

Rubella virus nomenclature update: 2013

Virological surveillance data on rubella viruses are used to track progress towards the goal of eliminating rubella, to help with case classification, and to document transmission pathways. This report updates the existing documentation and incorporates the increase in rubella data generated by the WHO Measles and Rubella Laboratory Network (LabNet) in the past 5 years. The nomenclature to describe the genetic diversity of wild-type rubella viruses was proposed initially in 2005,¹ and was revised in 2007.²

The nomenclature for wild-type rubella viruses facilitates virological surveillance by defining standard methods for the genetic characterization and naming of viruses and organizes rubella virus sequences into recognized and provisional genotypes. The nomenclature system must be able to accommodate the discovery of new rubella genotypes, the evolution of known viruses, and periodic strain displacement. This flexibility is primarily obtained through the use of provisional genotype designations for virus groups that are small or diverse, for those that do not have the required reference viruses, or for those where the relationship with other genotypes is unclear. An update of the nomenclature is presented here; it describes the upgrading of 3 provisional genotypes, to give 12 recognized genotypes (1B, 1C, 1D, 1E, 1F, 1G, 1H, 1I, 1J, 2A, 2B, 2C) and retains only 1 provisional genotype (1a). The numbers refer to large distantly related groups of viruses designated as clade 1 and clade 2; letters represent genotypic groups within the clades

Nomenclature des virus rubéoleux – mise à jour 2013

Les données de la surveillance virologique des virus rubéoleux sont utilisées pour suivre les progrès réalisés en vue de l'objectif d'élimination de la rubéole, pour aider à la classification des cas et pour documenter les voies de transmission. Le présent rapport actualise la documentation existante et y intègre les données sur la rubéole obtenues au cours des 5 dernières années par le Réseau mondial OMS des laboratoires d'étude de la rougeole et de la rubéole (LabNet). La nomenclature pour décrire la diversité génétique des virus rubéoleux de type sauvage a été proposée à l'origine en 2005,¹ puis révisée en 2007.²

La nomenclature des virus rubéoleux de type sauvage facilite la surveillance virologique en définissant des méthodes standardisées de caractérisation génétique et de dénomination de ces virus et en classant les séquences virales en génotypes reconnus ou provisoires. Le système de nomenclature doit être en mesure d'intégrer la découverte de nouveaux génotypes, l'évolution des virus rubéoleux connus et le remplacement périodique des souches. Cette flexibilité est avant tout obtenue en recourant à des appellations provisoires des génotypes pour des groupes de virus petits mais très divers, pour ceux qui ne possèdent pas les virus de référence nécessaires ou pour ceux dont les rapports avec les autres génotypes ne sont pas très clairs. Nous présentons ici une actualisation de la nomenclature; elle décrit l'élévation de 3 des génotypes provisoires au rang de génotypes établis, ce qui en donne 12 au total (1B, 1C, 1D, 1E, 1F, 1G, 1H, 1I, 1J, 2A, 2B, 2C), et ne retient qu'un génotype provisoire (1a). Les chiffres font référence à des grands groupes de virus présentant une parenté lointaine et classés dans les clades 1 et 2; les lettres représentent les groupes

**WORLD HEALTH
ORGANIZATION**
Geneva

**ORGANISATION MONDIALE
DE LA SANTÉ**
Genève

Annual subscription / Abonnement annuel
Sw. fr. / Fr. s. 346.–

08.2013
ISSN 0049-8114
Printed in Switzerland

¹ See No. 14, 2005, pp. 125–132.

² See No. 24, 2007, pp. 216–222.

¹ See No. 14, 2005, pp. 125–132.

² Voir N° 24, 2007, pp. 216–222.

with lower case letters representing provisional genotypes. This update of the nomenclature provides a summary of the classification of rubella viruses including those that have been identified and analyzed since the last report in 2007.

This update retains basically the original system for naming viruses but includes some modifications to make the system consistent with other viruses surveyed in LabNet.

Information about the global distribution of wild-type rubella viruses, covering the years 2005–2010, has been published.³

Current and past genotypes

Provisional genotypes 1h, 1i, and 1j were established in 2007.² All 3 genotypes are now upgraded to accepted genotypes 1H, 1I, and 1J, since the criteria for becoming a recognized genotype have been fulfilled for all 3.¹ Briefly, the sequences of the full-length structural protein open reading frame (3192 nucleotides) from 2 reference viruses for each genotype are available (*Table 1*) and these viruses have been deposited in the WHO virus strain bank, maintained at the US Centers for Disease Control and Prevention (CDC) and Public Health England (PHE).

Four of the rubella genotypes, 1D, 1F, 1I, and 2A, have not been reported in circulation within the past 10 years: the last report for 1D was in 1996, for 1F in 2002, for 1I in 1994 and for wild-type 2A in 1980 (although 2A vaccine-derived viruses have been reported more recently). Thus, these genotypes are considered inactive and probably extinct.

Viruses of 4 genotypes (1E, 1G, 1J and 2B) have been reported most frequently, representing 71% of the sequences currently available for analysis. A dataset of 872 sequences was generated by requiring a unique 739 nucleotide sequence from viruses collected since 1961 for inclusion. There were sequences from only 30 viruses of genotype 1J in this dataset, which was insufficient for detailed analysis. A larger number of sequences from viruses of the other 3 genotypes were available (ranging from 75 for 1G to 292 for 1E) and these viruses were from widely distributed geographic locations. In addition, viruses of genotype 1G and 2B showed significant intra-genetic diversity (maximally 2.6% and 2.4%, respectively). As virological surveillance increases and more genetic data become available, it is clear that a description of circulating viruses more precise than the genotype will be necessary to track rubella viruses of genotypes 1E, 1G, and 2B.

génomiques à l'intérieur des clades, les lettres en minuscule indiquant des génotypes provisoires. Cette mise à jour fait une synthèse de la classification des virus rubéoleux en y intégrant ceux qui ont été identifiés et analysés depuis le dernier rapport en 2007.

Fondamentalement, le système original de désignation des virus a été retenu, mais en y incluant certaines modifications pour le rendre cohérent avec d'autres virus étudiés dans le réseau LabNet.

Des informations sur la répartition mondiale des virus rubéoleux de type sauvage, couvrant les années 2005–2010, ont été publiées.³

Génotypes actuels et du passé

Les génotypes provisoires 1h, 1i, et 1j ont été établis en 2007.² Les 3 ont été désormais élevés au rang de génotypes acceptés 1H, 1I, et 1J, les critères pour devenir un génotype reconnu ayant été remplis pour tous les 3.¹ En bref, on dispose des séquences correspondant à la totalité de la protéine de structure en cadre de lecture ouvert (3192 nucléotides) pour 2 virus de référence de chaque génotype (*Tableau 1*) qui ont été déposés dans la banque de souches de l'OMS, gardée par les *Centers for Disease Control and Prevention* (CDC) des États-Unis et *Public Health England* (PHE).

Pour 4 génotypes, 1D, 1F, 1I et 2A, aucune circulation n'a été signalée au cours des 10 dernières années, le dernier rapport remontant à 1996 pour 1D, 2002 pour 1F, 1994 pour 1I et 1980 pour le 2A de type sauvage (bien qu'on ait signalé plus récemment des virus 2A dérivés d'une souche vaccinale). On considère donc que ces génotypes sont inactifs et ont probablement disparu.

Ce sont les virus de 4 génotypes (1E, 1G, 1J et 2B) qui ont été signalés le plus fréquemment et ils représentent 71% des séquences actuellement disponibles pour l'analyse. Un ensemble de données comportant 872 séquences a été obtenu en demandant pour l'inclusion la même séquence de 739 nucléotides provenant des virus recueillis depuis 1961. Il n'y avait dans cet ensemble de données que des séquences de 30 virus de génotype 1J, ce qui est insuffisant pour une analyse détaillée. Pour les virus des 3 autres génotypes, il y avait un plus grand nombre de séquences (allant de 75 pour 1G à 292 pour 1E), ces virus provenant de localisations géographiques très largement réparties. De plus, on a mis en évidence une grande diversité intra-génomique pour les virus des génotypes 1G et 2B (au maximum 2,6% et 2,4% respectivement). Avec le renforcement de la surveillance virologique et la mise à disposition de plus en plus de données génétiques, il est clair qu'une description plus précise des virus en circulation deviendra nécessaire pour suivre les virus des génotypes 1E, 1G et 2B.

³ Emily S. Abernathy, et al. Status of global virologic surveillance for rubella viruses. *Journal of Infectious Diseases*, 2011, 204 (suppl 1):S524–532.

³ Emily S. Abernathy, et al. Status of global virologic surveillance for rubella viruses. *Journal of Infectious Diseases*, 2011, 204 (suppl 1):S524–532.

Standard naming for rubella sequences

The proposed naming convention is a slight modification of the existing rubella nomenclature^{1,2} and is designed to harmonize the naming of rubella, measles⁴ and mumps viruses.⁵ The criteria for naming are outlined below.

Rubella virus RNA sequences can be generated from either viral isolates or RNA extracted directly from clinical material. Sequences are designated as either:

RVi – sequence derived from RNA extracted from a rubella virus isolate

RVs – sequence derived from RNA extracted directly from clinical material

Other data that are required to be included in the sequence name are:

- City or state/province where the case occurred (*required*). The inclusion of both city and state information is no longer included in a standard WHO name. Additional geographic information can be collected but will be held independent of the name. Unabbreviated city/state names should be used. Only standard ASCII letters are to be used, and spaces included if needed rather than hyphens or underscores.
- Country, ISO-3 letter designation (*required*). The full ISO-3 list can be accessed from the United Nations website.⁶
- Date of onset of disease by epidemiological week and year (*required*). Week 1 is defined as the week with the first Monday of the year. For example, Sunday 1st January 2012 would be 52.11 and Monday 2nd January 2012 would be 1.12. If the rash onset date is not known, the date of specimen collection should be used. If onset and collection dates are unavailable, the date the sample was received in the laboratory should be used. For sequences where only the year is known, “0” should be used as the epidemiological week. If the year and month are known but not the epidemiological week or day, then the epidemiological week should be defined as the second full week in that month.
- If the WHO name is identical for 2 or more entries, the subsequent sequences should be numbered by 2, 3 etc. Only 1 sequence should be reported from each case even if sequence data are available from multiple specimens.

Dénomination standard des séquences rubéoleuses

La convention proposée modifie légèrement la nomenclature existante pour la rubéole^{1,2} et elle est conçue pour harmoniser la dénomination des virus de rubéole, de la rougeole⁴ et des oreillons.⁵ Les critères de dénomination retenus sont décrits ci-après.

Les séquences d'ARN du virus rubéoleux peuvent être soit obtenues à partir d'isolements viraux, soit extraites directement de matériels cliniques. Les séquences ont donc 2 appellations possibles:

RVi – séquence établie à partir d'un ARN extrait d'un isolement de virus rubéoleux

RVs – séquence établie à partir d'un ARN extrait directement d'un échantillon clinique

Les autres données qui doivent figurer dans le nom de la séquence sont les suivantes:

- Ville ou État/province où le cas s'est produit (*exigé*). Le nom OMS standard n'inclut plus les informations comportant à la fois le nom de la ville et de l'État. Des informations géographiques supplémentaires peuvent être consignées mais resteront indépendantes du nom. On utilise les noms complets, non abrégés des villes ou des États. Seuls les caractères ASCII standards sont employés en y incluant, si nécessaire, des espaces plutôt que des traits d'union ou de soulignement.
- Pays, désignation par les lettres ISO-3 (*exigé*). La liste complète ISO-3 peut être consultée sur le site des Nations Unies.⁶
- Date d'apparition de la maladie par semaine et année épidémiologique (*exigée*). La semaine 1 est définie comme celle comportant le premier lundi de l'année. Par exemple, pour le dimanche 1^{er} janvier 2012 on indiquera 52.11 alors que, pour le lundi 2 janvier 2012, ce sera 1.12. Si l'on ignore la date d'apparition de l'éruption, on utilise la date de prélèvement de l'échantillon. Si ces 2 dates ne sont pas disponibles, on utilise alors celle à laquelle l'échantillon est arrivé au laboratoire. Pour les séquences dont on ne connaît que l'année, on utilisera « 0 » pour indiquer la semaine épidémiologique. Lorsque l'année et le mois sont connus, mais pas la semaine épidémiologique ni le jour, on définit alors la semaine épidémiologique comme la deuxième semaine complète du mois en question.
- En cas de noms OMS identiques pour 2 entrées ou plus, les séquences ultérieures seront numérotées par ordre croissant, 2, 3 etc. Une seule séquence doit être notifiée par cas, même si l'on dispose de données provenant de multiples échantillons.

⁴ See No. 9, 2012, pp. 73–80.

⁵ See No. 22, 2012, pp. 217–224.

⁶ See <http://unstats.un.org/unsd/methods/m49/m49alpha.htm>, accessed June 2013.

⁴ Voir N° 9, 2012, pp. 73-80.

⁵ Voir N° 22, 2012, pp. 217-224.

⁶ Voir <http://unstats.un.org/unsd/methods/m49/m49alpha.htm>, consulté en juin 2013.

The following data are *optional* for inclusion in a WHO name:

- Genotype in square brackets; for genotype assignment, sequencing of the minimum 739 nucleotide window of E1 is required.
- There are special designations for sequences derived from cases with congenital rubella syndrome (CRS) and from newborns with congenital rubella infection (CRI). In both instances the geographic localization is the place of birth and onset date of disease is the date of birth. Sequences from vaccine viruses have a special designation (VAC). Submission of vaccine sequences from recent vaccinees is not encouraged.

The following examples illustrate the current nomenclature:

- RVi/Ho Chi Minh.VNM/41.11/[2B](CRS) designates a genotype 2B rubella virus isolate from a CRS patient in Ho Chi Minh, Vietnam, in the 41st week of 2011.
- RVs/Hong Kong.CHN/20.12/2[1E] designates a genotype 1E rubella virus derived directly from a clinical sample collected in Hong Kong, China, in the 20th week of 2012. This is the second sequence from the same week and location.

Note that a slash is to be used after the year, but not after the sequence number if there is more than 1 sequence from the same location and week.

Data reporting

Network laboratories are requested to share data, including the time and place of detection as well as the genetic identity of rubella viruses identified through national surveillance activities. These data help national and regional surveillance programmes to determine whether outbreaks have been caused by indigenous circulation or importation of viruses, and assist in the global mapping of the circulation of rubella viruses of different genotypes. Laboratories are encouraged to report genetic data on rubella viruses at least monthly, preferably directly to the WHO database RubeNS.

RubeNS database

Initially, a spreadsheet-based genotype database of global rubella viruses was established at WHO to allow timely sharing of genetic information within the WHO LabNet. However, because of the need for increased functionality and capacity to deal with larger numbers of rubella sequences, RubeNS (Rubella Nucleotide Sur-

La figuration des renseignements suivants dans un nom OMS est *facultative*:

- Génotype entre crochets; pour l'attribution à un génotype, il est nécessaire d'avoir le séquençage de la fenêtre minimum de 739 nucléotides du fragment E1.
- Il y a des dénominations spéciales pour les séquences établies à partir des cas de syndrome de rubéole congénitale (CRS pour *congenital rubella syndrome*) et de nouveau nés présentant une infection rubéoleuse congénitale (CRI pour *congenital rubella infection*). Dans les 2 cas, la localisation géographique est le lieu de naissance et la date d'apparition de la maladie, la date de naissance. Les séquences provenant de virus vaccinaux ont une désignation spéciale (VAC). La soumission de séquences vaccinales provenant de sujets récemment vaccinés n'est pas encouragée.

Les exemples qui suivent illustrent la nomenclature actuelle:

- RVi/Ho Chi Minh.VNM/41.11/[2B](CRS) désigne un virus rubéoleux de génotype 2B isolé à partir d'un patient ayant eu un syndrome de rubéole congénitale à Ho Chi Minh Ville (Vietnam), au cours de la 41^e semaine de 2011.
- RVs/Hong Kong.CHN/20.12/2[1E] désigne un virus rubéoleux de génotype 1E obtenu directement à partir d'un échantillon clinique prélevé à Hong Kong (Chine) au cours de de la vingtième semaine de 2012. C'est la seconde séquence obtenue au même endroit au cours de la même semaine.

Il est à noter que s'il y a plusieurs séquences au même endroit et pendant la même semaine, on insère une barre oblique après le nombre indiquant l'année, mais pas après le numéro de la séquence.

Notification des données

Il est demandé aux laboratoires du réseau d'échanger leurs données, notamment concernant le moment et le lieu de la détection, ainsi que l'identité génétique des virus rubéoleux identifiés au cours des activités de la surveillance nationale. Ces données aident les programmes nationaux et régionaux de surveillance à déterminer si les flambées ont été causées par une circulation autochtone ou par l'importation de virus; elles contribuent à l'établissement de la cartographie mondiale de la circulation des virus rubéoleux appartenant aux différents génotypes. Les laboratoires sont invités à notifier les données génétiques sur les virus rubéoleux au moins une fois par mois, de préférence directement à la base de données RubeNS de l'OMS.

Base de données RubeNS

Au début, une base de données sur les génotypes sous forme de tableur a été créée à l'OMS pour permettre les échanges d'informations génétiques en temps voulu au sein du LabNet. Toutefois, avec le besoin d'étendre les fonctionnalités et d'augmenter la capacité de traiter des nombres importants de séquences rubéoleuses, RubeNS (Rubella Nucleotide Sur-

veillance) was developed. This sequence database initiative is a LabNet activity under technical leadership from PHE, with a steering committee comprising representatives from the WHO LabNet Global Specialized or Regional Reference Laboratories. The corresponding database and web-application⁷ are maintained by PHE and sequence data are contributed by LabNet member laboratories or downloaded from GenBank. The main objective of the database is to collect sequence information from the 739 nucleotide fragment (the minimum fragment required for genotyping), the complete E1 coding region, or the complete structural polyprotein, or the complete genome, for comparison with other sequences collected globally.

Additional information, including epidemiological information on the case, is also collected if available. An example of important epidemiological information for inclusion is the importation status. If a virus can be identified as an importation using standard epidemiological means and the name of the exporting country is known, this information should be provided. Information must be sufficient to create a “WHO name” as described in the standard format defined in this document, or a standard format name may be submitted directly. Individual sequences can be assigned for public or private viewing. Sequences designated as public are viewable by all submitters to RubeNS. Those sequences designated as private will only be viewable by the sequence submitter, the RubeNS steering committee and the rubella programme at WHO.

The website features the ability to search and filter sequences submitted to the database using geographical, temporal and epidemiological parameters. RubeNS has the facility to enable sequence submitters to upload their sequence data from RubeNS to GenBank. Bioinformatics tools in RubeNS allow users to find identical or similar sequences and identify a preliminary genotype designation. RubeNS also allows reference sequences (*Table 1*) to be downloaded so that rubella genotyping can be performed using common phyloge-

Surveillance) a été mise au point. Cette initiative, une base de données sur les séquences, est une activité de LabNet sous la direction technique de PHE, avec un comité d'orientation constitué des représentants des laboratoires mondiaux spécialisés ou régionaux de référence du LabNet de l'OMS. La base de données et l'application sur le Web correspondantes⁷ sont tenues à jour par PHE et les données sur les séquences sont fournies par les laboratoires membres du LabNet ou téléchargées à partir de GenBank. La base de données a pour objectif principal de recueillir les informations sur la séquence du fragment de 739 nucléotides (le fragment minimum requis pour le génotypage), c'est-à-dire la région codante E1 complète, sur la polyprotéine de structure complète ou sur le génome complet, pour comparaison avec d'autres séquences collectées dans le monde entier.

Des informations supplémentaires, notamment de nature épidémiologique sur le cas, sont également collectées s'il y en a. Par exemple, une information épidémiologique importante à inclure dans les données concerne l'état de l'importation. Si, en appliquant les moyens épidémiologiques standards, on peut identifier un virus comme résultant d'une importation et si l'on connaît le pays d'origine, ces renseignements doivent être indiqués. Les informations fournies doivent être suffisantes pour créer un « nom OMS », tel que décrit dans la dénomination standardisée définie dans le présent document, sinon une dénomination sous forme standard peut être soumise directement. Les séquences individuelles peuvent être classées comme accessibles au public ou privées. Celles qui sont classées comme accessibles au public peuvent être consultées par tous ceux qui soumettent des données dans RubeNS. En revanche, celles qui sont classées comme privées ne sont accessibles qu'à ceux qui les ont soumises, au comité d'orientation de RubeNS et au programme de lutte contre la rubéole à l'OMS.

Le site Web a des fonctions permettant de rechercher et de filtrer les séquences soumises à la base de données selon des paramètres géographiques, temporels et épidémiologiques. RubeNS donne la possibilité à ceux qui soumettent des séquences de télécharger de RubeNS à GenBank celles dont ils sont propriétaires. Des outils de bioinformatique dans RubeNS permettent aux utilisateurs de trouver des séquences identiques ou similaires et d'identifier une désignation préliminaire du génotype. La base permet aussi de télécharger des séquences

预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=5_28109

