Progress in Nucleic Acid Testing and Vaccine Research for the Novel Coronavirus

Shu Yuqing1,a, Li Yajun2,b, Jiang Haiyin3,c, Luan Yifei3,d,*

1Hangzhou Medical College, School of Public Health, Hangzhou, Zhejiang, 311399, China
2The First People’s Hospital of Guiyang, Guiyang, Guizhou, 550002, China
3Hangzhou Medical College, Hangzhou, Zhejiang, 311300, China
a2173962741@qq.com, b1512713923@qq.com, c2020000298@hmc.edu.cn,
dluanyifei163@163.com
*Corresponding author

Keywords: Novel coronavirus pneumonia; progress; nucleic acid testing; vaccine; virus detection

Abstract: Since the outbreak of COVID-19, nucleic acid testing has become a critical means for epidemic control, playing an irreplaceable role in containment. Firstly, testing techniques have evolved from initial real-time fluorescence PCR to advanced high-throughput detection technologies including isothermal nucleic acid amplification, nucleic acid mass spectrometry, and next-generation sequencing. These new techniques have enabled highly efficient nucleic acid detection and significantly shortened turnaround time, with improved sensitivity. However, the stability and accuracy of detection still need improvement, requiring comprehensive optimization in aspects like sample processing and reagent selection. In the meantime, remarkable progress has also been achieved in vaccine development, with multiple COVID-19 vaccines approved for use. Vaccines can be categorized into inactivated vaccines, recombinant protein vaccines, DNA vaccines, mRNA vaccines, and viral vector vaccines based on technological approaches. All these vaccines can induce immune responses to some extent, but protection durability and efficacy against emerging variants remain to be investigated. Optimizing vaccine design to enhance safety and immunogenicity is the current priority.

1. Introduction

Compared to influenza viruses, the Omicron variant poses greater risks, even though it exhibits lower virulence, as it causes a higher number of asymptomatic infections and mild cases with upper respiratory tract involvement. In the past, repeated COVID-19 outbreaks led to a grim situation characterized by a rapid increase in domestic confirmed cases and asymptomatic infections worldwide, while community transmission remained uncontrolled in certain regions. The epidemic reached a critical level, with widespread outbreaks and sporadic cases intermingled, and both imported cases and local clusters occurred frequently. Nucleic acid testing has become an essential method for the rapid screening of suspected COVID-19 infections.

Mass nucleic acid testing and proactive screening are essential strategies aimed at achieving "early
detection, early isolation, early diagnosis, and early treatment" of COVID-19 infected patients. These efforts, in conjunction with the development and administration of multiple COVID-19 vaccines, have injected new momentum into pandemic prevention and control. In the initial phase of the pandemic, nucleic acid testing emerged as the primary and feasible method for detecting the virus. It enabled medical facilities to rapidly screen suspected COVID-19 cases and individuals from high-risk areas, buying precious time for emergency and critical care interventions. While significant progress has been made in pandemic prevention and control with the introduction of multiple COVID-19 vaccines, nucleic acid testing played an indispensable role during the early stages of the pandemic. It remains a crucial tool for identifying and containing outbreaks.

Notably, to date, no comprehensive review on this topic has been published. Therefore, conducting a literature review in this area would provide novel insights into the importance and impact of nucleic acid testing during the early phase of the pandemic and its continued relevance alongside vaccination efforts.

2. Characteristics of Pathology

The novel coronavirus pneumonia (coronavirus disease 2019, COVID-19) is an acute respiratory infectious disease caused by the novel coronavirus [1], SARS-CoV-2 is a virus with a diameter of 60-140 nm. It contains a single-stranded RNA genome, nucleocapsid proteins, envelope proteins, and outer membrane components. The genome is 29.8 to 29.9 kb in length and comprises 14 open reading frames (ORFs) that encode 27 proteins[2]. Among the ORFs, ORF1ab is the largest gene located in the untranslated region (UTR) and encodes various proteins required for viral transcription and replication, including multiple non-structural proteins (NSPs). The genes located in the 3'-UTR encode four major structural proteins: the spike protein (S protein), membrane protein (M protein), envelope protein (E protein), and nucleocapsid protein (N protein) [3]. The novel coronavirus spreads through respiratory droplets, direct contact, and fecal-oral transmission[4,5]. Infected individuals typically develop symptoms within 8.2-15.6 days, with an average of 11.2 days. The elderly progress faster than the young. The virus deposits in the respiratory tract, gradually reaches deeper into the lungs, and can damage the nervous system (e.g. brain), digestive system (e.g. liver, stomach, intestines), urinary system (e.g. kidneys) and cardiovascular system in severe cases [6].

China reported its first confirmed case on January 23, 2020. Clinical observations have classified the symptoms into mild, moderate, severe and critical types. Moderate symptoms are most common, mainly presenting as fever, dry cough and fatigue. Patients may also experience sore throat, muscle ache, diarrhea, and loss of taste. Severe cases can develop breathing difficulties, hypoxemia, acute respiratory distress syndrome, septic shock, multi-organ failure[7], and even death[8]. The rapid diagnosis of novel coronavirus infection is crucial for controlling clustered outbreaks. To achieve this goal, it is imperative to develop high-throughput and accurate nucleic acid testing technologies.

3. Similarities and Differences between Nucleic Acid Sampling and Testing Methods

Detection of the virus in upper and lower respiratory tract specimens is effective for diagnosing active clinical infection of the novel coronavirus. Respiratory specimens have higher detection rates than blood samples. However, there are differences in detection rates among specimens collected from different sites. Bronchoalveolar lavage fluid from the lower respiratory tract has the highest detection rate, followed by sputum from deep cough, nasopharyngeal swabs, and oropharyngeal swabs. In summary, lower respiratory tract specimens are preferred for accurate diagnosis of coronavirus infection. Sample collection is a critical first step for COVID-19 testing. Sample collectors obtain different types of swabs from patients and extract high purity nucleic acids for testing. There are a few swabbing strategies tailored for different types of patients. Nasopharyngeal
Swabs are often used for people in centralized isolation or close contacts. For patients with cough or late stage disease, anal swabs are recommended. Additionally, positive anal swabs correlate with disease severity and can serve as an early warning sign for severe illnesses. By using suitable swabbing approaches for different patients, sample collectors can obtain optimal specimens for COVID-19 testing[9,10]. Pooled sampling and pooled testing are two approaches for mass nucleic acid testing that may improve efficiency to some extent. However, pooled sampling introduces large amounts of human genomic DNA, slowing down subsequent nucleic acid extraction and PCR amplification. On the other hand, pooled testing reduces the amount of each sample, lowering the detection rate for positive samples. Therefore, for routine surveillance or testing of people traveling across provinces, multiple consecutive sampling and testing are recommended to increase overall detection rate. The sampling and testing methods should be chosen rationally according to the situation to ensure effectiveness. For example, repeated testing of suspected cases with initial negative results helps reduce the risk of false negatives and improve detection.

4. Principles of Nucleic Acid Testing

Nucleic acid testing enables diagnosis of novel coronavirus infection by detecting viral RNA in patient upper respiratory tract samples. The feasibility depends on sufficient viral load. Traditional methods like real time digital PCR take hours, insufficient for rapid accurate diagnosis. Thus, researchers worldwide have developed various testing principles. Quantitative PCR incorporates fluorescent probes for quantification. Loop-mediated isothermal amplification, LAMP utilizes continuous DNA synthesis under constant temperature for rapid amplification. Nucleic acid mass spectrometry uses mass spectrometers to analyze nucleic acid fragments, simplifying workflows. High-throughput sequencing performs large-scale parallel sequencing to directly detect viral genomes. These technologies improve sensitivity and specificity while greatly shortening testing time, enabling rapid precise diagnosis of novel coronavirus infection.

Based on technical principles, real time digital PCR fluorescence techniques can be divided into two categories. One utilizes intercalating fluorescent dyes to monitor PCR amplification in real-time for quantification. The other employs sequence-specific fluorescent probes to monitor fluorescence signals from target sequences in real-time. Due to its high sensitivity and low miss rate, quantitative PCR fluorescence is widely used for novel coronavirus confirmation globally. Specific primers exponentially amplify target sequences, allowing detection of the virus at early infection stages. In summary, real time digital PCR fluorescence techniques are rapid and sensitive, allowing early detection of positive cases. However, they require specialized laboratory equipment and trained personnel. Operational errors or suboptimal laboratory conditions can cause aerosol contamination leading to false positive results.

Despite the limitations, real time digital PCR fluorescence remains an important tool for novel coronavirus testing. Optimizing experimental conditions and strictly following protocols will maximize the strengths of this technique for rapid accurate diagnosis. Targeting key viral genes including ORF1ab (RdRp), N, E, S and ORF8 of the novel coronavirus is recommended. The RdRp gene is located in ORF1ab and N and E genes are highly conserved. Detection of RdRp and E has fewer limitations and higher sensitivity compared to the N gene. Thus, RdRp and E are preferred targets for novel coronavirus testing[11]. The World Health Organization has published primer sequences targeting the E and RdRp genes, which were first used globally to screen and confirm SARS-CoV-2. Methods designed based on these can successfully distinguish between SARS-CoV and SARS-CoV-2[12]. The Chinese Center for Disease Control and Prevention designed primers targeting the N gene and ORF1ab for detecting novel coronavirus RNA.

LAMP techniques use target gene-specific primers and strand displacement DNA polymerases to
generate loop-structured amplicons, achieving exponential amplification. LAMP techniques have been applied to develop SARS-CoV-2 diagnostic kits, such as the respiratory virus detection array co-developed by Biogerm and Tsinghua University. They have the advantage of high specificity. LAMP techniques are valuable additions to the nucleic acid testing toolkit for combating the pandemic. However, compared to PCR, LAMP has higher primer design requirements and increased difficulty in application. Aerosol contamination can also lead to false negatives. Therefore, further optimization of key factors for LAMP techniques is crucial for improving detection performance.

Nucleic acid mass spectrometry techniques offer advantages like rapid detection speed, high accuracy and sensitivity. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is an emerging biotechnology that analyzes and identifies nucleic acids by sequence. It has been widely applied in novel coronavirus detection. Primers are designed to amplify viral gene sequences and single nucleotide polymorphism (SNP) sites. MALDI-TOF MS detects the amplicons and identifies the virus by matching the peak pattern to reference spectra of specific gene fragments. Amplicons spanning SNP sites generate spectra with variable peak masses, allowing virus identification. By amplifying sequences across SNPs and analyzing the resultant mass spectra, MALDI-TOF MS can reliably detect and identify viruses.

SMRT Sequencing, a type of NGS, analyzes nucleic acids from clinical samples to detect, genotype and trace infectious diseases. Advantages include increased throughput, lower costs, and shortened testing turnaround. The workflow involves collecting patient samples, amplifying and purifying cDNA, quantifying nucleic acids, generating sequencing libraries from templates, and performing genomic sequencing. Recently, novel coronavirus has mutated with about 2 nucleotide substitutions accumulating every month.

A key advantage of SMRT Sequencing is monitoring viral genomic changes to promptly detect variants before large-scale spread. In samples with extremely low viral loads, NGS enables high genomic coverage for detecting mutations in all regions, overcoming limitations of PCR-based methods missing mutations due to insufficient viral copies. SMRT Sequencing has some key limitations, including inaccurate detection of minority variants, expensive instruments, and need for expert interpretation. Similarly, while NGS provides powerful tools for viral detection and variant surveillance, it also has limitations to be addressed. Specifically, NGS workflows involve complex library preparation, require high-quality input DNA, and remain difficult to implement clinically compared to PCR-based approaches.

5. Vaccine Research

The goal of developing COVID-19 vaccines is to prevent disease caused by the novel coronavirus SARS-CoV-2, providing an effective preventive measure for people worldwide. By establishing herd immunity through vaccination, the spread of the virus can be controlled and incidence and mortality reduced. Furthermore, COVID-19 vaccination protects those at high risk of severe illness or death, including the elderly, immunocompromised individuals, and healthcare workers. In this way, vaccines can curb the global public health crisis.

SARS-CoV-2 infection is mediated by the viral surface spike protein binding to the human ACE2 receptor. The S protein comprises functional subunits S1 and S2, where S1 is responsible for ACE2 binding and S2 facilitates virus-cell membrane fusion. As the S protein is critical for viral entry, it is a key target for vaccines against COVID-19. Vaccine-induced antibodies recognize the S protein and block its interaction with ACE2, neutralizing the virus and preventing cell entry. Through this mechanism, vaccines provide protection against SARS-CoV-2[18]. In addition, vaccines elicit T lymphocyte responses that release cytokines and clear infected cells. By stimulating humoral and cellular immunity against the S protein, vaccines block the critical step of viral entry, aligning with
the mechanism of SARS-CoV-2 infection[19].

Currently, various COVID-19 vaccine platforms are under research and clinical trials globally. Each platform has relative advantages and disadvantages for inducing virus-specific antibodies to block infection and promoting cell-mediated clearance of infected cells. Relying on one platform alone is unlikely to meet diverse population needs. Inactivated vaccines are well-established with proven safety and production capacity, suitable for rapid pandemic control. However, they have low immunogenicity and require repeated boosters for protection. In contrast, recombinant protein vaccines contain only viral antigens and have controllable safety, but poor efficacy alone, needing immunostimulant adjuvants. DNA and RNA vaccines have theoretical advantages like persistent stimulation, but also have safety concerns over possible gene integration and instability. Finally, viral vector vaccines potently activate immunity but have risks needing further evaluation.

- **Inactivated vaccines**: involving chemically or physically inactivating pathogenic viruses to remove replicability while retaining some antigenicity. This activates B cells to produce antibodies that can neutralize the target virus. Production methods include chemical inactivation with formaldehyde or β-propiolactone, or physical treatments like UV and heat. Key advantages of inactivated vaccines include high safety without risk of virulent strains, and the ability to leverage mature production processes for rapid mass manufacturing. However, some antigenic epitopes can be altered during inactivation, reducing immunogenicity. Inactivated COVID-19 vaccines developed by Sinovac and Sinopharm in China demonstrate these advantages, enabling rapid pandemic response through multiple doses. Their manufacturability plays a vital role in COVID-19 control.

- **Recombinant protein vaccines**: involving producing viral surface protein antigens using gene engineering and recombinant DNA technology. The gene sequences encoding target antigens are inserted into bacteria or yeast to generate recombinant proteins. These proteins are then purified and formulated into vaccines without containing whole virus particles. Recombinant protein vaccines have several key advantages. Most notably, they confer good safety with no risk of viral replication or infection, since they do not contain whole viruses. Additionally, they allow precise immune targeting of specific viral antigens. In contrast to some other vaccines, recombinant proteins pose no risk of reverting to virulent strains. However, recombinant protein vaccines often require adjuvants and multiple doses to potentiate immunity[20].

As standalone protein antigens have low immunogenicity, adjuvants and multiple doses are needed to improve protection. Novavax’s COVID-19 recombinant protein vaccine has entered Phase 3 trials, generating neutralizing antibody levels higher than convalescent sera[21]. Recombinant protein vaccines provide precisely defined antigens and are relatively easy to standardize for mass production after the initial development process. With further optimizations to immunogenicity, they can serve as an important component of the COVID-19 vaccine landscape. To enhance the protective effects of recombinant protein vaccines, further optimizations to vaccine design and formulation are needed.

- **DNA vaccines**: involving injecting plasmid DNA encoding viral antigens to be expressed in human cells, stimulating immune responses. Key advantages include simple preparation, high stability, and the ability to modulate different arms of immunity. DNA vaccines are also highly preservable. However, they face some challenges. Most importantly, there are concerns about safety, including potential DNA integration causing mutations or activating oncogenes[22]. Additionally, DNA vaccines currently have weak immunogenicity, needing optimized dosing and administration schedules. The plasmid DNA is also prone to degradation by nucleases. Despite these obstacles, DNA vaccines represent an important next-generation approach that can potentially overcome limitations through research on new delivery methods, adjuvants, antigen design, and manufacturing optimizations.

- **RNA vaccines**: containing viral antigen-encoding mRNA as the active drug substance. Host cells translate the mRNA into target antigens to elicit immune responses. Key advantages of RNA
vaccines include rapid development, design flexibility, and ease of large-scale production. There are two major types of RNA vaccine platforms: non-replicating mRNA and self-amplifying mRNA. In particular, Moderna’s mRNA-1273 COVID-19 vaccine comprises of modified mRNA encoding the spike protein antigen encapsulated in lipid nanoparticles. This mRNA is delivered into human cells for translation and expression of the target antigen, activating humoral and cellular immunity. Additionally, the mRNA design allows precise targeting of key viral antigens. Messenger RNA is inherently unstable. This instability necessitates optimized delivery systems to protect and stabilize the mRNA. Consequently, these systems must provide efficient storage stability at room temperature for at least one week. In particular, the mRNA must remain thermostable to meet transportation needs without relying on cold-chain storage. Addressing the challenges of mRNA instability is critical for developing effective mRNA vaccine formulations and delivery systems.

- Viral vector vaccines: involving introducing antigen-encoding genes into replication-deficient viral vectors to deliver and express target antigens, eliciting immune responses. They stimulate both humoral and cellular immunity with high immunogenicity. Booster shots can further enhance effects. However, viral vectors also have some potential drawbacks. Although risks are theoretically low, concerns exist over possible reversion to virulence. Pre-existing vector immunity may also impact efficacy. Companies are optimizing vector selection and engineering strategies to overcome these challenges. Further optimizations to vectors and attenuation methods can continue improving safety.

COVID-19 vaccine development has abandoned the previous serial approach for a faster parallel model with concurrent trial phases and real-time monitoring. Specifically, vaccination strategies need to differ based on age. Children have lower COVID-19 incidence, milder symptoms, and less severe disease courses compared to adults. This reflects underlying physiological factors that confer relative viral resistance in children. For instance, previous observations of lower SARS rates in children also indicate possible structural or immunological protectiveness against coronaviruses. Given these age-based differences in susceptibility, customized vaccination approaches are likely needed for pediatric populations. In addition to age, pre-existing medical conditions also require tailored vaccination approaches. For example, individuals with autoimmunity often have overactive immune systems and inflammation that could potentially be exacerbated by vaccines. Cohorts with inflammatory eye diseases have endogenous immune factors that may be prone to flare-ups. Therefore, customized strategies or contraindications need to be considered for vaccinating those with autoimmunity, with careful testing in these subgroups. Countries have worked diligently to efficiently develop COVID-19 vaccines by adhering to regulations and technical requirements, with profound impacts on pandemic resistance. This momentum has greatly catalyzed the evolution of healthcare systems towards precision medicine – tailoring treatments to individuals based on real-time assessment of changing factors, surveilling for risks, and advancing the sector overall. For instance, the rapid development of COVID vaccines demonstrates real-time assessment and adjustment to match the emerging virus. Strategies like age-based regimens also showcase precision by tailoring vaccines to specific groups. Overall, the global COVID-19 vaccine effort has accelerated advances in data-driven, personalized medicine.

6. Conclusion

At the initial outbreak, nucleic acid testing was the primary diagnostic method. However, the high transmissibility of the novel coronavirus led to false negative results, with accuracy affected by sampling, quality and storage factors. Moreover, it only detects viral load at the time of testing, missing early infections and recovered individuals. Given the drawbacks of nucleic acid testing, vaccination has been progressively implemented for pandemic control. First, it helps assess overall transmission extent and total infected numbers in populations. Second, vaccination characterizes
population immunity levels. It also enables screening of blood donors and convalescent plasma candidates. Finally, monitoring vaccine coverage provides data on epidemiological patterns and trends. In these ways, vaccination has become a vital tool in the global fight against COVID-19.

Within the first 5.5 days of SARS-CoV-2 infection, nucleic acid testing has a positivity rate up to 90%, higher than that of antibody testing. Therefore, combining these two methods improves overall virus detection to 98.6%, with complementary strengths. First, since antibody testing catches cases missed by early nucleic acid testing, it helps validate nucleic acid results to reduce misdiagnosis and missed diagnoses. Retrospective serological surveys especially benefit, clearly delineating epidemiological patterns of infection. Second, serology enables workforce resumption for those with antibodies. Finally, it plays a vital role in reinforcing pandemic control through these multifaceted applications[25]. Nucleic acid testing enables early identification of positive cases, technically empowering pandemic control through "early detection". Antibody testing complements nucleic acid testing for preventive screening. For clinical applications, combining both methods along with multiple specimen types like throat swabs, stool, blood provides robust diagnostic support. However, the current reliance on nucleic acid and antibody testing for clinical criteria like discharge makes proper sampling and test selection crucial. Inadequate viral loads and improper sample handling can lead to false negative results. Different sampling sites should be used based on disease stage, as positivity varies across sites for a given patient. National COVID-19 diagnosis and treatment protocols provide authoritative details, along with laboratory testing guidelines. In summary, careful adherence to protocols is critical given the reliance on testing for clinical decision-making.

As emerging SARS-CoV-2 mutations continue to pose challenges, both nucleic acid testing and vaccines are facing new obstacles. In testing, factors such as rheumatoid factors and heterophilic antibodies can lead to non-specific binding and false positive results. Moreover, cross-reactivity between different coronaviruses can complicate the interpretation of test outcomes. To tackle these issues, it is crucial to develop highly specific antibodies or employ sandwich assays that can minimize non-specific binding. Additionally, optimizing assay formulations can help in suppressing interfering factors and improving the accuracy of test results. Upgrading instrumentation and implementing multi-layered quality control measures are also essential to enhance the precision of testing.

In the case of vaccines, the emergence of new mutations may impact their efficacy and the overall immune response. Therefore, it is essential to continue research and development efforts to adapt vaccine technologies to address the evolving virus. An integrated approach that fosters innovations in both testing and vaccine development is vital in the ongoing fight against COVID-19. By proactively addressing the challenges presented by the mutating virus, we can better equip ourselves to combat the pandemic effectively.

Author Contributions

Shu YuQing designed the study, collected the data, analyzed and interpreted the data and wrote the manuscript. LiYajun was the second reviewer; Jiang Haiyin was the third reviewer; Luan Yifei was the Fourth reviewer. All authors have read and agreed to the published version of the manuscript.

Acknowledgement

Funding: National Innovation and Entrepreneurship Training Program for College Students (202313023005)

References

[1] Xin Guo, Sun Rong. Prescription-syndrome relationship in traditional Chinese medicine and pharmacology research


