

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active. Contents lists available at ScienceDirect



International Journal of Infectious Diseases



journal homepage: www.elsevier.com/locate/ijid

# Rapid and qualitative identification of SARS-CoV-2 mutations associated with variants of concern using a multiplex RT-PCR assay coupled with melting analysis



Giuseppe Sberna<sup>1</sup>, Lavinia Fabeni<sup>1</sup>, Giulia Berno, Fabrizio Carletti, Eliana Specchiarello, Francesca Colavita, Silvia Meschi, Giulia Matusali, Anna Rosa Garbuglia, Licia Bordi<sup>\*</sup>, Eleonora Lalle

Laboratorio di Virologia, Istituto Nazionale per le Malattie Infettive (INMI) "Lazzaro Spallanzani" - IRCCS, Rome, Italy

## ARTICLE INFO

Article history: Received 24 March 2022 Revised 20 May 2022 Accepted 21 June 2022

Keywords: SARS-CoV-2 Variants of concern Rapid identification RT-PCR Melting analysis

# ABSTRACT

*Objectives:* Considering the spread of new genetic variants and their impact on public health, it is important to have assays that are able to rapidly detect SARS-CoV-2 variants.

*Methods:* We retrospectively examined 118 positive nasopharyngeal swabs, first characterized by the Sanger sequencing, using the Simplexa® SARS-CoV-2 Variants Direct assay, with the aim of evaluating the performance of the assay to detect N501Y, G496S, Q498R, Y505H, E484K, E484Q, E484A, and L452R mutations.

*Results:* A total of 111/118 nasopharyngeal swabs were in complete agreement with the Sanger sequencing, whereas the remaining seven samples were not amplified due to the low viral load. The evaluation of the ability of the assay to detect the E484Q mutation was performed using a viral isolate of the SARS-CoV-2 Kappa variant, showing concordance in 15/15 samples. Simplexa® SARS-CoV-2 Variant Direct assay was able to detect mutation pattern of Alpha, Beta, Gamma, Delta, and Omicron variants with 100% specificity and 94% sensitivity, whereas 100% sensitivity and specificity for the Kappa variant was observed. *Conclusion:* The assay can be useful to obtain faster results, contributing to a prompt surveillance of SARS-CoV-2 variants; however, it requires to be confirmed by the Sanger method, especially in the case of pattern of mutations that are different from those expected and also requires updates as new variants emerge.

© 2022 The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

#### Introduction

COVID-19 has been declared a pandemic by the World Health Organization (WHO); more than 520 million cases of COVID-19 have been confirmed (WHO, https://covid19.who.int/, accessed on May 19, 2022) and new genetic variants of SARS-CoV-2 have emerged, spreading worldwide rapidly and becoming a global threat. During the last year, WHO classified five variants as variant of concern (VOC) due to their greater transmissibility or virulence: Alpha (UK, B.1.1.7), Beta (South Africa, B.1.351), Gamma (Brazil,

\* Corresponding author: Licia Bordi, Laboratory of Virology, National Institute of Infectious Diseases "L Spallanzani" – IRCCS, Via Portuense 292, 00149 Rome, Italy

E-mail address: licia.bordi@inmi.it (L. Bordi).

P1), Delta (India, B.1.617.2), and more recently, Omicron (Multiple countries, B.1.1.529) (WHO, https://www.who.int/en/activities/ tracking-SARS-CoV-2-variants/, accessed on May 19, 2022). All the other variants were classified (or reclassified) as variants of interest, variants under monitoring, and formerly monitored variants (Kappa [India, B.1.617.1]) (Chakraborty *et al.*, 2022; Fernandes *et al.*, 2022; Hoffmann *et al.*, 2021; Koyama *et al.*, 2020; Li *et al.*, 2021; Micheli *et al.*, 2022; Qin *et al.*, 2021; Reardon, 2021; Saxena *et al.*, 2022; Tao *et al.*, 2021).

Therefore, many states have decided to closely monitor changes in the SARS-CoV-2 genome and the consequences of those changes and to start a research screening to verify the current situation in their regional territories, including the Italian Lazio region. Sequencing of (partial) genes and whole genomes has been proven as gold standard for the identification of virus variants,

<sup>&</sup>lt;sup>1</sup> These authors have contributed equally to this work.

https://doi.org/10.1016/j.ijid.2022.06.032

<sup>1201-9712/© 2022</sup> The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

to understand outbreak transmission dynamics and spill-over events, and screen for mutations that potentially have an impact on transmissibility, pathogenicity, and/or countermeasures (e.g., diagnostics, antiviral drugs, and vaccines) (ECDC, https://www.ecdc. europa.eu/en, accessed on March 6, 2022). Although the cost of sequencing has significantly fallen over the past decades, sequencing still requires substantial investment in resources (financial, infrastructure, and human), and it is time-consuming (Thompson and Milos, 2011).

Considering the rapid spread of new genetic variants and their impact on public health, it is extremely important to have an alternative assay that is able to perform a prompt identification of mutations that are associated with circulating variants. The aim of the study was to evaluate the performance of the RUO Simplexa® SARS-CoV-2 Variants Direct assay from DiaSorin S.p.A. (Gerenzano, Italy) in detecting and differentiating N501Y, G496S, Q498R, Y505H, E484K, E484Q, E484A, and L452R mutations of the spike protein, using the Sanger sequencing as the reference method.

#### Materials and methods

## Clinical samples

Nasopharyngeal Swabs (NPSs) from 118 patients (68 males and 50 females) who were selected from routine clinical purposes between March 31, 2021, and February 11, 2022, at National Institute for Infectious Diseases "Lazzaro Spallanzani" - IRCCS (INMI) in Rome were first analyzed for the presence of SARS-CoV ribonucleic acid (RNA) by the Simplexa Direct assay, sequenced by the Sanger method, and retrospectively tested with the Simplexa SARS-CoV-2 Variant Direct assay. All samples were stored at  $-80^{\circ}$ C from their arrival in the laboratory.

#### Virus

The evaluation of the ability of the Simplexa® SARS-CoV-2 Variants Direct assay to detect the E484Q mutation, less represented in the population of Lazio region, was performed using a viral stock, named hCoV-19/Italy/LAZ-INMI-642isl/2021 (EVAg, https://www.european-virus-archive.com/virus/ sars-cov-2-strain-hcov-19italylaz-inmi-642isl2021-clade-g-lineageb16171-vum-former-voi-kappa, accessed on March 6, 2022). The sample came from a patient who was hospitalized at INMI, and it was isolated in the biosafety level 3 facility. The infectious titer, calculated by the Reed and Muench method (Reed LJ et al., 1938) on VeroE6 cells, was 1  $\times$   $10^{5.83}$  TCID\_{50}/ml. Viral stock was spiked into oral swab UTM matrix and several dilutions (from 1000 to 1 TCID<sub>50</sub>/ml) were tested in replicates using the Simplexa® SARS-CoV-2 Variants Direct. The dilution was chosen with the intent of being in a range of viral load similar to that observed in clinical sample (roughly between 10<sup>3</sup>-10<sup>6</sup> copies/ml).

#### Simplexa<sup>®</sup> SARS-CoV-2 Variants Direct assay

The run was performed using a Direct Amplification Disk, consisting of eight simultaneous reactions (one positive control and seven samples). For the amplification. we used 50  $\mu$ l of the mix per reaction, 50  $\mu$ l of positive control, and 50  $\mu$ l for each seven sample, loaded onto the LIAISON® MDX instrument (Dia-Sorin S.p.A., Gerenzano, Italy).

The software generates an amplicon melting temperature analysis and fluorescence melting curve pattern visualization to allow identification of the SARS-CoV-2 mutations N501Y, G496S, Q498R, Y505H, E484K, E484Q, E484A, and L452R, which are potentially associated with variants: Alpha, Beta, Gamma, Delta, Kappa, Omicron, and others (DiaSorin Molecular, https://molecular.diasorin.com/ us/kit/simplexa-sars-cov-2-variants-direct/, accessed on March 6, 2022) (Table 1).

## Simplexa<sup>™</sup> SARS-CoV-2 Direct assay

All 118 NPS were routinely tested for the presence of SARS-CoV-2 RNA using the Simplexa<sup>TM</sup> SARS-CoV-2 Direct assay.

The run was performed by using a Direct Amplification Disk, consisting of eight simultaneous sample reactions. For the amplification, we used 50  $\mu$ l of the mix per reaction and 50  $\mu$ l for each sample, loaded onto the LIAISON® MDX instrument (DiaSorin S.p.A., Gerenzano, Italy).

Samples with cycle threshold (Ct) values <40 were considered positive; for statistical calculations, an arbitrary value of 45 Ct was assigned to negative samples.

#### Sanger sequencing

Viral RNA was extracted from NPS by the automated extraction system QiaSymphony (Qiagen Instruments AG Switzerland, Hilden, Germany), according to the manufacturer's instructions. The sequencing strategy was adjusted over time on the basis of the appearance of different variants characterized by specific pattern of mutations. Three set of primers have been in-house designed, which are able to amplify specific regions of SARS-CoV-2 gene S, whose sequences allowed the typing of the variants currently circulating. Primer and amplifications are here described for the first time: M6970-FW 5'TGACAAAGTTTTCAGATCCTCAGT3', M6970-RW 5' GGTCCATAAGAAAAGGCTGAGA 3' (aa. coverage: 47-171), VAR1-L FW 5' TCTCTGCTTTACTAATGTCTATGCAGA 3', VAR1-L RW 5' AACAGGGACTTCTGTGCAGT 3' (aa coverage: 399-616), VAR2 FW 5' GGTTTAACAGGCACAGGTGT 3', VAR2 RW 5' GACACTGGTA-GAATTTCTGTGGTA 3' (aa coverage: 552-722).

Reverse transcription-polymerase chain reaction (RT-PCR) conditions for M6970 primers were 52°C for 30 minutes, 95°C for 15 minutes, 45 cycles at 94°C for 35 seconds, 57°C for 35 seconds, and 72°C for 1 minute, and a final extension at 72°C for 10 minutes. RT-PCR conditions for VAR1-L and VAR2 primers were 52°C for 30 minutes, 95°C for 15 minutes, 45 cycles at 94°C for 40 seconds, 60°C for 50 seconds, and 72°C for 1 minute, and a final extension at 72°C for 10 minutes. Amplification products were sequenced bidirectionally by the 3130 or 3500XL Genetic analyzer with BigDye Terminator v3.1 chemistry.

Only the amplicon VAR1-L was used for comparisons with the Simplexa® SARS-CoV-2 Variants Direct assay because all mutations detected by Simplexa fall within the region amplified by those set of primers.

Sequences obtained were compared, by alignment, with the original Wuhan virus sequence (accession number: NC\_045512.2). Lineage and clade data, quality controls, mutations, missing nucleotides, and gaps were analyzed and confirmed with the support of Pangolin (https://pangolin.cog-uk.io/ accessed on March 6, 2022) and Nextclade (https://clades.nextstrain.org/ accessed on March 6, 2022).

#### Statistical analysis

Specificity and sensibility were calculated using the MedCalc statistical software (MedCalc, https://www.medcalc.org/, accessed on March 6, 2022) (MedCalc Software Ltd., Ostend, Belgium).

## Results

A total of 118 samples collected from March 31, 2021, and February 11, 2022, were routinely tested for the presence of SARS-CoV-2 RNA using the Simplexa<sup>TM</sup> SARS-CoV-2 Direct assay

#### Table 1

Interpretation of VOCs on the basis of the detection of mutations.

WHO Nomenclature	Alpha	Beta	Gamma	Delta	Kappa	Omicron
N501Y, G496S, Q498R. Y505H	Only N501Y	Only N501Y	Only N501Y	Absent	Absent	Presents
E484K	Absent	Present	Present	Absent	Absent	Absent
E484Q	Absent	Absent	Absent	Absent	Present	Absent
E484A	Absent	Absent	Absent	Absent	Absent	Present
L452R	Absent	Absent	Absent	Present	Present	Absent

#### Table 2

Mutations associated with the main VOCs detected by Simplexa® SARS-CoV-2 Variants Direct assay and Sanger sequencing on SARS-CoV-2 Spike region.

	Sanger sequencing	Sanger sequencing and Simplexa® SARS-CoV-2 Variants Direct
Alpha	Del69-70, Del144, A570D, D614G, P681H, T716I	N501Y
Beta	D80A, K417N, D614G	E484K, N501Y
Gamma	D138Y, K417T, D614G, H655Y	E484K, N501Y
Delta	Del157-158, T478K, D614G, P681R	L452R
Карра	T95I, G142D, D614G, P681R	L452R, E484Q
Omicron	A67V, Del69-70, T95I, G142D, Del143-145, K417N, N440K, G446S, S477N, T478K, Q493R,	E484A, G496S, Q498R, N501Y, Y505H
	T547K, D614G, H655Y, N679K, P681H	

(Bordi et al., 2020), obtaining median Ct values of 22.4 (Ct minimum-maximum (min-max): 12.4-29.3) in S and 22.6 (Ct minmax: 13.5-29.7) in ORF1ab genes. Sequences obtained by the Sanger method from these samples were compared with the original Wuhan virus sequence (accession number: NC\_045512.2). The specific pattern of mutations observed (Table 2) identified the following VOCs: 18 Alpha (B.1.1.7), 2 Beta (B.1.351), 18 Gamma (P.1), 33 Delta (B.1.617.2), and 40 Omicron (B.1.1.529). The same NPS were tested by the Simplexa® SARS-CoV-2 Variants Direct assay that was able to detect 111/118 samples: 18 NPS showed the N501Y mutation (referable to Alpha variant), 20 NPS showed the pattern N501Y+E484K (referable to Beta/Gamma variants), 33 NPS showed only the L452R mutation (referable to Delta variant), and 40 NPS showed the pattern N501Y, G496S, Q498R, and Y505H + E484A (referable to Omicron variant), in agreement with the Sanger results. Table 2 summarizes the mutations considered by sequencing and those common with the Simplexa Direct Variant assay. The 7 NPS samples that were not amplified by the Simplexa® SARS-CoV-2 Variants Direct assay were retested with the Simplexa<sup>TM</sup> SARS-CoV-2 Direct assay to verify the presence or absence of SARS-CoV-2 RNA after thawing. Notably, these samples resulted positive, although with high Ct: median Ct was 32.5 (Ct min-max: 29.2-34.3) in S and 36.3 (Ct min-max: 32.0-45.0) in ORF1ab genes, corresponding to low viral load. Overall, the Simplexa® SARS-CoV-2 Variant Direct assay was able to detect mutation pattern characteristic of Alpha, Beta, Gamma, Delta, and Omicron variants in clinical samples with 100% specificity and 94% sensitivity (95% confidence interval 88.2-97.6). Due to the absence of clinical samples that were positive for SARS-CoV-2 Kappa variant, the ability of the Simplexa® SARS-CoV-2 Variants Direct assay to detect E484Q mutation was established using oral swab UTM matrix, spiked with hCoV-19/Italy/LAZ-INMI-642isl/2021 viral isolate. Replicates of analyzed dilution showed the presence of E484Q mutation in 100% of cases (15/15), showing 100% of sensitivity and specificity.

#### Discussion

COVID-19 has wreaked havoc across the world and has overwhelmed many healthcare systems and economies of many countries. The emergence and spreading of SARS-CoV-2 VOC with the potential for increased transmission, disease severity, and resistance to treatment and vaccines call for fast and accurate monitoring of the population (Rubin, 2021). Sequencing is the reference method used to monitor the diffusion of emerging SARS-CoV-2 variants using the Sanger or NGS method; nevertheless, these techniques require substantial investment in resources (financial, infrastructure, and human) and a relatively long turnaround time (De Pace *et al.*, 2022). A prompt and widespread epidemiologic surveillance of COVID-19 is not feasible using only these methods.

Our results demonstrate that the Simplexa® SARS-CoV-2 Variants Direct assay was able to detect mutation pattern characteristic of Alpha, Beta, Gamma, Delta, and also Omicron variants in 111/118 NPS with 100% specificity and 94% sensibility compared with the Sanger sequencing. Moreover, 100% sensitivity and specificity for the Kappa variant was observed using swab UTM matrix spiked with hCoV-19/Italy/LAZ-INMI-642isl/2021 viral isolate, showing the presence of E484Q in 15/15 samples. Similar results were obtained by De Pace et al. (2022) in a comparative analysis of different multiplex RT-PCR assays, despite mainly limited on Alpha and Delta variants due to the low incidence of the other variants described in their study. Notably, we were able to analyze groups of all variants (Alpha, Beta, Gamma, Delta, and Kappa) with comparable size and we had the opportunity to extend the analysis also to the Omicron variant, which was not evaluated in the aforementioned study (De Pace et al., 2022).

Further analysis highlighted that the absence of amplification with the Simplexa® SARS-CoV-2 Variants Direct assay for seven samples was due to the low viral load, indicating that a possible limit of the system is that for Ct  $\geq$ 30, the success of the amplification should not be guaranteed. However, it is already known that the main factor of discrepant results between sequencing and screening tests for SARS-CoV-2 variants is a low viral load (Migueres *et al.*, 2021); in fact, in the study of De Pace *et al.*, they only tested samples with a high viral load (mean Ct value of 23.6  $\pm$  3.8) (De Pace *et al.*, 2022).

Another important consideration is that the assay is able to identify the presence or absence of these specific mutations but does not allow to uniquely classify all viral strains because some of the mutations are common to several variants. For instance, the presence of N501Y + E484K mutations is common to Beta and Gamma variants; in this situation, knowledge of variants circulating in a given period can provide an indication, but sequencing becomes mandatory. On the other hand, the Simplexa® SARS-CoV-2 Variants Direct assay is fast, easy to use, does not need extra equipment, such as centrifuges or an extraction system, and can be handled by laboratory personnel with little extra training for procedures and interpretation of results. Therefore, the assay represents a simple, accurate, and fast method, enabling high speed detection in just 107 minutes.

Overall, our results indicate that the Simplexa® SARS-CoV-2 Variants Direct assay can be helpful to obtain faster results, contributing to prompt surveillance of SARS-CoV-2 variants and give preliminary indication that is useful to the patient's management; nevertheless, the assay requires to be confirmed by the Sanger method, especially in the case of patterns of mutations that are different from those expected, and also requires continuous updates as new variants emerge.

## Funding

This research was supported by funds to the National Institute for Infectious Diseases "Lazzaro Spallanzani" IRCCS from Ministero della Salute (Ricerca Corrente, linea 1; COVID-2020-12371817), the European Commission – Horizon 2020 (EU project 101003544 – CoNVat; EU project 101003551 – EXSCALATE4CoV; EU project 101005111 – DECISION; EU project 101005075 – KRONO), and the European Virus Archive – GLOBAL (grants no. 871029).

## Author contributions

G.S., L.B.: conceptualization, analysis of results, writing, laboratory testing; G.B., L.F.: analysis of results, laboratory testing, writing; E.L, F.C., E.S.: analysis of results, laboratory testing; F.C., G.M., S.M.: provided materials, laboratory testing; A.R.G.: review and editing.

#### **Institutional Review Board Statement**

The study was conducted in accordance with the Declaration of Helsinki and with the protocol code no. 70, approved on December 17, 2018, by the institutional review board of the National Institute for Infectious Diseases, L. Spallanzani, IRCCS, according to which the study protocol described here did not provide for the signing of an informed consent by the patients because no further samples were taken other than those performed for diagnostic purposes. The data of biological samples collected for diagnostic purposes were used only after their complete anonymization and the results of the tests had no impact on the clinical management of the patients. Furthermore, the analysis of genetic data was not provided.

# **Conflicts of interest**

The authors have no competing interests to declare.

#### References

- Bordi L, Piralla A, Lalle E, Giardina F, Colavita F, Tallarita M, Sberna G, Novazzi F, Meschi S, Castilletti C, Brisci A, Minnucci G, Tettamanzi V, Baldanti F, Capobianchi MR. Rapid and sensitive detection of SARS-CoV-2 RNA using the Simplexa<sup>™</sup> COVID-19 direct assay. J Clin Virol 2020;128.
- Chakraborty C, Bhattacharya M, Sharma AR. Present variants of concern and variants of interest of severe acute respiratory syndrome coronavirus 2: their significant mutations in S-glycoprotein, infectivity, re-infectivity, immune escape and vaccines activity. Rev Med Virol 2022;32:e2270.
- De Pace V, Bruzzone B, Orsi A, Ricucci V, Domnich A, Guarona G, Randazzo N, Stefanelli F, Battolla E, Dusi PA, Lillo F, Icardi G. Comparative analysis of five multiplex RT-PCR assays in the screening of SARS-CoV-2 variants. Microorganisms 2022;10:306.
- Fernandes Q. Inchakalody VP, Merhi M, Mestiri S, Taib N, El-Ella MA D, Bedhiafi T, Raza A, Al-Zaidan L, Mohsen MO, Yousuf Al-Nesf MA, Hssain AA, Yassine HM, Bachmann MF, Uddin S, Dermime S. Emerging COVID-19 variants and their impact on SARS-CoV-2 diagnosis, therapeutics and vaccines. Ann Med 2022;54(1):524-40 Dec.
- Hoffmann M, Arora P, Groß R, Seidel A, Hörnich BF, Hahn AS, Krüger N, Graichen L, Hofmann-Winkler H, Kempf A, Winkler MS, Schulz S, Jäck HM, Jahrsdörfer B, Schrezenmeier H, Müller M, Kleger A, Münch J, Pöhlmann S. SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies. Cell 2021;184:2384–93 e12.
- Koyama T, Platt D, Parida L. Variant analysis of SARS-CoV-2 genomes. Bull World Health Organ 2020;98:495–504.
- Li G, Zhou Z, Du P, Yu M, Li N, Xiong X, Huang H, Liu Z, Dai Q, Zhu J, Guo C, Wu S, Baptista-Hon DT, Miao M, Ming LW, Wu Y, Zeng F, Zhang CL, Zhang ED, Song H, Liu J, Lau JY, Xiang AP, Zhang K. The SARS-CoV-2 spike L452R-E484Q variant in the Indian B.1617 strain showed significant reduction in the neutralization activity of immune sera. Precis Clin Med 2021;4:149–54.
- Micheli V, Bracchitta F, Rizzo A, Mancon A, Mileto D, Lombardi A, Stefanelli P, Gismondo MR. First identification of the new SARS-CoV-2 Omicron variant (B.1.1.529) in Italy. Clin Infect Dis 2022:ciab1044.
- Migueres M, Lhomme S, Trémeaux P, Dimeglio C, Ranger N, Latour J, Dubois M, Nicot F, Miedouge M, Mansuy JM, Izopet J. Evaluation of two RT-PCR screening assays for identifying SARS-CoV-2 variants. J Clin Virol 2021;143.
- Qin S, Cui M, Sun S, Zhou J, Du Z, Cui Y, Fan H. Genome characterization and potential risk assessment of the novel SARS-CoV-2 variant omicron (B.1.1.529). Zoonoses:1.
- Reardon S. How the Delta variant achieves its ultrafast spread. Nature 2021.
- Reed LJ, Muench H. A simple method of estimating fifty per cent endpoints12. Am J Epidemiol 1938;27:493–7.
- Rubin R. COVID-19 vaccines vs variants-determining how much immunity is enough. JAMA 2021;325:1241–3.
- Saxena SK, Kumar S, Ansari S, Paweska JT, Maurya VK, Tripathi AK. Abdel-Moneim AS. Characterization of the novel SARS-CoV-2 Omicron (B.1.1.529) variant of concern and its global perspective. J Med Virol 2022;94:1738–44.
- Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, Fera D, Shafer RW. The biological and clinical significance of emerging SARS-CoV-2 variants. Nat Rev Genet 2021;22:757–73.
- Thompson JF, Milos PM. The properties and applications of single-molecule DNA sequencing. Genome Biol 2011;12:217.