Development of metrological support for nucleic acid measurements

M S Vonsky¹,², A L Runov¹,²

¹ D.I. Mendeleev Institute for Metrology, Moskovsky ave., 19, St.-Petersburg 190005, Russia
² Almazov National Medical Research Centre, Akkuratova str., 2, St.-Petersburg 197341, Russia

E-mail: m.s.vonsky@vniim.ru

Abstract. Measurements of nucleic acid have become an essential part of analytical applications in many fields including life science, medicine, veterinary medicine and biotechnology. Peculiarities of these measurements are largely determined by the uniqueness of quantity and analyte definitions. Quantitative PCR, applied for routine nucleic acids measurements, is a relative method demanding external calibrators with quantity values traceable to corresponding reference materials. Direct nucleic acid measurements became possible after introduction of the digital PCR (dPCR) method. Ensuring accurate, comparable and traceable to SI nucleic acids measurements is the focus of Nucleic Acid Analysis Working Group of the Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology. A number of pilot and key comparisons carried out by Working Group have shown that application of dPCR allows participating National Metrological Institutes to obtain comparable results of measurements of the sequence copy number concentrations and copy number ratios.

1. Introduction

Measurements of quantitative characteristics of nucleic acids, i.e. the content of certain nucleotide sequences, are widely used in biotechnological and biopharmaceutical laboratories, clinical diagnostic and veterinary laboratories, forensic medical examination, sanitary and epidemiological and environmental control. The peculiarities of these measurements are largely determined by the uniqueness of the subject, its properties and definitions of analyte and measured quantity.

2. Properties and quantities in nucleic acid sequence measurements

Nucleic acids are presented as extended linear polymers, the simplified structural formula of which can be represented as a sequence of four monomers - nucleotides A, C, G, and T for DNA or A, C, G, and U for RNA. The primary structure of a nucleic acid, or a specific sequence of nucleotides, uniquely identifies a biomolecular object - a fragment of a DNA or RNA molecule and represents a nominal property - a property of a phenomenon, body, or substance, where the property has no magnitude [1], no size [2] and can be compared only by equivalence. The best characterized sequences of various organisms and viruses have been deposited in the database of reference sequences RefSeq [3], which is maintained by the international cooperation of the nucleotide sequence databases INSDC. Manufacturers of reference materials whose certified characteristic is a nucleotide sequence provide them with references to RefSeq entries. For example, in SRM 2392 produced by NIST (USA) the certified nucleotide sequences of the human mitochondrial genome are given with reference to the Revised Cambridge Reference Sequence [4], deposited as NCBI Reference Sequence: NC_012920.1 [5].

The quantity expressing the content of nucleic acids fragments is the number of copies of the sequence encoding them. This quantity – "number of copies of sequence" – cannot be described in
terms of seven base SI quantities, but represents a counting quantity with the associated unit "one". The unit "one" is a neutral element of any system of units and enters any system of units automatically [6]. Thus, "number of copies of a sequence" is a dimensionless SI quantity. Section 5.4.7 of the "SI Units" brochure prescribes that only numbers should be used to express values of quantities with the unit "one". Although in practice "copy" or "cp" unit is used to express the results of nucleic acid sequence copy number measurements, this result should be expressed simply as a number, and sequence copy concentration as \( \mu \text{L}^{-1} \) (not "cp/\mu\text{L}"). In the same way, the results of sequence copy number ratio measurements should be expressed using the internationally recognized symbol “%”, not “cp/cp”.

While a specific nucleotide sequence is defined formally, physically a given nucleotide sequence can be part of a whole set of nucleic acid macromolecules of different sizes. Theoretically, when recalculating the number of nucleotide base pairs into length units, genomic DNA molecules of human chromosomes have linear length from 16 mm for chromosome 21 to 85 mm for chromosome 1 [7]. In practice it is impossible to isolate full-length DNA macromolecules from biological material. Beside biological mechanisms that cause rapid destruction of the macromolecules integral structure during cell death, the procedures of cell destruction, genomic DNA isolation and purification are accompanied with hydrodynamic shocks leading to DNA fragmentation [8]. Impossibility of clear description of occasionally fragmented macromolecules rendered the concept of mole used to express the amount of substance in inorganic and organic chemistry inapplicable for expression of nucleic acid sequence quantity. Indeed, according to the new mole definition, «The mole, symbol mol, is the SI unit of amount of substance. One mole contains exactly 6.022 140 76 x 10^{23} elementary entities» [9], and entities can be represented by atoms, molecules, ions, or electrons [10]. A formally defined but physically non-existent nucleotide sequence does not satisfy this requirement, although in some cases the content of well-defined short nucleic acid fragments is expressed in moles [11].

3. A number of copies as a dimensionless SI quantity

Until recently, measurements of DNA target sequences content were performed using the real-time polymerase chain reaction method, or quantitative PCR (qPCR). This method is relative and requires application of external standards (calibrators) with assigned values, traceable to the appropriate reference materials (CRM). The certified value of the DNA content in CRMs can be expressed in a mass fraction (g/kg) or in percents and in WHO International Standards - in international units (IU) For example, 3rd International Standard, 2011 of Hepatitis B virus DNA, NIBSC code 10/264, contains 850,000 IU/vial [12].

Accuracy of measurements results obtained with qPCR is largely determined by the equivalence of amplification efficiencies in all reactions – with the calibrator and with the sample DNA. Even minor differences in amplification efficiencies can lead to significant bias in results – it was shown, that when PCR efficiencies in assay vary from 1.70 to 1.91, actual measurement result vary from 0.25 to 139 times [13].

Introduction of the digital PCR (dPCR) method makes it possible to perform direct measurements of the target DNA sequences copy number concentration. A number of international pilots and key comparisons at the highest level of accuracy were conducted by the Nucleic Acid Analysis Working Group under the auspices of the Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology (CCQM). Application of dPCR method by participating National Metrological Institutes ensured acquisition of comparable measurement results of target sequence copy number concentration with values close to 10 \( \mu \text{L}^{-1} \), i.e., near the method quantification limit and sequence copy number ratios up to 1:1000 (0.1 %) [14].

ID-MS, isotope dilution mass spectrometry, is an acknowledged primary method of measurement, i.e. a method having the highest metrological qualities, whose operation is completely described and understood, for which a complete uncertainty statement can be written down in terms of SI units, and whose results are, therefore, accepted without reference to a standard of the quantity being measured [15]. Application of so called “orthogonal” methods for target sequence copy number measurements,
including mass spectrometry with isotope dilution (ID-MS) and single molecule flow cytometry confirmed dPCR results in recent CCQM-P199b pilot study on amount of the SARS-CoV-2 viral RNA measurements [16]

As a result of research project carried out at D.I. Mendeleev Institute for Metrology (VNIIM) in the period 2019-2020, laboratory infrastructure was formed that ensured successful participation in a number of CCQM pilot and key comparisons. Results of VNIIM participation in CCQM-K86c [14] were used to support the first calibration and measurement capability (CMC) claim for the Russian Federation in the field of DNA measurements. The capability was registered in the BIPM CMC Capabilities Database.

The Federal Agency on Technical Regulating and Metrology task to VNIIM to create a state primary quantity unit standard for DNA sequence copy number will make it possible to ensure on the national level traceability of nucleic acid measurement results to the dimensionless SI quantity unit “sequence number of copies”. Transfer of DNA quantity units from the state primary standard to customers will be realized through the development and production of certified reference materials, the metrological characteristics of which include values of the sequence copy number concentration and/or the sequences copy number ratio.

4. Conclusion
Development of metrological support for nucleic acid sequence copy number measurements is highly stimulated by current pandemic situation worldwide. Collaboration of National Metrological Institutes in the CCQM Nucleic Acid Analysis Working Group promotes global comparability of nucleic acid measurements worldwide. Traceability of measurement results to SI dimensionless unit sequence number of copies on national level will be established via creation of state primary standard followed by elaboration of reference materials meeting the needs of customers.

References
[1] International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM) 3rd Edition 2012 JCGM, https://www.bipm.org/en/publications/guides
[2] Nordin G et al Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences (IFCC-IUPAC Recommendations 2017) 2018 Pure and Applied Chemistry 90
[3] O'Leary N et al 2016 Nucleic Acids Res. 44(D1) D733-45
[4] Andrews R et al 1999 Nat. Genet. 23 p 147
[5] Homo sapiens mitochondrion, complete genome. NCBI Reference Sequence: NC_012920.1 Internet resource, https://www.ncbi.nlm.nih.gov/nucleotide/NC_012920.1 Retrieved 03 November 2021
[6] The International System of Units (SI) 2019 , https://www.bipm.org/en/measurement-units
[7] Piovesan A et al 2019 BMC Research Notes 12 (1) 106
[8] Matniyazov R et al 2014 Biomics 6 (3) pp 155-166
[9] Roberto M et al 2018 Pure Appl. Chem. 90 (1), pp. 175-180
[10] “On the revision of the International System of Units - International Union of Pure and Applied Chemistry”. IUPAC | International Union of Pure and Applied Chemistry. 16 November 2018. Retrieved 03 November 2021
[11] Kaczmarczyk A Bowra S Elek Z and Vincze E 2012 BMC Plant Biol. 12 184
[12] Fryer J Heath A Wilkinson D and Minor P 2017 Biologicals 46 57-63
[13] Ruijter J et al 2021 Efficiency-correction is required for accurate qPCR analysis and reporting Clinical chemistry 67
[14] Mester Z et al 2020 Metrologia 57(1A) 0804
[15] Milton M and Quinn T 2001 Metrologia 38 289
[16] National Measurement Institutes demonstrate high accuracy reference measurement system for SARS-CoV-2 testing. https://www.bipm.org/en/~2020-nmi-covid Retrieved 03 November 2021