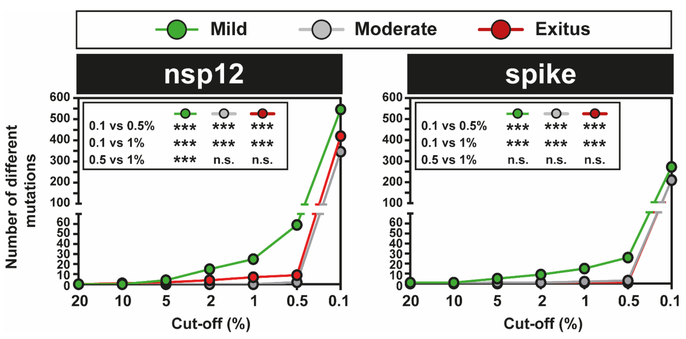
SARS-CoV-2 MUTANT SPECTRA REVEAL DIFFERENCES BETWEEN COVID-19 SEVERITY CATEGORIES

Brenda Martínez-González[1,2], María Eugenia Soria[2,3,4], Lucía Vázquez-Sirvent[2], Cristina Ferrer-Orta[5], Rebeca Lobo-Vega[2], Pablo Mínguez[1,6], Lorena de la Fuente[1,6], Carlos Llorens[7], Beatriz Soriano[7], Ricardo Ramos-Ruíz[8], Marta Cortón[1,6], Rosario López-Rodríguez[1,6], Carlos García-Crespo[3,4], Pilar Somovilla[3,9], Antoni Durán-Pastor[3], Isabel Gallego[3,4], Ana Isabel de Ávila[3,4], Soledad Delgado[10], Federico Morán[11], Cecilio López-Galíndez[12], Jordi Gómez[4,13], Luis Enjuanes[2], Llanos Salar-Vidal[2], Mario Esteban-Muñoz[2], Jaime Esteban[2], Ricardo Fernandez-Roblas[2], Ignacio Gadea[2], Carmen Ayuso[1,6], Javier Ruíz-Hornillos[14], Nuria Verdaguer[5], Esteban Domingo[3,4] and Celia Perales[1,2,4]

[1]Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM) Av. Reyes Católicos 2, 28040 Madrid, Spain. [2]Centro Nacional de Biotecnología (CNB-CSIC), Consejo Superior de Investigaciones Científicas (CSIC), Campus de Cantoblanco, 28049 Madrid, Spain. [3]Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Consejo Superior de Investigaciones Científicas (CSIC), Campus de Cantoblanco, 28049 Madrid, Spain. [4]Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Instituto de Salud Carlos III, 28029 Madrid, Spain. [5]Institut de Biología Molecular de Barcelona CSIC, 08028 Barcelona, Spain. [6]Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain. [7]Biotechvana, "Scientific Park", Universidad de Valencia, 46980 Valencia, Spain. [8]Scientific Park of Madrid, Campus de Cantoblanco, 28049 Madrid, Spain. [9]Departamento de Biología Molecular, Universidad Autónoma de Madrid, Campus de Cantoblanco, 28049, Madrid, Spain. [10]Escuela Técnica Superior de Ingeniería de Sistemas Informáticos (ETSISI), Universidad Politécnica de Madrid, Madrid, Spain. [11]Departamento de Bioquímica y Biología Molecular, Universidad Complutense de Madrid, Madrid, Spain. [12]Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain. [13]Instituto de Parasitología y Biomedicina 'López-Neyra' (CSIC), Parque Tecnológico Ciencias de la Salud, Armilla, 18016 Granada, Spain. [14]Hospital Infanta Elena, Valdemoro, Madrid, Spain.

RNA virus populations are composed of complex mixtures of genomes that are termed mutant spectra. SARS-CoV-2 replicates as a viral quasispecies, and mutations that are detected at low frequencies in a host can be dominant in subsequent variants. We have studied mutant spectrum complexities of SARS-CoV-2 populations derived from thirty nasopharyngeal swabs of patients infected during the first wave (April 2020) in the Hospital Universitario Fundación Jiménez Díaz. The patients were classified according to the COVID-19 severity in mild (non-hospitalized), moderate (hospitalized) and exitus (hospitalized with ICU admission and who passed away due to COVID-19). Using ultra-deep sequencing technologies (MiSeq, Illumina), we have examined four amplicons of the nsp12 (polymerase)-coding region and two amplicons of the spike-coding region. Ultra-deep sequencing data were analyzed with different cut-off frequency for mutation detection. Average number of different point mutations, mutations per haplotype and several diversity indices were significantly higher in SARS-CoV-2 isolated from patients who developed mild disease. A feature that we noted in the SARS-CoV-2 mutant spectra from diagnostic samples is the remarkable absence of mutations at intermediate frequencies, and an overwhelming abundance of mutations at frequencies lower than 10%. Thus, the decrease of the cut-off frequency for mutation detection from 0.5% to 0.1% revealed an increasement (50- to 100-fold) in the number of different mutations. The significantly higher frequency of mutations in virus from patients displaying mild than moderate or severe disease was maintained with the 0.1% cut- off frequency. To evaluate whether the frequency repertoire of amino acid substitutions differed between SARS-CoV-2 and the well characterized hepatitis C virus (HCV), we performed a comparative study of mutant spectra from infected patients using the same bioinformatics pipelines. HCV did not show the deficit of intermediate frequency substitutions that was observed with SARS-CoV-2. This difference was maintained when two functionally equivalent proteins, the corresponding viral polymerases, were compared. In conclusion, SARS-CoV-2 mutant spectra are rich reservoirs of mutants, whose complexity is not uniform among clinical isolates. Virus from patients who developed mild disease may be a source of new variants that may acquire epidemiological relevance.

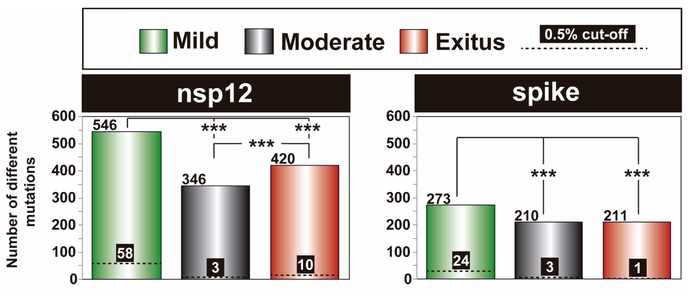
Overwhelming abundance of mutations a frequencies < 10%



The number of different mutations was significantly higher at 0.1% than at higher cut-off levels.

The higher number of mutations in virus from mild patients is maintained at different depth levels.

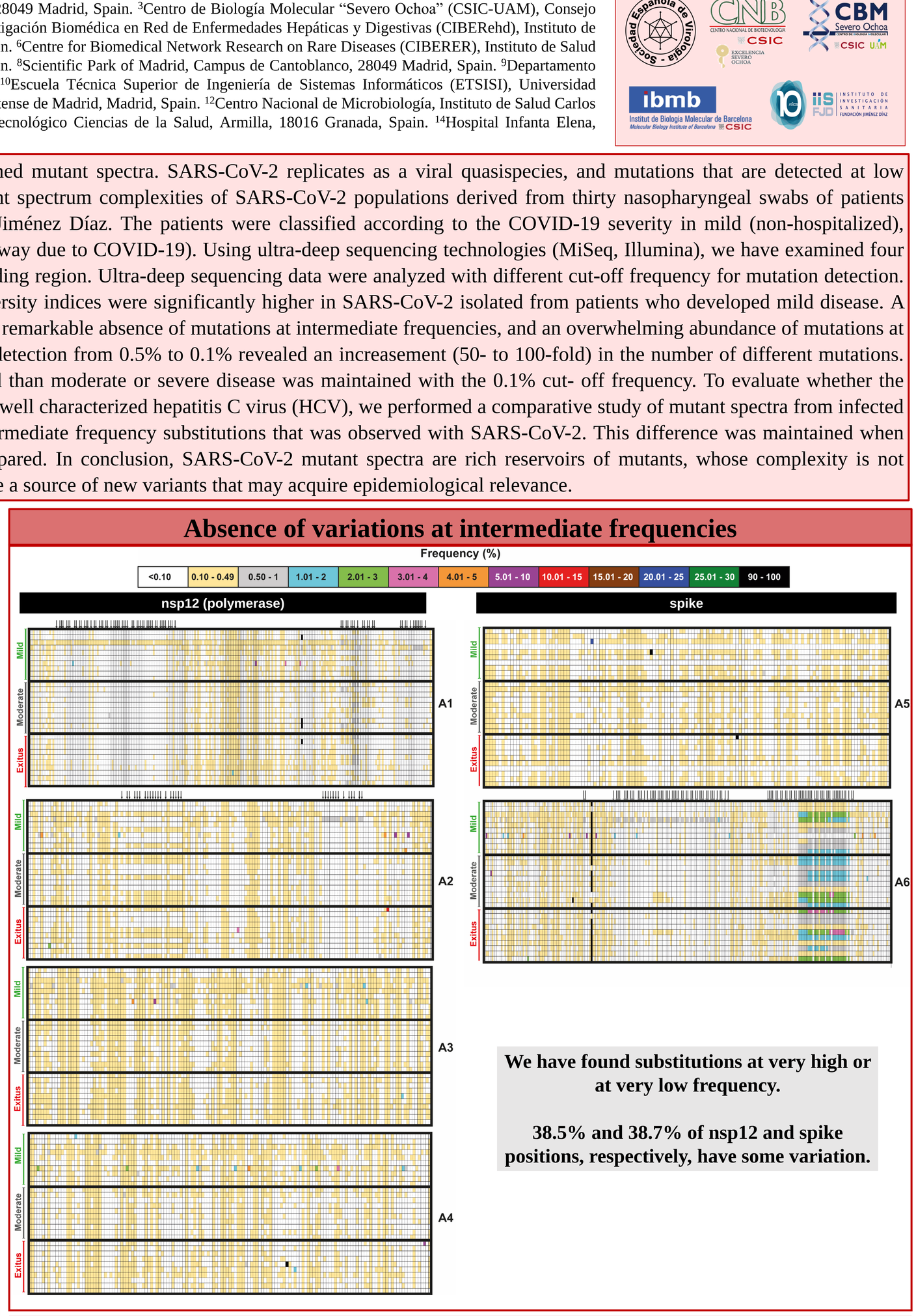
Higher frequency of mutations in virus from mild patients



A significantly higher number of different mutations was found in virus isolated from patients who developed mild COVID-19.

This tendency was observed at 0.5% cut-off level and at 0.1% cut-off level.

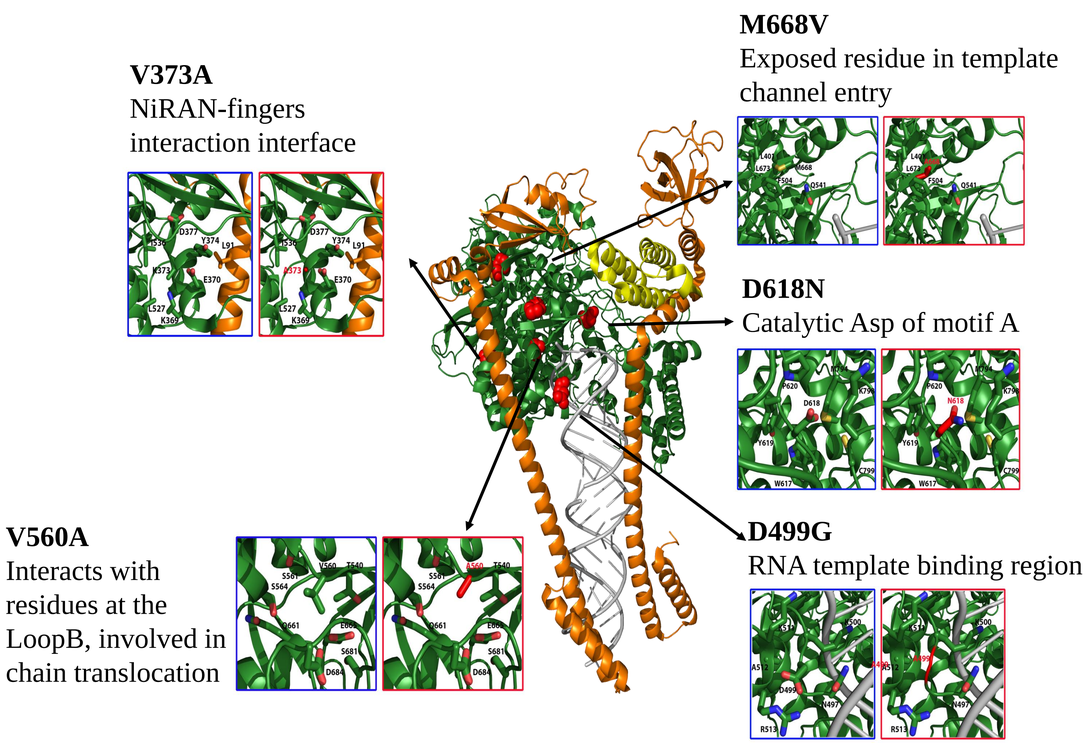
Absence of variations at intermediate frequencies



We have found substitutions at very high or at very low frequency.

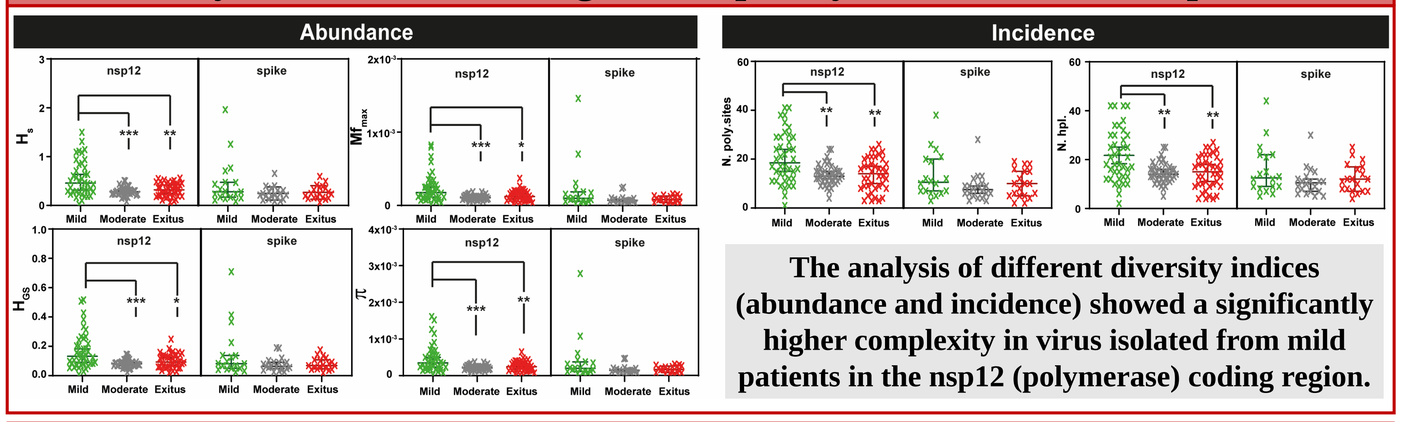
38.5% and 38.7% of nsp12 and spike positions, respectively, have some variation.

Prediction of possible structural and functional changes



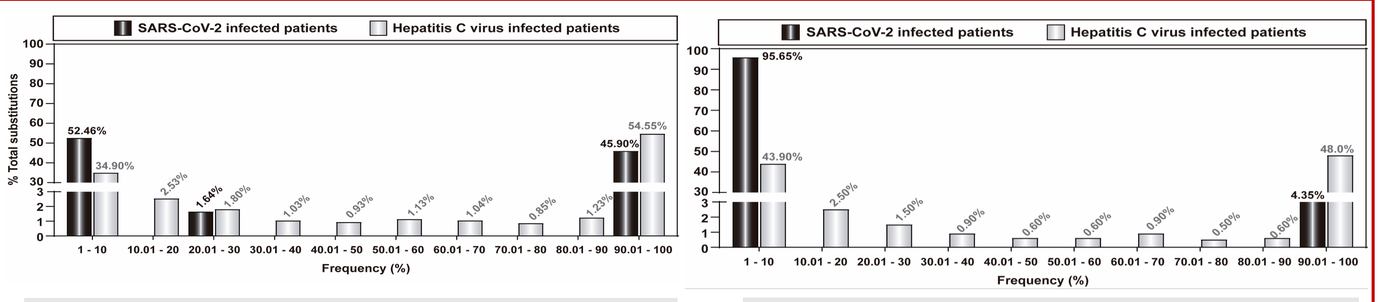
Some structural and functional changes in the nsp12 (polymerase) have been predicted

Diversity indices showed a higher complexity in virus from mild patients



The analysis of different diversity indices (abundance and incidence) showed a significantly higher complexity in virus isolated from mild patients in the nsp12 (polymerase) coding region.

Differences between SARS-CoV-2 and HCV



SARS-CoV-2 presents an absence of substitutions at intermediate frequencies compared to HCV.

The same tendency was observed comparing both polymerases.

- **A higher number of substitutions was observed in virus isolated from mild patients at different cut-off levels.**
- **This increased complexity of virus isolated from mild patients was confirmed studying different diversity indices.**
- **Some substitutions predicted structural and functional changes in the nsp12 protein.**
- **Most substitutions appeared at frequencies between 0.1% - 0.5%.**
- **A different distribution of mutant spectra was observed comparing SARS-CoV-2 with HCV.**

Martínez-González et al, Pathogens 2022 Jun 8;11(6):662
Martínez-González et al, Microbiol Spectr. 2022Apr 27;10(2):e0022122

This work was supported by Instituto de Salud Carlos III, Spanish Ministry of Science and In-novation (COVID-19 Research Call COV20/00181), and co-financed by European Development Regional Fund 'A way to achieve Europe'. The work was also supported by grants CSIC-COV19-014 from Consejo Superior de Investigaciones Científicas (CSIC), project 525/C/2021 from Fundació La Marató de TV3, PID2020-113888RB-I00 from Ministerio de Ciencia e Innovación, BFU2017-91384-EXP from Ministerio de Ciencia, Innovación y Universidades (MCIU), PI18/00210 and PI21/00139 from Instituto de Salud Carlos III, and S2018/BAA-4370 (PLATESA2 from Comu-nidad de Madrid/FEDER). C.P., M.C., and P.M. are supported by the Miguel Servet programme of the Instituto de Salud Carlos III (CPII19/00001, CPII17/00006, and CP16/00116, respectively) co-financed by the European Regional Development Fund (ERDF). CIBERehd (Centro de Investi-gación en Red de Enfermedades Hepáticas y Digestivas) is funded by Instituto de Salud Carlos III. Institutional grants from the Fundación Ramón Areces and Banco Santander to the CBMSO are also acknowledged. The team at CBMSO belongs to the Global Virus Network (GVN). B.M.-G. is supported by predoctoral contract PFIS FI19/00119 from Instituto de Salud Carlos III (Ministerio de Sanidad y Consumo) cofinanced by Fondo Social Europeo (FSE). R.L.-V. is supported by predoctoral contract PEJD-2019-PRE/BMD-16414 from Comunidad de Madrid. C.G.-C. is sup-ported by predoctoral contract PRE2018-083422 from MCIU. BS was supported by a predoctoral research fellowship (Doctorados Industriales, DI-17-09134) from Spanish MINECO.