAGGACTTGTTCTTACCTT3' and Rev-S21796 5'TAAATGGTAGGACAGGGTTATCAAAC3'), and 200 nM of the following probes (S_Probe6970in 5'FAM- TCCATGCTATACATGTCTCTGGGACCAATG BHQ1- 3' and S_Probe6970del 5'HEX- TTCCATGCTATCTCTGGGACCAATGGTACT BHQ1-3'). RT-qPCR mastermix was prepared using PrimeScript™ One Step RT-PCR Kit

(Takara Bio, USA), and temperature program was 10 min at 50ºC, 3 min at 95ºC, and 45 cycles of 3 seconds at 95ºC and 30 seconds at 60ºC.

RT-qPCR analysis for each target included the analysis of duplicate wells containing undiluted RNA, and duplicate wells containing a ten-fold dilution to monitor the presence of inhibitors. Every RT-qPCR assay included 4 wells corresponding to negative controls (2 nuclease-free water and 2 negative extraction controls). Commercially available Twist Synthetic SARS-CoV-2 RNA Controls (Control 2, MN908947.3; and Control 14, EPI_ISL_710528) were used to prepare standard curves for genome quantitation. Both synthetic RNA controls were quantified by droplet-based digital PCR using One-Step RT-ddPCR Advanced kit for probes in a QX200™ System (Bio-Rad), to estimate the exact concentration of genome copies (GC)/µl, prior to construction of RT-qPCR standard curves. Limit of detection (LOD) and limit of quantification (LOQ) were determined for each specific target by running a series of dilutions of the target with 4-10 replicates per dilution. Parameters of all standard curves and estimated LOD and LOQ for the 4 participating laboratories are summarized in supplementary **Table S1**.

RT-qPCR data analysis and interpretation. The following criteria were used to estimate SARS-CoV-2 gene viral titers. For each specific target, Cq values ≤40 were converted into GC/L using the corresponding standard curve and volumes tested. Occurrence of inhibition was estimated by comparing average viral titers obtained from duplicate wells tested on undiluted RNA with duplicate wells tested on ten-fold diluted RNA. Inhibition was ascertained when difference in average viral titers was higher than 0.5 $\log_{10}$, and viral titers inferred from the ten-fold RNA dilution. The percentage of

SARS-CoV-2 genomes containing the ΔHV69/70 deletion within the S gene was calculated using the following formula:

% = GC/L (Probe6970del) / [GC/L (Probe6970del) + GC/L (Probe6970in)] x 100. In cases with one of the concentrations <LOQ, percentage was calculated using the corresponding LOQ of the assay. Data with both concentrations <LOQ were not considered.

S gene sequencing and single nucleotide polymorphism (SNPs) identification. Full-length S gene sequencing of wastewater samples was performed following the ARTIC Network protocol (https://artic.network/ncov-2019, with minor modifications) with selected v3 primers (Integrated DNA Technologies) for genome amplification and KAPA HyperPrep Kit (Roche Applied Science) for library preparation [23]. Libraries were loaded in MiSeq Reagent Kit 600v3 cartridges and sequenced on MiSeq platform (Illumina). The raw sequenced reads were cleaned from low-quality segments, and mapped against the Wuhan-Hu-1 reference genome sequence to find out variant-specific signature mutations (mutations and indels).

RESULTS

Duplex SARS-CoV-2 S gene allelic discrimination RT-qPCR assay validation. To confirm the ability of the RT-qPCR assay to discriminate and estimate the proportion of targets containing the ΔHV69/70 deletion in different scenarios, 9 preparations containing 90:10, 50:50 or 10:90 proportions of B.1.1.7 and Wuhan-Hu-1 synthetic control RNAs at 3 different total concentration levels ($1x10^4$, $1x10^3$ and $1x10^2$ GC/rxn) were made and analyzed (**Figure 1**). Assays performed using B.1.1.7 and Wuhan-Hu-1 synthetic control RNAs did not show cross-reactivity (data not shown).

Results showed the ability of the method to detect and quantitatively discriminate sequences containing the ΔHV69/70 deletion from wildtype sequences in mixed samples through a wide range of total RNA concentrations often observed in extracted wastewater samples.

Additionally, to confirm that sequences detected in natural samples collected during the study period containing the ΔHV69/70 deletion corresponded to the B.1.1.7 variant, a subset of 8 samples (4 from week 21_11 and 4 from week 21_12), with B.1.1.7 proportion estimates ranging between 35% and 100% were analyzed by NGS sequencing of the S gene. Variant analysis showed a total of 12 nucleotide substitutions and 4 deletions in comparison with the reference genome of SARS-CoV-2 isolate Wuhan-Hu-1 (MN908947.3), most of them being specific for B.1.1.7 variant. No strong correlation was observed between coverage and specific amplicons. Between 3 and 7 mutation markers out of 9 markers specific for B.1.1.7 variants in the spike region (ΔHV69-70, ΔY144, N501Y, A570D, D614G, P681H, T716I, S982A and D1118H) were detected in all samples, confirming that the RT-qPCR assay could be used to trace the occurrence of B.1.1.7 variant, as suggested by the recently published EU recommendation [4] (**Figure 2**). Seven additional amino acid substitutions/deletions, which were not specific for B.1.1.7 variant, were also detected: G142V, A222V, G257V, S375P, ΔT376, F377L and K537E. Of these substitutions, as of April 28th, G142V and G257V had already been reported in 1128 and 259 sequences published at GISAID database, respectively, but the others had only been reported at low frequencies. Substitutions S375P/ΔT376/F377L affecting 3 consecutive residues within the Receptor Binding Domain (RBD) and identified in 35% of sequences from WWTP-31 in Madrid, have been individually reported less than 15 times in several countries,

238 including Spain. However, our data report the occurrence of the 3 mutations within the
239 same variant. Of note, mutations in these residues have been related to antigenic drift
240 [24,25]. Substitution K537E had been reported once in a nasopharyngeal specimen
241 belonging to B.1.177 (EPI_ISL_1547898; hCoV-19/Slovakia/UKBA-2586/2021).
242 Marker A222V is also present in B.1.777 variant, which originated in Spain became the
243 predominant variant in most European countries during the second pandemic wave [26].
244 Sequence data obtained in this study are available at Genbank (SAMN19107574 and
245 PRJNA728923).
246
247 **Temporal and geographical emergence of B.1.1.7 variant in the Spanish territory.**
248 Wastewater samples from 32 Spanish WWTPs from mid December 2020 to end of
249 March 2021 were weekly analyzed to monitor the emergence of B.1.1.7 in the territory.
250 Total levels of SARS-CoV-2 RNA were determined by RT-qPCR using N1 target as
251 well as the S discriminatory RT-qPCR, without normalization by the population number
252 (**Figure 3**). A moderate correlation was observed between both SARS-CoV-2 genome
253 concentration measures between N1 and total S gene titers (R^2=0.303, data not shown).
254 From all samples analyzed for different SARS-CoV-2 RT-qPCR assays, inhibition was
255 observed in 17.4%, 19.5% and 16.1% of samples for N1, S target (wildtype) and S
256 target (B.1.1.7) assays, respectively, without significant differences between targets.
257 Regarding recoveries of the 3 animal coronaviruses used as process control viruses by
258 different participating laboratories, recovery percentages (mean±standard deviation)
259 were of 23.4±16.0%, 24.9±22.6%, and 52.9±29.8%, for TGEV, MHV, and PEDV,
260 respectively. The Kruskal-Wallis test showed significant differences for PEDV
261 ($p<0.001$).

Lockdown measures in Spain during the study period were remarkable (mandatory use of face mask, nighttime curfews, restrictions regarding bar and restaurant opening times, social gathering restrictions, restricted opening hours and attendance, municipality of residence confinement implemented in most regions, etc.) and were associated to the nationwide State of Alarm, in place since October 2020. As most European countries, at the clinical level, a peak in COVID-19 cases occurred between the end of December 2020 (week 20_52) and early February 2021 (week 21_05). As measured through the N1 target, a peak in SARS-CoV-2 genome levels in wastewater was observed in week 21_01 in 9 regions including Zaragoza (**Fig 3B**; WWTP-15), Las Palmas in Canary Islands (**Fig 3D**; WWTP-18), 3 cities in Catalonia (**Fig 3H**; WWTP-26, 27 and 28) and Madrid (**Fig 3I**; WWTP 07, 08, 31 and 32); in week 21_02 in regions including Córdoba and Granada in Andalucía (**Fig 3A**; WWTP03 and WWTP-04), Palma de Mallorca (**Fig 3C**; WWTP-17), Santander (**Fig 3E**; WWTP-20), Cuenca (**Fig 3F**; WWTP-29), Madrid (**Fig 3I**; WWTP-30) and Logroño (**Fig 3M**; WWTP-14); in week 21_03 in Guadalajara (**Fig 3F**; WWTP-25), Segovia (**Fig 3G**; WWTP-21), Soria (**Fig 3G**; WWTP-22), Madrid (**Fig 3I**; WWTP-09), Valencia (**Fig 3J**; WWTP-01) and Ourense (**Fig 3L**; WWTP-05); in week 21_04 in Málaga (**Fig 3A**; WWTP_06), Albacete (**Fig 3F**; WWTP-24), Valladolid (**Fig 3G**; WWTP-23), Bilbao (**Fig 3N**; WWTP-13) and Oviedo (**Fig 3O**; WWTP-16) and, and in week 21_05 in Sevilla (**Fig 3A**; WWTP-10) and Vitoria (**Fig 3N**; WWTP-12). SARS-CoV-2 genome titers in Tenerife (Canary Islands) (**Fig 3D**; WWTP-19), were already at peak titers in the first week of the study (**Figure 3**).

First detection of B.1.1.7 variant in wastewater samples occurred in the Southern city of Málaga (Andalucía) in week 20_52 (**Figure 4**). The first week of January 2021, it could

also be detected in the 2 largest cities (Madrid and Barcelona), in 2 Northern cities (Santander and Vitoria), another city in Andalucía (Córdoba) and in Tenerife (Canary Islands), suggesting that multiple introductions occurred during Christmas holidays in different parts of the Spanish territory, and representing 20% of all sampled WWPTs. B.1.1.7 levels detected in week 21_01 were lower than 10% in most WWTPs, with the exception of WWTP-32 in Madrid, which was of 22.9%. Percentages of WWTPs with B.1.1.7 detection increased progressively, up to 56% by week 21_04, 91% by week 21_08, and 97% by week 21_13. Our data also showed that while total level of SARS-CoV-2 RNA genomes measured by N1 target showed a slight decline for several weeks after the first peak observed during early 2021, this negative trend was slowed or reversed in most WWTPs from the time when the proportion of B.1.1.7 became more abundant (**Figure 3**). In 6 cities, a significant increase of N1 RNA titers higher than 1 $\log_{10}$ with respect to preceding week, as adopted by the VATar COVID-19 Spanish Reporting System (available at https://www.miteco.gob.es/es/agua/temas/concesiones-y-autorizaciones/nota-tecnica-vatar-miterd_tcm30-517518.pdf), was observed in the last week of the study (**Fig 3E** WWTP-20, **Fig 3F** WWTP 29, **Fig 3G** WWTP 23, **Fig 3I** WWTP07, 08 and 30).

The relative proportion of the B.1.1.7 variant in wastewater could be estimated for 91% of positive samples. **Figure 4** shows the heatmap of the evolution of B.1.1.7 prevalence in wastewater over time. Predominance of B.1.1.7 variant with prevalences ≥90% was reached in all WWTPs except in WWTP 9 and 32 in Madrid, and WWTP-15 (Zaragoza). In Madrid B.1.1.7 reached 48% and 86%, and in Zaragoza 51% at the end of the study period, although it kept progressively increasing thereafter (data not shown). When considering data from 9 WWTPs, which showed at least 2 consecutive

312 weeks with B.1.1.7 percentages near fixation (90-100%) as a confirmation of
313 predominance (**Table S2**), we could estimate that approximately 8.1±2.0 weeks-time
314 were required for B.1.1.7 variant to become predominant in sewage. This would
315 correspond to an average increase rate of 11.7% (9.6-15.1%) per week.
316
317 The proportion of B.1.1.7 was compared to the prevalence detected at the clinical level.
318 The abundance of B.1.1.7, as a fraction of all sequenced clinical specimens by the local
319 authorities in each Autonomous Community, was obtained from update reports of the
320 epidemiological situation of the variants of SARS-CoV-2 of importance published by
321 the Spanish Ministry of Health [27]. When comparing the proportion of B.1.1.7 estimated
322 from wastewater with the proportion estimated from sequencing of clinical isolates, a
323 good correlation was observed (R^2=0.5012) (**Figure 5A**). Analysis comparing B.1.1.7
324 prevalence estimate from clinical specimens with prevalence estimated from wastewater
325 one or two weeks before did not increase the correlation determinant (data not shown),
326 suggesting that wastewater analysis did not allow us to anticipate the increase in B.1.1.7
327 proportion. On a geographical/temporal analysis, wastewater testing allowed to confirm
328 circulation of B.1.1.7 variant before its identification in clinical specimens, especially in
329 part due to a noticeable clinical undertesting in most regions. Data for 3 selected weeks
330 are shown in **Figure 5B**. By week 21_04 at the end of January, at the clinical level the
331 variant was only detected, on clinical samples, in Galicia and País Vasco, mainly due to
332 the low number of sequenced isolates in most regions, while it was detected in 18/32
333 (56%) WWTPs. At the end of the study, all Autonomous Communities Public Health
334 departments reported percentages higher than 60%, but while some WWTPs showed
335 very high percentages some others showed relatively lower proportions, indicating
336 differences at the local level.

DISCUSSION

As a cost-effective approach to screen thousands of inhabitants, wastewater-based epidemiology (WBE) is a valuable tool to anticipate the circulation of specific pathogens in a community and to closely track their incidence evolution through space and time [28,29]. This surveillance has proven to be extremely useful to monitor the circulation of total SARS-CoV-2 in different parts of the world [1–3], and may be effective for tracking novel SARS-CoV-2 VOC. Despite NGS would allow the definitive identification of specific variants and has already been applied on sewage samples [30,31], it is resource and time-consuming, thus limiting the number of samples that can be processed and the number of labs which can implement this approach on a regular basis. In addition, most deep sequencing protocols allow the identification of signature mutations within short individual reads, which when detected in wastewater samples containing a mixture of different isolates, may not be proof of co-occurrence of such mutations within the same genome, thus providing only an indirect evidence of the presence of a certain variant [32]. Finally, the use of NGS is also challenged by the presence of inhibitors in samples, limiting its success rate and depth coverage. Given its higher tolerance to inhibitors, droplet digital RT-PCR has been acknowledged as a suitable approach to simultaneous enumerate the concentration of variants with the N501Y mutation and wildtype in wastewater [33], but droplet digital RT-PCR widespread use may be limited nowadays by the high economic investment in instrumentation. On the other hand, the use of RT-qPCR methods offers the advantage of rapid turnaround time, lower cost, and immediate availability in most public health laboratories. In the current study, we validated a duplex RT-qPCR assay to discriminate and enumerate SARS-CoV-2 variants containing the ΔHV69/70 deletion from variants lacking it.

Among molecular markers specific for B.1.1.7 variant, the 6-nucleotide deletion corresponding to residues 69/70 was chosen because it offers the possibility to design highly-specific robust probes to be used in wastewater samples, which unlike clinical specimens, will contain mixed sequences in most cases. Similar to the TaqPath COVID-19 assay (Thermo Fisher Scientific) and other RT-qPCR protocols designed for clinical diagnosis [34], the novel duplex RT-qPCR assay developed in this study proved highly specific and discriminatory.

The ΔHV69/70 deletion is located within the N-terminal domain of the S glycoprotein and has been described to be located at a recurrent deletion region (RDR), and phylogenetic studies showed that it has arisen independently at least 13 times [35]. In addition to being a signature mutation of highly transmissible B.1.1.7 variant, it has also been described in other lineages, including cluster-5 variant, identified both in minks and humans in Denmark, some isolates belonging to 20A/S:439K variant, which emerged twice independently in Europe, B.1.258 and B.1.525 lineages [35–39], although none of these other lineages have been shown to spread widely. According to GISAID, from a dataset of 442,175 sequences collected from 1 December 2020 to 31st March 2021 containing Δ69, as a hallmark of ΔHV69/70 deletion, the proportion of sequences which were classified as B.1.1.7 was of 92.4% (for clinical sequences isolated in Spain during the same period, this percentage was of 98.1%). Among sequences containing Δ69 not classified as B.1.1.7 variant, other lineages including B.1.258, B.1.525, B.1.177, B.1.429+B.1.427, P1, B.1.351 and B.1.617 were observed in a minority of cases. Among sequences belonging to the predominant lineage in Spain at the onset of this study, B.1.177, only 0.23% of sequences deposited in GISAID contained Δ69, confirming that detection ΔHV69/70 is highly indicative of a genome belonging to

B.1.1.7 lineage. Finally, in the NGS analysis performed on 8 selected samples with a high proportion of ΔHV69/70 containing genomes, between 3-8 additional B.1.1.7 mutation signatures were identified, confirming that the detected genomes very likely correspond to the B.1.1.7 variant.

During the study period, weekly wastewater estimates of the proportion of B.1.1.7, representing a larger and more comprehensive proportion of typed cases including both symptomatic and asymptomatic cases, well reflected the trends in the reported sequenced clinical cases in most regions. Despite the number of clinical specimens sequenced from public health laboratories was not high during the study period, and showed strong geographic differences, a correlation was observed between the proportion of B.1.1.7 cases observed at the clinical level and data estimated from sewage when using samples from the same week (**Figure 5A**). Of note, this association was not more robust when using data from 1-2 previous weeks (data not shown). The lack of anticipation ability could be due to several unknown factors, including differences in shedding and kinetic levels between variants, differences in the proportion of asymptomatic infections, and differences in environmental stability. Sewage surveillance allowed the identification of B.1.1.7 circulation in the Spanish territory in the Southern city of Málaga before it was confirmed at the clinical level by National Public Health Authorities, and allowed us to infer multiple simultaneous introductions during Christmas and New Year's holidays in distant parts of the country (Madrid, Barcelona, Santander, Vitoria, Córdoba and Tenerife). By the end of January 2021, only 13% (2/15) Autonomous Communities had reported B.1.1.7 clinical cases, while circulation in sewage had been confirmed in 67% (10/15) of them, confirming its use as an early warning approach. Data from 11 WWTPs which reached B.1.1.7 near

412 fixation rates, defined as higher than 90% for ≥ 2 consecutive weeks, showed that
413 8.1±1.8 weeks were required to reach B.1.1.7 predominance, which would be a slightly
414 shorter time than what has been locally observed at the clinical level. A research
415 publication reported first detection of imported B.1.1.7 clinical cases in Madrid in week
416 20_52 (December 2020), and a proportion of 62% of total newly diagnosed COVID-19
417 cases 10 weeks later [40]. Data from other studies are also the UK reported by ECDC
418 show that B.1.1.7 cases went from less than 5% of all positive cases to more than 60%
419 in less than six weeks during November to mid-December 2020 [41], and Davies et al.
420 demonstrated that it became dominant throughout the country [15]. Estimates from the US
421 indicate that B.1.1.7 would become dominant in most states 4 months after its first
422 identification in late November 2020 [14].
423
424 Our data also showed that predominance of B.1.1.7 variant appeared to correspond to a
425 slowdown in the negative trend of total SARS-CoV-2 wastewater levels, which had
426 been observed from early 2021 in most cities, probably due to lockdown measures and
427 the mass-vaccination campaign, which was initiated the last week of 2020. Even in
428 some cities, including Santander (**Fig 3E**; WWTP-20), Cuenca (**Fig 3F**; WWTP-29),
429 Valladolid (**Fig 3G**; WWTP-23) and Madrid (**Fig 3I**; WWTP-07, 08 and 30), total
430 SARS-CoV-2 levels showed a positive trend at the end of the study. These results
431 suggest that the emergence of B.1.1.7 cases could have produced a higher transmission
432 rate and a slight increase in COVID-19 incidence, as confirmed by clinical
433 epidemiological data reported by the Spanish Ministry of Health, reporting an incidence
434 peak between the end of March and April 2020 [42]. Despite this positive trend markedly
435 observed in some regions, the smooth running of the mass vaccination campaign
436 starting on December 27th 2020, in addition to non-pharmaceutical interventions, likely

437 contributed to minimizing the impact of B.1.1.7 emergence. As of end of March 2021,
438 the percentage of the Spanish population who had been partially immunized or totally
439 vaccinated were of 13.2 % and 6.8 %, respectively.
440
441 Finally, despite the state of alarm decreed by the Spanish government, as a measure to
442 unify confinement and restriction measures across the country, was maintained
443 throughout the study period, predominance of B.1.1.7 variant was not homogeneous,
444 and dynamics were variable among cities across the country. For instance, a rapid
445 predominance of B.1.1.7 variant was observed in Granada (WWTP-4) and Cáceres
446 (WWTP-36), while in other cities, B.1.1.7 reached prevalences higher than 90% by
447 week 3-6 after first positive detection and decreased thereafter for 2-3 weeks. Several
448 reasons could explain these waves, including differences in regional social distancing
449 behaviors, repetitive B.1.1.7 case imports, introduction of additional variants, climatic
450 effect on sewage composition, size of WWTP, or variability related to the use of grab
451 samples instead of composite samples.
452
453 This study highlights the use of WBE as a cost-effective, non-invasive and unbiased
454 approach which may complement clinical testing during the COVID-19 pandemic, and
455 demonstrates the applicability of duplex RT-qPCR assays on sewage surveillance as a
456 rapid, attractive, and resourceful method to track the early circulation and emergence of
457 known VOC in a population, especially at times when clinical typing is insufficient and
458 when signature mutations can be unequivocally assigned to a specific VOC. The current
459 strategy could be readily adaptable to track specific mutations of other VOC as soon as
460 they are identified by clinical genomic sequencing in the future and integrated into
461 existing wastewater surveillance programs.

462

463 **Supporting Information:** Additional data on parameters defining standard curves,
464 limit of detection (LOD), and limit of quantification (LOQ) for RTqPCR assays used in
465 the study (Table S1), and data used to estimate the time required to reach B.1.1.7
466 prevalence >90% in wastewater (Table S2).

467

468

469 **ACKNOWLEDGEMENTS**

470 This work was partially supported by the COVID-19 wastewater surveillance project
471 (VATar COVID19), funded by the Spanish Ministry for the Ecological Transition and
472 the Demographic Challenge of and the Spanish Ministry of Health; grants from CSIC
473 (202070E101) and MICINN co-founded by AEI FEDER, UE (AGL2017-82909); grant
474 ED431C 2018/18 from the Consellería de Educación, Universidade e Formación
475 Profesional, Xunta de Galicia (Spain); Direcció General de Recerca i Innovació en Salut
476 (DGRIS) Catalan Health Ministry Generalitat de Catalunya through Vall d'Hebron
477 Research Institute (VHIR), and Centro para el Desarrollo Tecnológico Industrial
478 (CDTI) from the Spanish Ministry of Economy and Business, grant number IDI-
479 20200297. Pilar Truchado is holding a Ramón y Cajal contract from the Ministerio de
480 Ciencia e Innovación. Adan Martinez is holding a predoctoral fellowship FI_SDUR
481 from Generalitat de Catalunya.
482 We gratefully acknowledge all the staff involved in the VATar COVID-19 project,
483 working with sample collection and logistics. The authors are grateful to Promega
484 Corporation (Madison, US) for technical advice, and thank Andrea Lopez de Mota for
485 her technical support.

486

CONFLICT OF INTEREST

None declared.

AUTHORS' CONTRIBUTIONS

Study design: Rosa M Pintó, Albert Bosch, Susana Guix, Ana Allende, Gloria Sanchez, Jesus L Romalde. Data generation and interpretation: Albert Carcereny, Adan Martínez-Velázquez, Pilar Truchado, Jenifer Cascales, Marta Lois, David Polo, Alba Pérez-Cataluña, Azahara Díaz-Reolid, Josep Gregori, Damir Garcia-Cehic. Resources: Margarita Palau, Cristina González Ruano. Funding acquisition: Albert Bosch, Ana Allende, Gloria Sánchez, Jesus L Romalde, Andres Anton, Josep Quer. Conceptualization: Rosa M Pintó, Albert Bosch, Susana Guix. Manuscript writing: Rosa M Pintó, Susana Guix. Critically revision of the manuscript: All authors.

REFERENCES

(1) Medema, G.; Been, F.; Heijnen, L.; Petterson, S. Implementation of Environmental Surveillance for SARS-CoV-2 Virus to Support Public Health Decisions: Opportunities and Challenges. *Current Opinion in Environmental Science and Health*. Elsevier B.V. October 1, 2020, pp 49–71. https://doi.org/10.1016/j.coesh.2020.09.006.

(2) Prado, T.; Fumian, T. M.; Mannarino, C. F.; Resende, P. C.; Motta, F. C.; Eppinghaus, A. L. F.; Chagas do Vale, V. H.; Braz, R. M. S.; de Andrade, J. da S. R.; Maranhão, A. G.; Miagostovich, M. P. Wastewater-Based Epidemiology as a Useful Tool to Track SARS-CoV-2 and Support Public Health Policies at Municipal Level in Brazil. *Water Research* **2021**, *191*, 116810. https://doi.org/10.1016/j.watres.2021.116810.

(3) Wastewater Surveillance of SARS-CoV-2 across 40 U.S. States. https://doi.org/10.1101/2021.03.10.21253235.

(4) EC. COMMISSION RECOMMENDATION of 17.3.2021 on a Common Approach to Establish a Systematic Surveillance of SARS-CoV-2 and Its Variants in Wastewaters in the EU. 2021.

(5) Ahmed, W.; Angel, N.; Edson, J.; Bibby, K.; Bivins, A.; O'Brien, J. W.; Choi, P. M.; Kitajima, M.; Simpson, S. L.; Li, J.; Tscharke, B.; Verhagen, R.; Smith, W. J. M.; Zaugg, J.; Dierens, L.; Hugenholtz, P.; Thomas, K. v; Mueller, J. F. First Confirmed Detection of

SARS-CoV-2 in Untreated Wastewater in Australia: A Proof of Concept for the Wastewater Surveillance of COVID-19 in the Community. *The Science of the total environment* **2020**, *728*, 138764. https://doi.org/10.1016/j.scitotenv.2020.138764.

(6) la Rosa, G.; Iaconelli, M.; Mancini, P.; Bonanno Ferraro, G.; Veneri, C.; Bonadonna, L.; Lucentini, L.; Suffredini, E. First Detection of SARS-CoV-2 in Untreated Wastewaters in Italy. *Science of the Total Environment* **2020**, *736*, 139652. https://doi.org/10.1016/j.scitotenv.2020.139652.

(7) Randazzo, W.; Truchado, P.; Cuevas-Ferrando, E.; Simón, P.; Allende, A.; Sánchez, G. SARS-CoV-2 RNA in Wastewater Anticipated COVID-19 Occurrence in a Low Prevalence Area. *Water Research* **2020**, *181*, 115942. https://doi.org/10.1016/j.watres.2020.115942.

(8) Chavarria-Miró, G.; Anfruns-Estrada, E.; Martínez-Velázquez, A.; Vázquez-Portero, M.; Guix, S.; Paraira, M.; Galofré, B.; Sánchez, G.; Pintó, R. M.; Bosch, A. Time Evolution of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in Wastewater during the First Pandemic Wave of COVID-19 in the Metropolitan Area of Barcelona, Spain. *Applied and environmental microbiology* **2021**, *87* (7), e02750-20. https://doi.org/10.1128/AEM.02750-20.

(9) Wurtzer, S.; Marechal, V.; Mouchel, J. M.; Maday, Y.; Teyssou, R.; Richard, E.; Almayrac, J. L.; Moulin, L. Evaluation of Lockdown Effect on SARS-CoV-2 Dynamics through Viral Genome Quantification in Waste Water, Greater Paris, France, 5 March to 23 April 2020. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **2020**, *25* (50), 2000776. https://doi.org/10.2807/1560-7917.ES.2020.25.50.2000776.

(10) Polo, D.; Quintela-Baluja, M.; Corbishley, A.; Jones, D. L.; Singer, A. C.; Graham, D. W.; Romalde, J. L. Making Waves: Wastewater-Based Epidemiology for COVID-19 - Approaches and Challenges for Surveillance and Prediction. *Water research* **2020**, *186*, 116404. https://doi.org/10.1016/j.watres.2020.116404.

(11) Rooney, C. M.; Moura, I. B.; Wilcox, M. H. Tracking COVID-19 via Sewage. *Current Opinion in Gastroenterology* **2021**, *37* (1).

(12) Saththasivam, J.; El-Malah, S. S.; Gomez, T. A.; Jabbar, K. A.; Remanan, R.; Krishnankutty, A. K.; Ogunbiyi, O.; Rasool, K.; Ashhab, S.; Rashkeev, S.; Bensaad, M.; Ahmed, A. A.; Mohamoud, Y. A.; Malek, J. A.; Abu Raddad, L. J.; Jeremijenko, A.; Abu Halaweh, H. A.; Lawler, J.; Mahmoud, K. A. COVID-19 (SARS-CoV-2) Outbreak Monitoring Using Wastewater-Based Epidemiology in Qatar. *The Science of the total environment* **2021**, *774*, 145608. https://doi.org/10.1016/j.scitotenv.2021.145608.

(13) Meera Chand, Susan Hopkins, Gavin Dabrera, Christina Achison, Wendy Barclay, Neil Ferguson, Erik Volz, Nick Loman, Andrew Rambaut, J. B. *Investigation of Novel SARS-COV-2 Variant Variant of Concern 202012/01 (PHE)*; 2020.

(14) Washington, N. L.; Gangavarapu, K.; Zeller, M.; Bolze, A.; Cirulli, E. T.; Schiabor Barrett, K. M.; Larsen, B. B.; Anderson, C.; White, S.; Cassens, T.; Jacobs, S.; Levan, G.; Nguyen, J.; Ramirez 3rd, J. M.; Rivera-Garcia, C.; Sandoval, E.; Wang, X.; Wong, D.; Spencer, E.; Robles-Sikisaka, R.; Kurzban, E.; Hughes, L. D.; Deng, X.; Wang, C.; Servellita, V.; Valentine, H.; de Hoff, P.; Seaver, P.; Sathe, S.; Gietzen, K.; Sickler, B.; Antico, J.; Hoon,

K.; Liu, J.; Harding, A.; Bakhtar, O.; Basler, T.; Austin, B.; MacCannell, D.; Isaksson, M.; Febbo, P. G.; Becker, D.; Laurent, M.; McDonald, E.; Yeo, G. W.; Knight, R.; Laurent, L. C.; de Feo, E.; Worobey, M.; Chiu, C. Y.; Suchard, M. A.; Lu, J. T.; Lee, W.; Andersen, K. G. Emergence and Rapid Transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell* **2021**, *184* (10), 2587-2594.e7. https://doi.org/10.1016/j.cell.2021.03.052.

(15) Davies, N. G.; Jarvis, C. I.; van Zandvoort, K.; Clifford, S.; Sun, F. Y.; Funk, S.; Medley, G.; Jafari, Y.; Meakin, S. R.; Lowe, R.; Quaife, M.; Waterlow, N. R.; Eggo, R. M.; Lei, J.; Koltai, M.; Krauer, F.; Tully, D. C.; Munday, J. D.; Showering, A.; Foss, A. M.; Prem, K.; Flasche, S.; Kucharski, A. J.; Abbott, S.; Quilty, B. J.; Jombart, T.; Rosello, A.; Knight, G. M.; Jit, M.; Liu, Y.; Williams, J.; Hellewell, J.; O'Reilly, K.; Chan, Y.-W. D.; Russell, T. W.; Procter, S. R.; Endo, A.; Nightingale, E. S.; Bosse, N. I.; Villabona-Arenas, C. J.; Sandmann, F. G.; Gimma, A.; Abbas, K.; Waites, W.; Atkins, K. E.; Barnard, R. C.; Klepac, P.; Gibbs, H. P.; Pearson, C. A. B.; Brady, O.; Edmunds, W. J.; Jewell, N. P.; Diaz-Ordaz, K.; Keogh, R. H.; Group, C. C.-19 W. Increased Mortality in Community-Tested Cases of SARS-CoV-2 Lineage B.1.1.7. *Nature* **2021**. https://doi.org/10.1038/s41586-021-03426-1.

(16) 9510 DETECTION OF ENTERIC VIRUSES. In *Standard Methods For the Examination of Water and Wastewater*; Standard Methods for the Examination of Water and Wastewater; American Public Health Association, 2018. https://doi.org/doi:10.2105/SMWW.2882.202.

(17) Moreno, J. L.; Zúñiga, S.; Enjuanes, L.; Sola, I. Identification of a Coronavirus Transcription Enhancer. *JOURNAL OF VIROLOGY* **2008**, *82* (8), 3882–3893. https://doi.org/10.1128/JVI.02622-07.

(18) Manaker, R. A.; Piczak, C. v; Miller, A. A.; Stanton, M. F. A Hepatitis Virus Complicating Studies With Mouse Leukemia. *JNCI: Journal of the National Cancer Institute* **1961**, *27* (1), 29–51. https://doi.org/10.1093/jnci/27.1.29.

(19) Raaben, M.; Einerhand, A. W. C.; Taminiau, L. J. A.; van Houdt, M.; Bouma, J.; Raatgeep, R. H.; Büller, H. A.; de Haan, C. A. M.; Rossen, J. W. A. Cyclooxygenase Activity Is Important for Efficient Replication of Mouse Hepatitis Virus at an Early Stage of Infection. *Virology Journal* **2007**, *4* (1), 55. https://doi.org/10.1186/1743-422X-4-55.

(20) Vemulapalli, R.; Gulani, J.; Santrich, C. A Real-Time TaqMan RT-PCR Assay with an Internal Amplification Control for Rapid Detection of Transmissible Gastroenteritis Virus in Swine Fecal Samples. *Journal of virological methods* **2009**, *162* (1–2), 231–235. https://doi.org/10.1016/j.jviromet.2009.08.016.

(21) Zhou, X.; Zhang, T.; Song, D.; Huang, T.; Peng, Q.; Chen, Y.; Li, A.; Zhang, F.; Wu, Q.; Ye, Y.; Tang, Y. Comparison and Evaluation of Conventional RT-PCR, SYBR Green I and TaqMan Real-Time RT-PCR Assays for the Detection of Porcine Epidemic Diarrhea Virus. *Molecular and Cellular Probes* **2017**, *33*, 36–41. https://doi.org/https://doi.org/10.1016/j.mcp.2017.02.002.

(22) ISO15216-1:2017. Microbiology of the Food Chain - Horizontal Method for Determination of Hepatitis A Virus and Norovirus Using Real-Time RT-PCR - Part 1: Method for Quantification. 2017.

(23) Andrés, C.; Garcia-Cehic, D.; Gregori, J.; Piñana, M.; Rodriguez-Frias, F.; Guerrero-Murillo, M.; Esperalba, J.; Rando, A.; Goterris, L.; Codina, M. G.; Quer, S.; Martín, M. C.;

Campins, M.; Ferrer, R.; Almirante, B.; Esteban, J. I.; Pumarola, T.; Antón, A.; Quer, J. Naturally Occurring SARS-CoV-2 Gene Deletions Close to the Spike S1/S2 Cleavage Site in the Viral Quasispecies of COVID19 Patients. *Emerging Microbes & Infections* **2020**, *9* (1), 1900–1911. https://doi.org/10.1080/22221751.2020.1806735.

(24) Elbe, S.; Buckland-Merrett, G. Data, Disease and Diplomacy: GISAID's Innovative Contribution to Global Health. *Global challenges (Hoboken, NJ)* **2017**, *1* (1), 33–46. https://doi.org/10.1002/gch2.1018.

(25) Greaney, A. J.; Starr, T. N.; Gilchuk, P.; Zost, S. J.; Binshtein, E.; Loes, A. N.; Hilton, S. K.; Huddleston, J.; Eguia, R.; Crawford, K. H. D.; Dingens, A. S.; Nargi, R. S.; Sutton, R. E.; Suryadevara, N.; Rothlauf, P. W.; Liu, Z.; Whelan, S. P. J.; Carnahan, R. H.; Crowe, J. E.; Bloom, J. D. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain That Escape Antibody Recognition. *Cell host & microbe* **2021**, *29* (1), 44-57.e9. https://doi.org/10.1016/j.chom.2020.11.007.

(26) Hodcroft, E. B.; Zuber, M.; Nadeau, S.; Vaughan, T. G.; Crawford, K. H. D.; Althaus, C. L.; Reichmuth, M. L.; Bowen, J. E.; Walls, A. C.; Corti, D.; Bloom, J. D.; Veesler, D.; Mateo, D.; Hernando, A.; Comas, I.; González Candelas, F.; Stadler, T.; Neher, R. A. Emergence and Spread of a SARS-CoV-2 Variant through Europe in the Summer of 2020. *medRxiv* **2021**, 2020.10.25.20219063. https://doi.org/10.1101/2020.10.25.20219063.

(27) Sanidad, M. de. *Actualización de La Situación Epidemiológica de Las Variantes de SARS-CoV-2 de Importancia En Salud Pública En España*; 2021.

(28) Alfonsi, V.; Romanò, L.; Ciccaglione, A. R.; la Rosa, G.; Bruni, R.; Zanetti, A.; della Libera, S.; Iaconelli, M.; Bagnarelli, P.; Capobianchi, M. R.; Garbuglia, A. R.; Riccardo, F.; Tosti, M. E.; Group, C. Hepatitis E in Italy: 5 Years of National Epidemiological, Virological and Environmental Surveillance, 2012 to 2016. *Eurosurveillance* **2018**, *23* (41). https://doi.org/https://doi.org/10.2807/1560-7917.ES.2018.23.41.1700517.

(29) Bisseux, M.; Colombet, J.; Mirand, A.; Roque-Afonso, A.-M.; Abravanel, F.; Izopet, J.; Archimbaud, C.; Peigue-Lafeuille, H.; Debroas, D.; Bailly, J.-L.; Henquell, C. Monitoring Human Enteric Viruses in Wastewater and Relevance to Infections Encountered in the Clinical Setting: A One-Year Experiment in Central France, 2014 to 2015. *Eurosurveillance* **2018**, *23* (7).

(30) Izquierdo-Lara, R.; Elsinga, G.; Heijnen, L.; Munnink, B. B. O.; Schapendonk, C. M. E.; Nieuwenhuijse, D.; Kon, M.; Lu, L.; Aarestrup, F.; Lycett, S.; Medema, G.; Koopmans, M. P. G.; de Graaf, M. Monitoring SARS-CoV-2 Circulation and Diversity through Community Wastewater Sequencing, the Netherlands and Belgium. *Emerging Infectious Disease journal* **2021**, *27* (5), 1405. https://doi.org/10.3201/eid2705.204410.

(31) Nemudryi, A.; Nemudraia, A.; Wiegand, T.; Surya, K.; Buyukyoruk, M.; Cicha, C.; Vanderwood, K. K.; Wilkinson, R.; Wiedenheft, B. Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. *Cell Reports Medicine* **2020**, *1* (6). https://doi.org/10.1016/j.xcrm.2020.100098.

(32) Jahn, K.; Dreifuss, D.; Topolsky, I.; Kull, A.; Ganesanandamoorthy, P.; Fernandez-Cassi, X.; Bänziger, C.; Stachler, E.; Fuhrmann, L.; Jablonski, K. P.; Chen, C.; Aquino, C.; Stadler, T.; Ort, C.; Kohn, T.; Julian, T. R.; Beerenwinkel, N. Detection of SARS-CoV-2 Variants in

Switzerland by Genomic Analysis of Wastewater Samples. *medRxiv* **2021**, 2021.01.08.21249379. https://doi.org/10.1101/2021.01.08.21249379.

(33) Heijnen, L.; Elsinga, G.; de Graaf, M.; Molenkamp, R.; Koopmans, M. P. G.; Medema, G. Droplet Digital RT-PCR to Detect SARS-CoV-2 Variants of Concern in Wastewater. *medRxiv* **2021**, 2021.03.25.21254324. https://doi.org/10.1101/2021.03.25.21254324.

(34) Vogels, C. B. F.; Breban, M. I.; Ott, I. M.; Alpert, T.; Petrone, M. E.; Watkins, A. E.; Kalinich, C. C.; Earnest, R.; Rothman, J. E.; Goes de Jesus, J.; Morales Claro, I.; Magalhães Ferreira, G.; Crispim, M. A. E.; Network, B.-U. C. G.; Singh, L.; Tegally, H.; Anyaneji, U. J.; Africa, N. for G. S. in S.; Hodcroft, E. B.; Mason, C. E.; Khullar, G.; Metti, J.; Dudley, J. T.; MacKay, M. J.; Nash, M.; Wang, J.; Liu, C.; Hui, P.; Murphy, S.; Neal, C.; Laszlo, E.; Landry, M. L.; Muyombwe, A.; Downing, R.; Razeq, J.; de Oliveira, T.; Faria, N. R.; Sabino, E. C.; Neher, R. A.; Fauver, J. R.; Grubaugh, N. D. Multiplex QPCR Discriminates Variants of Concern to Enhance Global Surveillance of SARS-CoV-2. *PLOS Biology* **2021**, *19* (5), e3001236.

(35) McCarthy, K. R.; Rennick, L. J.; Nambulli, S.; Robinson-McCarthy, L. R.; Bain, W. G.; Haidar, G.; Duprex, W. P. Recurrent Deletions in the SARS-CoV-2 Spike Glycoprotein Drive Antibody Escape. *Science (New York, N.Y.)* **2021**, *371* (6534), 1139–1142. https://doi.org/10.1126/science.abf6950.

(36) Thomson, E. C.; Rosen, L. E.; Shepherd, J. G.; Spreafico, R.; da Silva Filipe, A.; Wojcechowskyj, J. A.; Davis, C.; Piccoli, L.; Pascall, D. J.; Dillen, J.; Lytras, S.; Czudnochowski, N.; Shah, R.; Meury, M.; Jesudason, N.; de Marco, A.; Li, K.; Bassi, J.; O'Toole, A.; Pinto, D.; Colquhoun, R. M.; Culap, K.; Jackson, B.; Zatta, F.; Rambaut, A.; Jaconi, S.; Sreenu, V. B.; Nix, J.; Jarrett, R. F.; Beltramello, M.; Nomikou, K.; Pizzuto, M.; Tong, L.; Cameroni, E.; Johnson, N.; Wickenhagen, A.; Ceschi, A.; Mair, D.; Ferrari, P.; Smollett, K.; Sallusto, F.; Carmichael, S.; Garzoni, C.; Nichols, J.; Galli, M.; Hughes, J.; Riva, A.; Ho, A.; Semple, M. G.; Openshaw, P. J. M.; Baillie, J. K.; Rihn, S. J.; Lycett, S. J.; Virgin, H. W.; Telenti, A.; Corti, D.; Robertson, D. L.; Snell, G. The Circulating SARS-CoV-2 Spike Variant N439K Maintains Fitness While Evading Antibody-Mediated Immunity. *bioRxiv* **2020**, 2020.11.04.355842. https://doi.org/10.1101/2020.11.04.355842.

(37) Kemp, S. A.; Meng, B.; Ferriera, I. A. T. M.; Datir, R.; Harvey, W. T.; Papa, G.; Lytras, S.; Collier, D. A.; Mohamed, A.; Gallo, G.; Thakur, N.; Carabelli, A. M.; Kenyon, J. C.; Lever, A. M.; de Marco, A.; Saliba, C.; Culap, K.; Cameroni, E.; Piccoli, L.; Corti, D.; James, L. C.; Bailey, D.; Robertson, D. L.; Gupta, R. K. Recurrent Emergence and Transmission of a SARS-CoV-2 Spike Deletion H69/V70. *bioRxiv* **2021**, 2020.12.14.422555. https://doi.org/10.1101/2020.12.14.422555.

(38) Bal, A.; Destras, G.; Gaymard, A.; Stefic, K.; Marlet, J.; Eymieux, S.; Regue, H.; Semanas, Q.; d'Aubarede, C.; Billaud, G.; Laurent, F.; Gonzalez, C.; Mekki, Y.; Valette, M.; Bouscambert, M.; Gaudy-Graffin, C.; Lina, B.; Morfin, F.; Josset, L.; the COVID-Diagnosis HCL Study Group. Two-Step Strategy for the Identification of SARS-CoV-2 Variant of Concern 202012/01 and Other Variants with Spike Deletion H69–V70, France, August to December 2020. *Eurosurveillance* **2021**, *26* (3). https://doi.org/https://doi.org/10.2807/1560-7917.ES.2021.26.3.2100008.

(39) Emma B. Hodcroft. CoVariants: SARS-CoV-2 Mutations and Variants of Interest https://covariants.org/.

(40) Pérez-Lago, L.; Sola Campoy, P. J.; Buenestado-Serrano, S.; Pharmacy, P. C.; Estévez, A.; de la Cueva Technician, V. M.; López, M. G.; Tecnichian, M. H.; Suárez-González, J.; Alcalá, L.; Comas, I.; González-Candelas, F.; Muñoz, P.; García de Viedma, D.; Group, on behalf of G. M. M.-I. C.-19 S. Epidemiological, Clinical and Genomic Snapshot of the First 100 B.1.1.7 SARS-CoV-2 Cases in Madrid. *Journal of Travel Medicine* **2021**. https://doi.org/10.1093/jtm/taab044.

(41) ECDC. *SARS-CoV-2 – Increased Circulation of Variants of Concern and Vaccine Rollout in the EU/EEA, 14th Update – 15 February 2021.*; 2021.

(42) (ISCIII), E. C.-19. RENAVE. CNE. C. *Informe Nº 77. Situación de COVID-19 En España. Casos Diagnosticados a Partir 10 de Mayo*; 2021.

703 **FIGURE LEGENDS**

704

705 **Figure 1**. Estimated genome copies (GC) corresponding to wild type SARS-CoV-2
706 sequences without ΔHV69/70 deletion (grey bars) and sequences containing ΔHV69/70
707 deletion in the S gene (yellow bars), from 9 preparations at 3 different total
708 concentration levels (**A**: $1x10^4$ GC/rxn, **B**: $1x10^3$ GC/rxn and **C**: $1x10^2$ GC/rxn), and 3
709 different proportions of Wuhan-Hu-1 and B.1.1.7 GC (90:10, 50:50, and 10:90). Data
710 correspond to mean values ± standard deviations from duplicate samples. Each sample
711 correspond to an independent preparation containing the indicated proportions of
712 B.1.1.7 and Wuhan-Hu-1 synthetic control RNAs. Samples at different proportions of
713 synthetic control RNAs were prepared in duplicate and were further diluted at the
714 indicated concentration levels.

715

716 **Figure 2.** Overview of the nucleotide substitutions detected in SARS-CoV-2 S gene
717 sequences from wastewater samples (n=8) as compared to the SARS-CoV-2 isolate
718 Wuhan-Hu-1 reference genome (MN908947.3). Percentages before each line indicate
719 the proportion of B.1.1.7 variant measured in each sample. B.1.1.7-specific markers are
720 shown in light orange; yellow markers show mutations described in B.1.777 variant,
721 and blue markers indicate others. RBD (Receptor Binding Domain) is indicated with a
722 dotted square. Amplicon numbers are shown at the bottom. Shaded green colors
723 indicate sequence coverage in logarithmic scale for each amplicon.

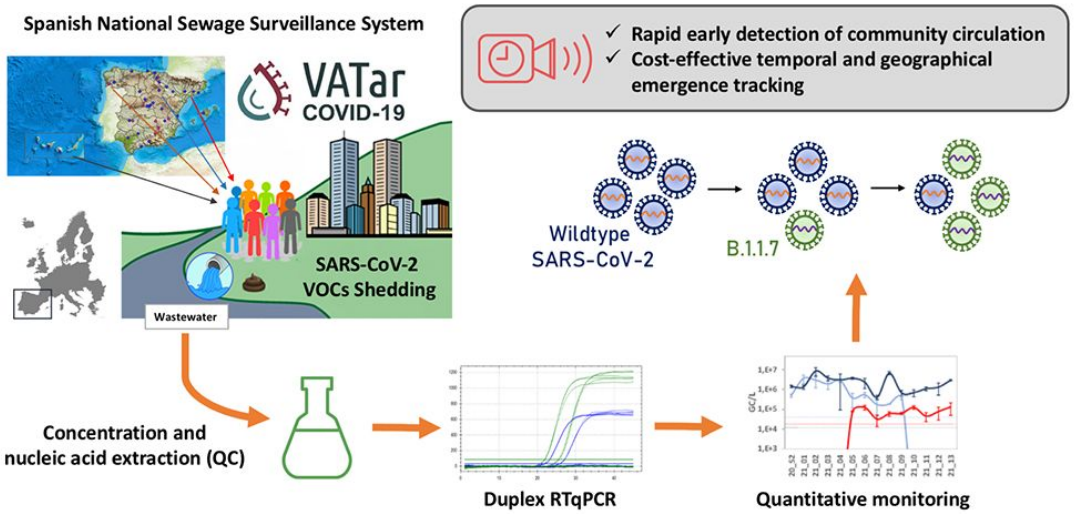
724

725 **Figure 3**. Concentration of SARS-CoV-2 RNA in wastewater samples collected in
726 Spain from December 2020 to March 2021, as measured by N1 RT-qPCR (dark blue),
727 and duplex S gene allelic discrimination RT-qPCR [wildtype S (light blue) and B.1.1.7

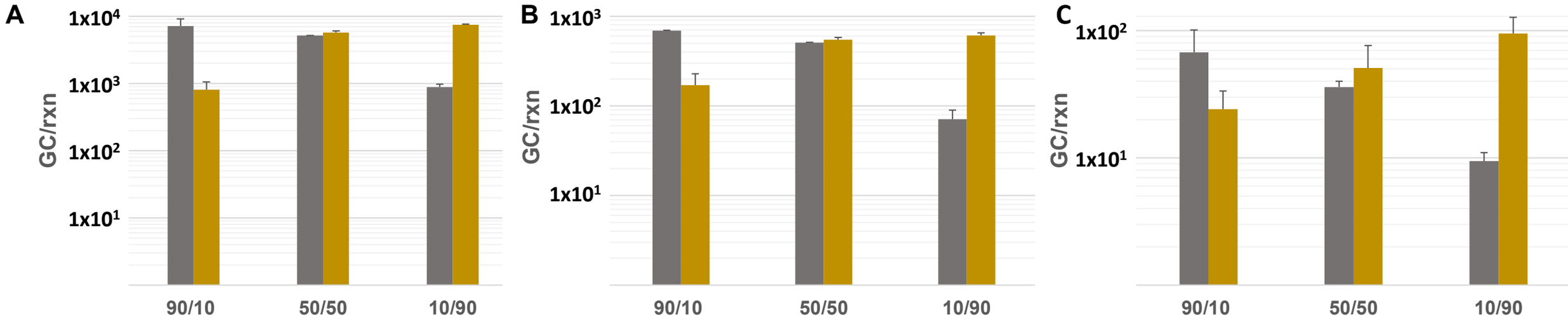
728 S (red)]. Wastewater treatment plants (WWTPs) are alphabetically grouped by
729 Autonomous Communities in Spain (**A**: Andalucía, **B**: Aragón, **C**: Baleares, **D**:
730 Canarias, **E**: Cantabria, **F**: Castilla-La Mancha, **G**: Castilla y León, **H**: Cataluña, **I**:
731 Com. De Madrid, **J**: Com. Valenciana, **K**: Extremadura, **L**: Galicia, **M**: La Rioja, **N**:
732 País Vasco, **O**: Pr. Asturias). Data represent average values and error bars standard
733 deviation of the RT-qPCR replicates used for calculation. Dotted lines correspond to the
734 limit of quantification of assays.
735
736 **Figure 4**. Evolution of B.1.1.7 SARS-CoV-2 prevalence over time, as measured by
737 duplex RT-qPCR in wastewater samples from 32 wastewater treatment plants
738 (WWTPs). As in Figure 3, data are alphabetically shown according to Autonomous
739 Community. * indicates samples with detection of a single variant, but with titers
740 <LOQ.
741
742 **Figure 5**. Comparison of B.1.1.7 estimates from wastewater testing and clinical
743 epidemiological surveillance. **(A)** Correlation between B.1.1.7 proportions estimated by
744 duplex RT-qPCR from wastewater and data reported by local authorities from clinical
745 specimens sequencing. **(B)** Geographic and temporal evolution of B.1.1.7 SARS-CoV-2
746 emergence in Spain during the study period, estimated from wastewater samples (left
747 panels) and reported in clinical data (right panels). For wastewater data, percentages are
748 indicated for each WWTP. * indicates samples with detection of a single variant, but
749 with titers <LOQ. For clinical data, percentages are indicated for each Autonomous
750 Community and number in parenthesis indicates the number of cases under sequence
751 study during that week. Communities in for which data were not available are depicted
752 colorless.

753

For Table of Contents Only



754



Wildtype	89.90±0.31	47.51±1.57	10.56±1.16
Δ69/70	10.10±0.31	52.49±1.57	89.44±1.16

80.52±5.25	48.04±1.68	10.32±1.89
19.48±5.25	51.96±1.68	89.68±1.89

73.09±2.41	43.22±10.16	9.27±1.51
26.91±2.41	56.78±10.16	90.73±1.51

Figure 1

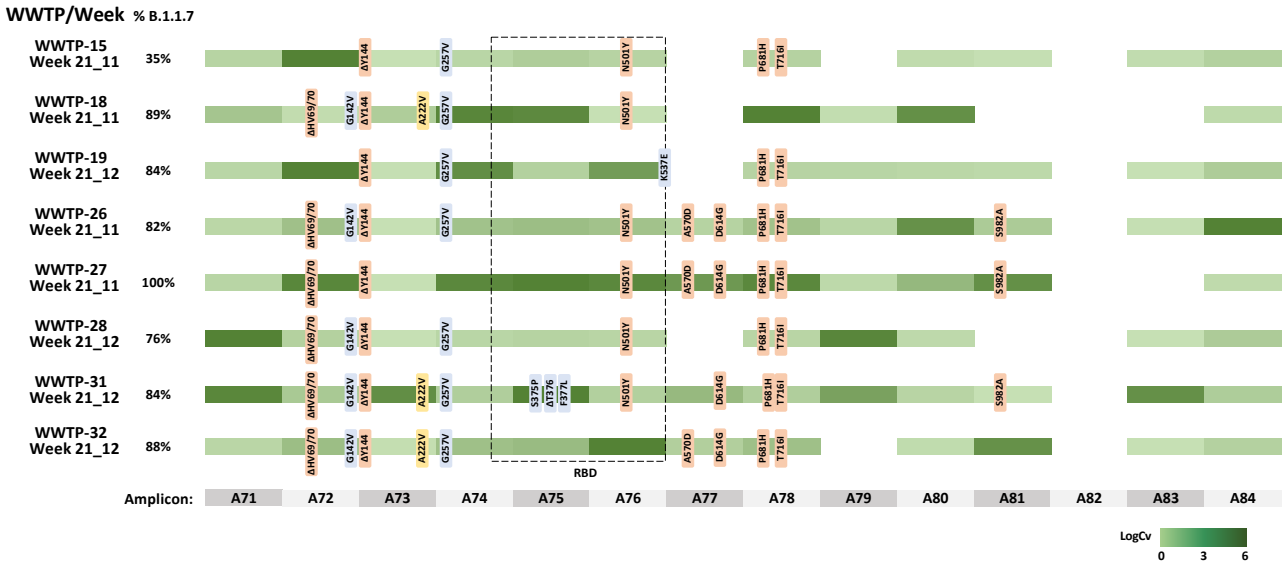

WWTP/Week
% B.1.1.7
WWTP-15 Week 21_11
35%
WWTP-18 Week 21_11
89%
WWTP-19 Week 21_12
84%
WWTP-26 Week 21_11
82%
WWTP-27 Week 21_11
100%
WWTP-28 Week 21_12
76%
WWTP-31 Week 21_12
84%
WWTP-32 Week 21_12
88%
ΔH69/70
G142V
ΔY144
A222V
G257V
S375P
ΔT376
F377L
N501Y
K537E
A570D
D614G
P681H
T716I
S982A
RBD
Amplicon:
A71
A72
A73
A74
A75
A76
A77
A78
A79
A80
A81
A82
A83
A84
LogCv
0
3
6

Figure 2

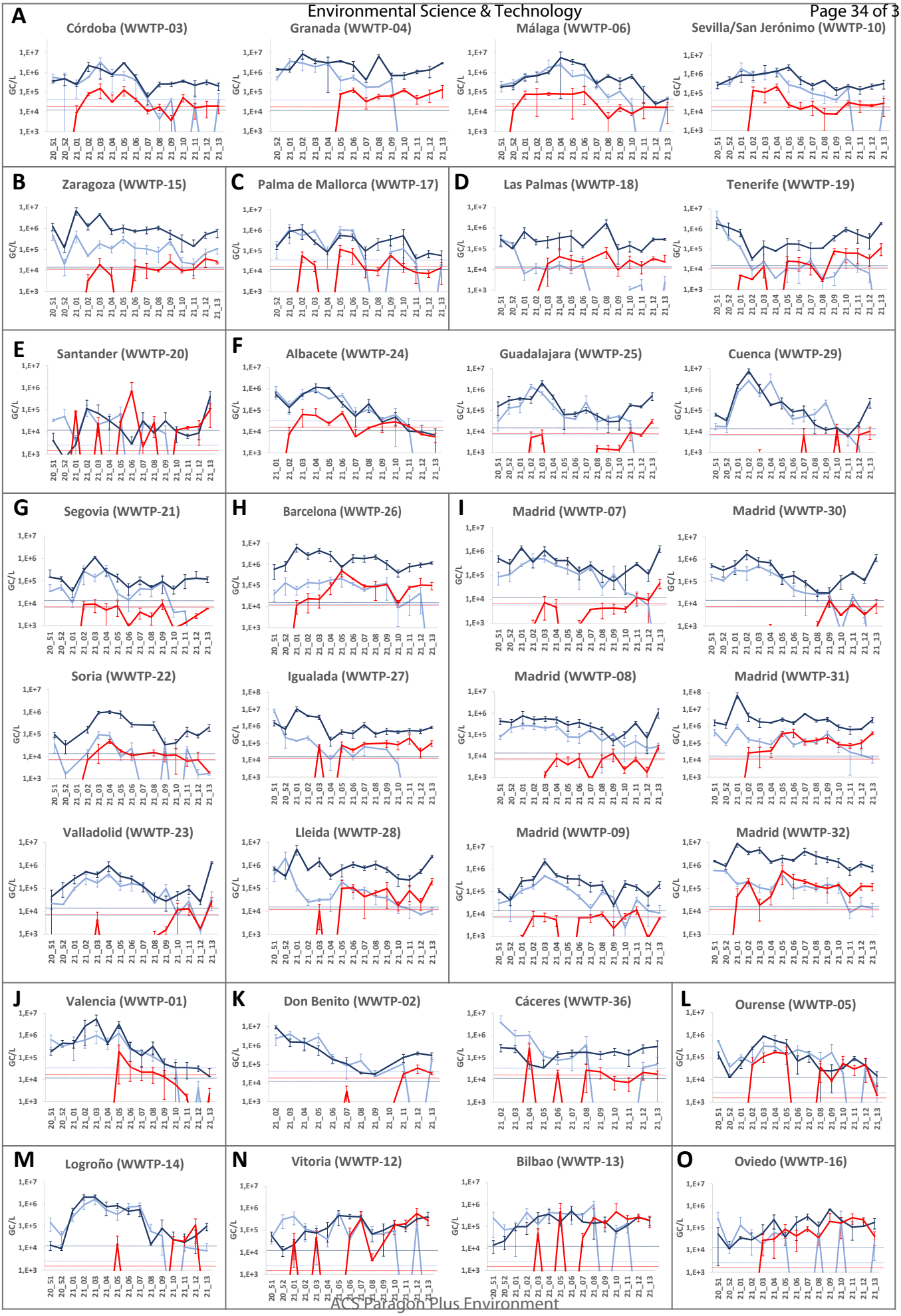

A
Córdoba (WWTP-03)
Granada (WWTP-04)
Málaga (WWTP-06)
Sevilla/San Jerónimo (WWTP-10)
B
Zaragoza (WWTP-15)
C
Palma de Mallorca (WWTP-17)
D
Las Palmas (WWTP-18)
Tenerife (WWTP-19)
E
Santander (WWTP-20)
F
Albacete (WWTP-24)
Guadalajara (WWTP-25)
Cuenca (WWTP-29)
G
Segovia (WWTP-21)
Soria (WWTP-22)
Valladolid (WWTP-23)
H
Barcelona (WWTP-26)
Igualada (WWTP-27)
Lleida (WWTP-28)
I
Madrid (WWTP-07)
Madrid (WWTP-30)
Madrid (WWTP-08)
Madrid (WWTP-31)
Madrid (WWTP-09)
Madrid (WWTP-32)
J
Valencia (WWTP-01)
K
Don Benito (WWTP-02)
Cáceres (WWTP-36)
L
Ourense (WWTP-05)
M
Logroño (WWTP-14)
N
Vitoria (WWTP-12)
Bilbao (WWTP-13)
O
Oviedo (WWTP-16)
GC/L

Figure 3

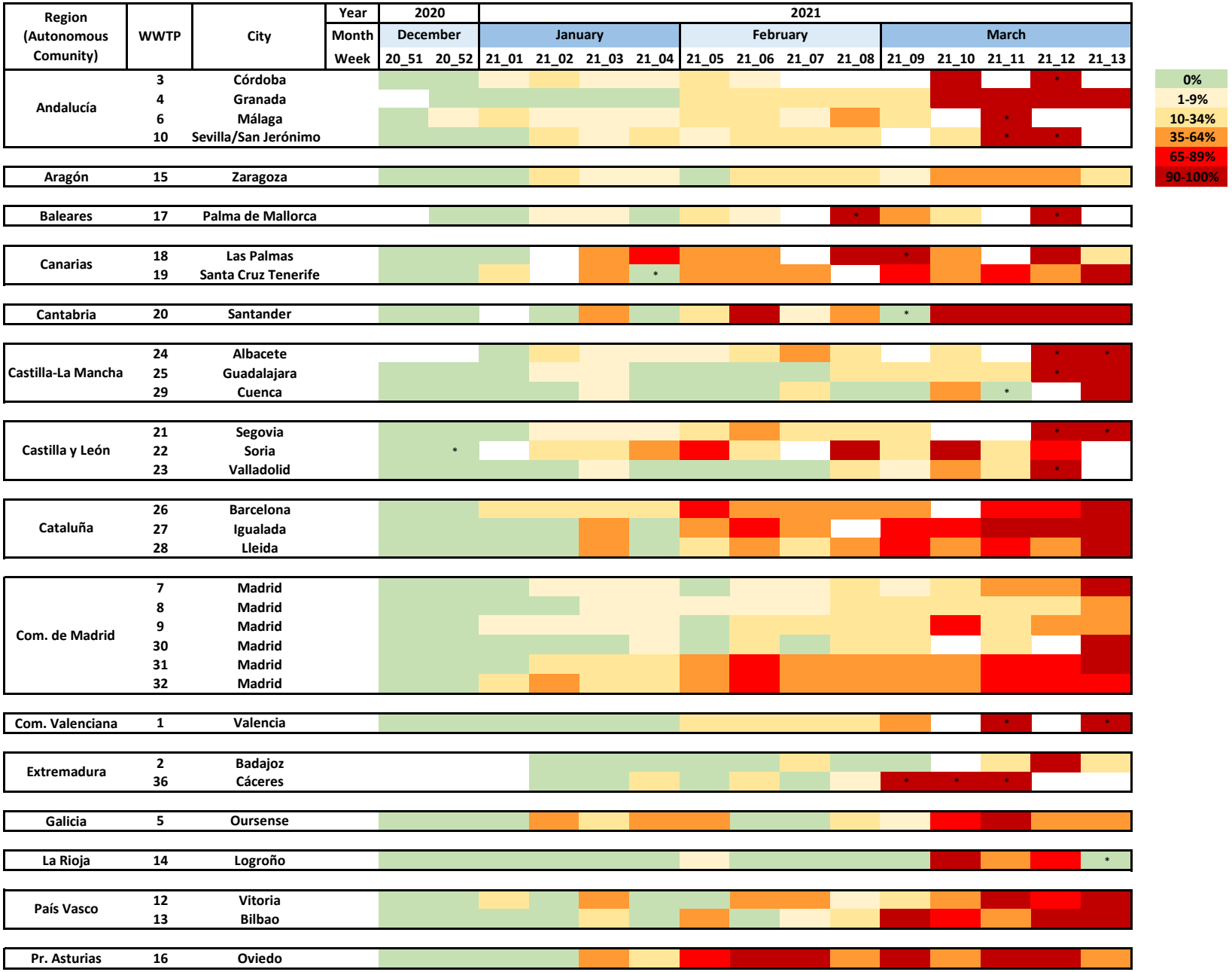


Figure 4

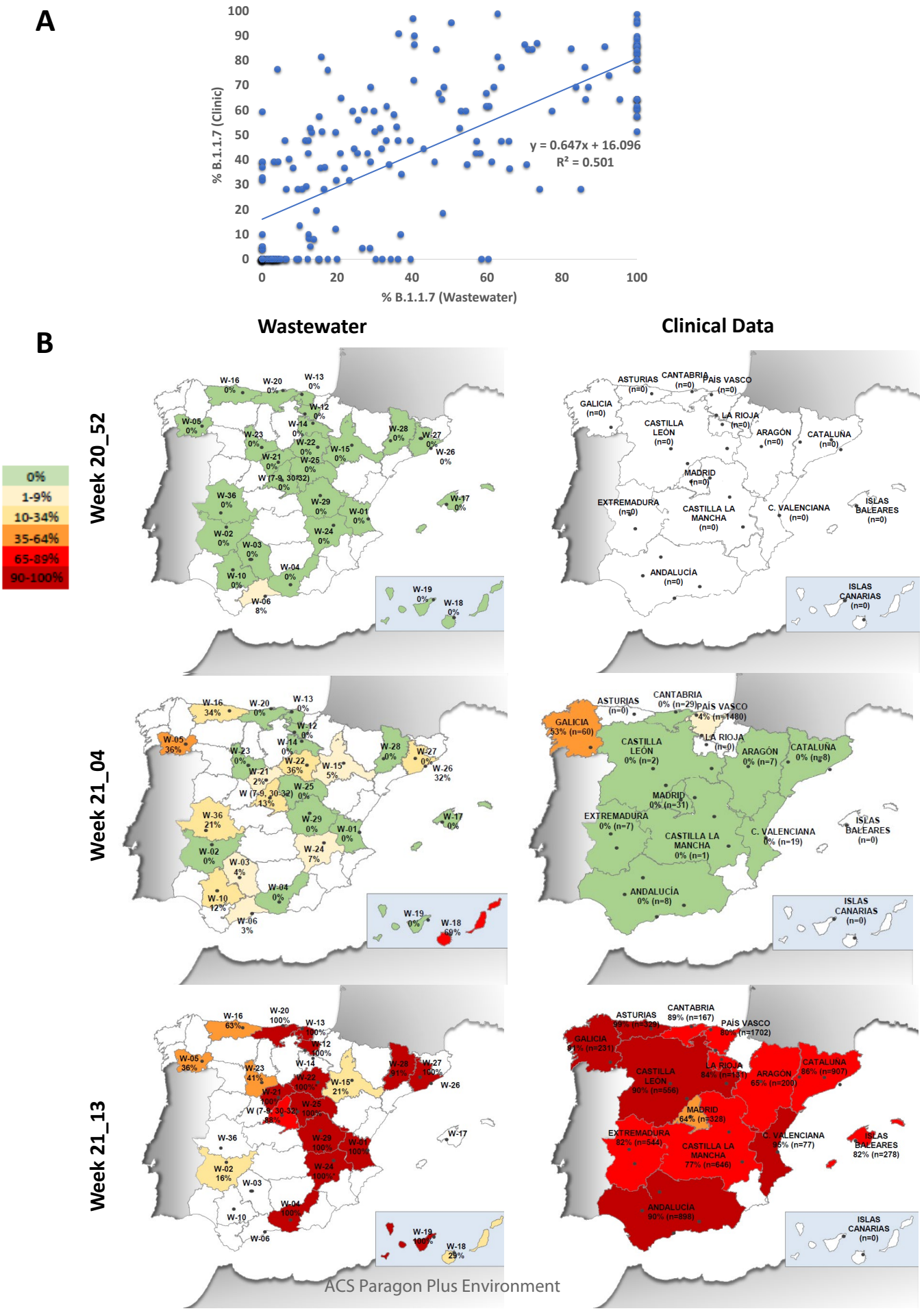

A
% B.1.1.7 (Clinic)
% B.1.1.7 (Wastewater)
y = 0.647x + 16.096
R² = 0.501
B
Wastewater
Clinical Data
Week 20_52
Week 21_04
Week 21_13
0%
1-9%
10-34%
35-64%
65-89%
90-100%
ASTURIAS (n=0)
CANTABRIA (n=0)
PAÍS VASCO (n=0)
GALICIA (n=0)
CASTILLA LEÓN (n=0)
LA RIOJA (n=0)
ARAGÓN (n=0)
CATALUÑA (n=0)
MADRID (n=0)
EXTREMADURA (n=0)
CASTILLA LA MANCHA (n=0)
C. VALENCIANA (n=0)
ISLAS BALEARES (n=0)
ANDALUCÍA (n=0)
ISLAS CANARIAS (n=0)
GALICIA 53% (n=60)
CANTABRIA 0% (n=29)
PAÍS VASCO 4% (n=1480)
CASTILLA LEÓN 0% (n=2)
LA RIOJA (n=0)
ARAGÓN 0% (n=7)
CATALUÑA 0% (n=8)
MADRID 0% (n=31)
EXTREMADURA 0% (n=7)
CASTILLA LA MANCHA 0% (n=1)
C. VALENCIANA 0% (n=19)
ANDALUCÍA 0% (n=8)
CANTABRIA 89% (n=167)
PAÍS VASCO 80% (n=1702)
GALICIA 91% (n=231)
CASTILLA LEÓN 90% (n=556)
LA RIOJA 84% (n=13)
ARAGÓN 65% (n=200)
CATALUÑA 86% (n=907)
MADRID 64% (n=328)
EXTREMADURA 82% (n=544)
CASTILLA LA MANCHA 77% (n=646)
C. VALENCIANA 95% (n=77)
ISLAS BALEARES 82% (n=278)
ANDALUCÍA 90% (n=898)

Figure 5