Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Differential prolonged multiomic responses to mRNA and inactivated virus COVID-19 vaccines

Hein M Tun
The Chinese University of Hong Kong
Shilin Zhao
The Chinese University of Hong Kong
Chunke Chen
The Chinese University of Hong Kong
Yuzhou Chen
The Chinese University of Hong Kong
Ye Peng
The Chinese University of Hong Kong
Huibin Lv
University of Illinois Urbana-Champaign
Jie Zhu
The Chinese University of Hong Kong
Yun Sang Tang
The Chinese University of Hong Kong
Ken KP Chan
The Chinese University of Hong Kong
Bonaventure Y Ip
The Chinese University of Hong Kong
Yuanxin Sun
The Chinese University of Hong Kong
Xin Liu
The Chinese University of Hong Kong
Karen Yiu
The Chinese University of Hong Kong
David SC Hui
The Chinese University of Hong Kong
Chris Ka Pun Mok
kapunmok@cuhk.edu.hk

The Chinese University of Hong Kong https://orcid.org/0000-0002-0525-6772

Research Article

Keywords:

Posted Date: June 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10011166/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at Cellular and Molecular Life Sciences on September 29th, 2026. See the published version at https://doi.org/10.1007/s00018-026-06455-z.

Abstract

Despite the significant impact of SARS-CoV-2 vaccines in curtailing the spread and severity of COVID-19, comprehensive omics-level understanding of longitudinal host responses remains elusive. We recruited 553 Hong Kong participants receiving two doses of CoronaVac (inactivated virus) or BNT162b2 (mRNA-based). Longitudinal blood samples (baseline, one-month, six-months post-vaccination) were collected. We assessed the longitudinal changes in omics landscape and their impact on host immunity by applying transcriptomic, metabolomic, lipidomic and cytokine datasets. Our data delineate pronounced differences in the multi-omics profiles between the two vaccine platforms examined. The mRNA-based BNT162b2 vaccine induced more profound changes in omics profiles compared to the inactivated CoronaVac, with notable alterations in adaptive immunity and metabolomic functions. Importantly, BNT162b2 displayed a durable immunological imprint persisting at six months, characterized by persistent immune gene modifications and metabolic reprogramming. Findings also reveal greater omics crosstalk in response to BNT162b2, which is related to the enhancing of its immunogenicity and durability. Furthermore, predictive modeling demonstrated that baseline multi-omics profiles forecast vaccine-specific performance (AUC: 0.65–0.84). Collectively, this work provides pivotal insights into the mechanistic distinctions between vaccine modalities. Our results advocate integrating multi-omics data to tailor vaccination strategies to individual immune profiles, enhancing public health outcomes.

Introduction

The deployment of vaccines marked a pivotal turning point in the battle to control the COVID-19 pandemic (1). To date, billions of vaccine doses have been administered worldwide, significantly mitigating the spread and severity of SARS-CoV-2 (2). Various types of COVID-19 vaccines have been developed and utilized across different countries, with inactivated virus vaccines and mRNA vaccines being the most prominently used during the early stages of the outbreak (3–5). Inactivated virus vaccines have a long history, dating back to their discovery by Louis Pasteur, and have been used for over a century (6–8). These vaccines are based on viruses that have been killed or inactivated, preventing them from replicating and causing disease while still being able to stimulate an effective immune response.

On the other hand, mRNA vaccines including Moderna mRNA-1273 and BioNTech/Pfizer BNT162b2 were first widely deployed during the COVID-19 pandemic (9). These vaccines work by delivering messenger RNA (mRNA) to instruct cells to produce a protein that triggers an immune response. Specifically, mRNA vaccines for COVID-19 encode the spike protein of the SARS-CoV-2 virus, educating the immune system to recognize and fight the virus if encountered (10). Studies comparing the immunogenicity of these two vaccine types have shown that mRNA vaccine induces higher levels of antibodies than the inactivated virus vaccine, although the T cell responses remain comparable between them (11). However, it has also been observed that the side effects following immunization with mRNA vaccines tend to be more pronounced than those seen with inactivated virus vaccines (12). This suggests that the two types of vaccines may elicit different host responses after their administration into the body.

Systems vaccinology is an emerging field that leverages comprehensive, high-throughput technologies to analyze the host's response to vaccination (13). This approach involves testing blood samples from vaccinated individuals to gather data from multiomic assays, such as transcriptomics, metabolomics, and lipidomics (14). These data can elucidate the pathways and mechanisms activated by vaccination and can provide insights into the host's immune response. Furthermore, analyzing the condition of individuals prior vaccination and associating these data with their subsequent immune responses can offer valuable information on vaccine efficacy and safety (15).

While most current studies focus on evaluating the immunogenicity and efficacy of vaccines, our understanding of the body's nuanced responses to these vaccines, especially mRNA vaccines, is still limited. To differentiate the host responses elicited by inactivated virus and mRNA vaccines, we conducted a cohort study in which participants who received either type of vaccine was recruited in Hong Kong. Blood samples were collected longitudinally before, one month, and six months after vaccination. Multiomic assays including transcriptomic, metabolic, lipidomic and cytokine profiling were then performed to analyze the data comprehensively. By comparing the host responses, this study aims to provide a deeper understanding of

the mechanisms underpinning the host responses to these two vaccine types and to inform future vaccine development and deployment strategies.

Results

Study design

A comprehensive summary of the study design and the flow of data analysis is shown in Figure 1. In this prospective study, 553 adults over 18 years old were recruited in Hong Kong SAR, China at vaccination centers at the Chinese University of Hong Kong Medical Centre, Prince of Wales Hospital and Kowloon Bay Vaccination Center between March 10 and August 31, 2021 as previously described (274 for BNT162b2 and 279 for CoronaVac group) (11,12). Heparin blood samples were collected from each participant at before (baseline), 1 month and 6 months after the second dose of vaccination. Additionally, demographic information of the participants was collected which includes age, gender, history of underlying disease, lifestyle factors (smoking, drinking, and regular exercise), vaccination history, and post-vaccination side effects through self-administered questionnaires (Table 1). Most of the samples from our cohort were mainly collected prior to the large Omicron BA.2 outbreak in 2022, a time when the overall infection rate in Hong Kong was very low (16). These samples were further confirmed as sero-negative to the ORF8, an indication of lacking breakthrough infection.

To gain insights into the molecular mechanisms underlying the responses to the vaccines, comprehensive analyses were conducted on samples, including metabolomic-, lipidomic-, transcriptomic-, and cytokine profiling. The antibody level that is against the ancestral strain (wild type Wuhan) of SARS-CoV-2 was measured using both surrogate virus neutralization test (sVNT) and ELISA using the receptor binding domain (RBD). All participants did not receive any COVID-19 vaccine prior to the recruitment and were naive to SARS-COV-2 infection as confirmed by our serology ELISA test against the ORF8 antigen.

The demographic characteristics of the two vaccine groups are compared in Table 1. Participants who received the CoronaVac vaccine were significantly older (median age: 54.0) than those vaccinated with BNT162b2 (median age: 46.0; $p<0.001$). Additionally, when compared against those who received the CoronaVac, a significantly larger proportion of BNT162b2 recipients reported a history of vaccinations against hepatitis A/B (53.7% vs 29.4%, $p<0.001$), mumps (20.6% vs 6.1%, $p<0.001$), pneumococcal conjugate (7.7% vs 3.2%, $p=0.024$), and hemorrhagic fever (1.8% vs 0%, $p=0.055$). These results were included as confounders in the statistical analysis for the comparison of the antibody level and host responses between the two vaccine groups.

Antibody responses induced by the two COVID-19 vaccines

Both vaccines elicited specific neutralizing antibody responses to the spike protein of wild-type SARS-CoV-2 one month after two doses of vaccination. However, CoronaVac recipients exhibited significantly lower antibody levels compared to those immunized with BNT162b2 (sVNT: 55.02% vs 96.08%, $p<0.001$; ELISA: 2.46 vs 1.01 area under curve (AUC), $p<0.001$) (Table 1, Fig. S1a). At six months of post-vaccination, those who had received BNT162b2 maintained a significantly higher level of neutralizing antibodies, whereas the antibody level in most of the participants with CoronaVac diminished to baseline. Antibody waning was commonly observed after receiving the COVID-19 vaccine. Instead of comparing the antibody level at the 6-month time-point, we also determined the durability of the antibody level by defining the relative drop of the antibody level from 1 month to 6 months after vaccination. BNT162b2 exhibits significantly longer durability than CoronaVac (Fig. S1b). Importantly, the durability of the antibody is independent of sVNT levels at M1 in the CoronaVac group ($R=0.02$, $p=0.818$, Fig. S1c). Consistent with the previous studies (*17, 18*), recipients of the BNT162b2 vaccines reported a higher incidence of adverse events, including fever, headache, body ache, fatigue, loss of appetite, and injection site pain, compared to the CoronaVac vaccines (Supplementary Table 1). Interestingly, females showed higher sVNT levels regardless of the vaccine type they had received (Supplementary Table 2). Moreover, a notable elevation in sVNT levels was observed in the CoronaVac cohort among participants with a history of influenza vaccination (Supplementary Table 2).

Differential expression of multi-omics profiles between the CoronaVac and BNT162b2 vaccines

The significant differences in the antibody levels and side effects between the two vaccine groups led us to hypothesize that host might respond differently to the two vaccines as reflected from the multiomic data. As participants were all recruited from Hong Kong during the same period of time while everyone could choose either one of the vaccines by their will, we thus analyzed the metabolomic, lipidomic, transcriptomic using linear mixed-effect modeling with Limma/Voom to model the counts of each host transcript/the $\log_2$ abundance of metabolites/lipids, after batch effect removal (Figure S2, See "Methods" section). This approach enabled us to attempt to delineate the omics-level changes associated with SARS-CoV-2 vaccination.

Comparing the results at baseline and one/six months following administration of the second dose, BNT162b2 vaccinees exhibited more pronounced changes in their omics profiles than those who had received CoronaVac (Figure 2). At one-month post-vaccination, we identified changes in transcripts (563 in BNT162b2 vs. 45 in CoronaVac), metabolites (50 vs. 25), lipids (48 vs. 14), and cytokine (22 vs. 22) between the two vaccine groups (Figure 2b-e, Fig. S3). Notably, 19 differentially expressed genes (DEGs), such as GIMAP1-GIMAP5, ZBED6, TMX2-CTNND1, LIMS3, were identified post-vaccination across both vaccine groups (Figure 2b, Fig. S4a, Supplementary table 3-4). Gene expressions related to cellular progression and immunological functions exhibited more extensive changes following BNT162b2 immunization compared to CoronaVac (Fig. S5). Interestingly, we observed a downregulation in several genes linked to neural functions post-BNT162b2 vaccination, including FOSL1, AREG, LHX3, and EP300. No significant differences in the expression of these genes before the first dose were observed between the two vaccine groups (Fig. S6). Furthermore, levels of D-Sphingosine and Sphingomyelins (SMs), crucial components in the ceramide synthesis pathway and markers of cell signaling and apoptosis (*19*), were significantly reduced after vaccination of BNT162b2 (Fig. S4b). This reduction was more pronounced in individuals vaccinated with BNT162b2, especially for SM(d39:1) (rep), SM(d40:1), SM(d41:1) (rep), SM(d42:1) (rep), and SM(d44:1). Additionally, the abundance of certain metabolites with known neurological roles, such as D-Erythro-sphingosine 1-phosphate, choline, D-Sphingosine, and Oleoylethanolamide, decreased following two-dose BNT162b2 (Fig. S4b).

Further, we performed Gene set enrichment analysis (GSEA) and over-representation analysis on transcriptomic and metabolomic data, respectively. While both vaccines triggered various biological function pathways, they did so in different directions, including cellular processes, immune system processes, and metabolic processes (Figure 3a). Intriguingly, all significant mapped pathways in CoronaVac recipients (N = 93) were activated, contrasting with more suppressed pathways, particularly in cellular processes, found at BNT162B2 (Figure 3a-b; Fig. S4). Pathways in neural function were suppressed including neuron projection morphogenesis, cell morphogenesis involved in neuron differentiation, synaptic signaling, dendrite development, dendritic tree, regulation of neuron projection development, and regulation of nervous system (Figure 3b, Supplementary Table 5). We next conducted blood transcription module (BTM) analysis on the differentially expressed genes between baseline and one-month post-vaccination. This analysis revealed a distinct polarization in the enrichment of immunological pathways. Specifically, the BNT162b2 vaccine was associated with an upregulation of pathways tied to adaptive immunity, including T cell activation and B cell enrichment, along with signatures suggestive of natural killer (NK) cell involvement (Figure 3c, Supplementary Table 6). For instance, pathways such as T cell activation and signaling, and enrichment in B cells, were markedly upregulated in the BNT162b2 group compared with the CoronaVac group (Figure 3c, Supplementary Table 6). Conversely, the CoronaVac vaccine was linked with a modest elevation in pathways associated with innate immunity, including monocyte and dendritic cell signatures (Figure 3c, Supplementary Table 6). Furthermore, analysis indicated a higher induction of antiviral responses, including toll-like receptor (TLR) and inflammatory signaling pathways in the BNT162b2 group, implicating a robust upregulation of type I interferon responses (Figure 3c, Supplementary Table 6). Additionally, the observed pathway changes were widespread among the participants, rather than being rare occurrences. More than 75% of those vaccinated against SARS-CoV-2 demonstrated upregulation in a minimum of 10 pathways associated with cellular processes and immunological functions. In contrast, neural function pathways showed a significant decline in frequency, suggesting a predominance of negative effect sizes within most BNT162b2 recipients (Figure 3d). The sensitivity analysis that directly compared the longitudinal omics alteration between the two vaccine groups identified 769 differentially expressed (DE) markers at M0-M1 (706 Transcripts, 38 metabolites, 25 lipids, 12 cytokines), Additionally, those markers playing immune functions, such as THBD, TNFAIP3, EGR3 and NR4A3, and playing neurological function (e.g.,

ATXN1L, ZMIZ1, NOTCH2) were overlapped with previous markers (Fig. S7). Metabolically, compared with CoronaVac, BNT162b2 vaccinees showed more elevated levels of a wide range of metabolites and lipids (Fig. S4b), many of which belonged to ABC transporters, purine metabolism, and sphingolipid signaling, and necroptosis pathways (Figure 3e). Overall, our data highlighted a divergent impact of SARS-CoV-2 vaccines on the immune system, with BNT162b2 enriching adaptive immune signatures and suppressing neural functions, while CoronaVac enriching innate immune signatures.

Convergence on neural function alterations and immune function changes

Further analysis of the various omics data identified key neural function-related networks and the potential mechanisms (Fig. S8a). Notably, gliclazide, known for its neurological effects (*20*), exhibited a positive correlation with 7 out of 11 neural function-associated genes (Fig. S8a). D-Sphingosine, which is implicated in neuronal survival and inflammation, also demonstrated significant interconnectivity within the network (Fig. S8a). STRING (Search Tool for Retrieval of Interacting Proteins/Genes) analysis of the gene network further substantiated these findings (Fig. S8b); These 11 genes were downregulated, predominantly in pathways related to nervous system development. And the alterations were independent of immunogenicity (Supplementary Table 7). This suggests a downregulation of neural functions following BNT162b2 vaccination. Additionally, the enrichment of receptor tyrosine kinase signaling pathways highlights a potential impact on neural development and function. For immune function assessment, a more intricate network was identified between the BNT162b2 and CoronaVac vaccines. STRING analysis of the BNT162b2 specifically revealed extensive interconnectivities among genes such as NR4A3 and MAFF, suggesting a potential modulation of immune functions following vaccination (Fig. S9). These insights provide a comprehensive view of the neural function and immune function alterations associated with different SARS-CoV-2 vaccines at the omics level (Supplementary Table 7-9).

Persistent changes in immunological functions and neural signatures

We then assessed whether the omics changes at 1-month post-vaccination could further persist in 6-month post-vaccination. We found that at 6 months post-vaccination, significant changes in metabolites, lipids persisted more in the BNT162b2 group (metabolites: 31/50, lipids: 23/48) than in the CoronaVac group (metabolites: 8/25, lipids: 2/17). However, only 52 out of 563 differentially expressed transcriptomic features between M0 and M1 remained persistently altered in BNT162b2, compared with 21 out of 45 in the CoronaVac group (Figure 2b-d). In terms of neural function, 8 out of 11 genes showed sustained downregulation, and within immunological function, 6 out of 13 genes remained downregulated (Figure 4a). Interestingly, several persistently downregulated transcriptomic markers, including PGLYRP1 (associated with cellular process), as well as LHX3 and EGR2 (involved in neural processes), were positively correlated with the relative decline in sVNT values between 1 and 6 months post-vaccination (Figure 4a). Pathway enrichment analyses revealed consistent changes in cellular processes, immunological functions, and neural functions, as well as in the biosynthesis of unsaturated fatty acids (Fig. S10, Supplementary Table 5-6). Notably, after 6 months of vaccination, BNT162b2-induced T cell activation and associated immune pathways showed a shift towards suppression (Fig. S10). Conversely, in the CoronaVac cohort, only one immunological pathway, Monocyte surface signature, exhibited such suppression (Fig. S10). Additionally, sphingolipid-related pathways continued to be enriched post-BNT162b2, with some neural-related metabolites returning to pre-vaccination levels (Fig. S10).

Different pattern of relationships between the crosstalk of multi-omics and immunogenicity in the two vaccine groups

Our investigation into the differences of vaccine-induced host responses revealed a differential omics crosstalk effect on immunogenicity and durability between BNT162b2 and CoronaVac (Figure 5, Fig. S11). For durability, both vaccine groups displayed more intense networks associated with longer durability (BNT162b2: 500 vs. 6 associations; CoronaVac: 105 vs. 14 associations, Figure 5, p value < 0.05). Certain correlations, particularly those involving LPCs-genes crosstalk, persistently appear in the network associated with longer durability compared to the network for shorter durability (Fig. S11). Regarding immunogenicity, in BNT162b2, a slightly denser network was observed in the higher immunogenicity group (top quantile of antibody level at M1) than non-higher immunogenicity group others (559 vs. 530 associations). Among those with robust immune responses, there was a notable complexity within the gene-metabolite/lipidomic network, especially for Lyso-

phosphatidylcholine (LPCs) and SMs (Figure S11). Conversely, the CoronaVac group did not demonstrate such marked differences in network intricacy (90 vs. 93 associations) (Fig. S11). Altogether, our findings provide insights into the specific multi-omics interactions at play, particularly highlighting the role of lipids in both immunogenicity and durability of the BNT162b2 vaccine.

Omics profile elucidates variation in immune responses to SARS-CoV-2 vaccines

In the BNT162b2 group, especially among those with robust immune responses, there was a notable complexity within the gene-metabolite/lipidomic network. Conversely, the CoronaVac group did not demonstrate such marked differences in network intricacy (Fig. S12a). To further identify the omics landscape in the variation of immune outcomes, we employed differential analysis using limma/voom for RNAseq and GLM for metabolites/lipids. These profiles suggested the diverse immunogenic responses elicited by the BNT162b2 and CoronaVac vaccines (Fig. S12b-d, Fig. S13, Supplementary Table 10-11). Specifically, BNT162b2 is associated with unique gene sets linked to immunogenicity (33 genes) and durability (47 genes), alongside a distinct metabolomic profile characterized by 14 metabolites related to immunogenicity and 22 to durability. Conversely, the CoronaVac group displayed a more limited metabolomic variance and a differing gene expression pattern, with 59 genes tied to immunogenicity and 29 to durability. Notably, LPC (18:2) emerged as the sole feature correlating with both heightened immunogenicity and durability in the BNT162b2 cohort (Fig. S12e).

Next, we investigated the immune-related genes and their interaction that are likely to contribute to higher immunogenicity and longer durability. We built gene-gene interaction networks using STRING database. Interestingly, most of the transcripts enriched in higher immunogenicity group clustered together and were involved in immune system, especially for innate immune response (Supplementary Table 12). For example, in BNT162b2 group, three clusters were fitted by the genes which were associated higher immunogenicity. In the first cluster (11 genes), genes such as LTF (lactotransferrin), DEFA1B, and DEFA4 (defensin alpha 1B and 4) were centered that are interconnected and have been identified to play significant roles in the immune response. Specifically, these genes are involved in the innate immune response within mucosal tissues and contribute to antimicrobial and antibacterial humoral responses. This clustering further suggested a coordinated action of these genes in bolstering the immunogenicity against BNT162b2. Moreover, the genes associated with longer durability may regulate the durability by cellular responses. Seven genes (e.g., SYT1, SCN5A, SLC5A11, MAP6, and KCNQ3) from the first cluster involved in calcium binding and ion transport activities. These insights into gene-gene interactions enhance our understanding of the molecular basis for the sustained immunogenic effects of the BNT162b2 vaccine, with further comparative details available in the supplementary materials for the CoronaVac group.

Lipids and metabolites also exerted modulation in the stronger immune responses. For BNT162b2, the lipidomic features, such as LPC (20:2) and PC (34:1), exhibit significant associations with higher immunogenicity, while a range of metabolites, including sphingomyelin and glycerophospholipids, demonstrate strong links to durability (Fig. S12e). This could suggest that lipid metabolism plays a crucial role in the immune response and longevity of vaccine efficacy for BNT162b2. In contrast, the CoronaVac vaccine shows a narrower spectrum of metabolites associated with immunogenicity, such as LPC (14:0), with fewer lipidomic features implicated. Durability is associated with a distinct set of metabolites, predominantly involved in energy and amino acid metabolism. The differences in the profiles of immunogenicity and durability between the vaccines might further imply why BNT162b2 elicits a more robust alteration in metabolic pathways, which correlates with its reported higher efficacy and longer-lasting immune response as compared to CoronaVac.

Multi-omics prediction modeling

We next sought to determine the relative contribution of pre-vaccination omics data in predicting post-vaccination immune responses (immunogenicity and durability) using machine learning. The models achieved the following areas under the curve: 0.73 (higher immunogenicity in BNT162b2), 0.72 (lower immunogenicity in BNT162b2), 0.72 (higher immunogenicity in CoronaVac), 0.80 (lower immunogenicity in CoronaVac), 0.84 (longer durability in BNT162b2), 0.65 (longer durability in CoronaVac) (Figure 6, Supplementary Table 13). Transcriptome was the most frequently selected omics type among the six independent models, indicating its significant role in predicting vaccine immune responses. Estradiol valerate were found

both in higher immunogenicity (top quantile) and lower immunogenicity (lowest quantile) prediction models in CoronaVac. Statistically, higher levels of estradiol valerate at baseline are associated with higher immunogenicity. However, a few common features among the six models may suggest potential different mechanisms in both the intensity and durability of immune pathways across different vaccine types. Collectively, we employed a machine learning approach that integrates multi-omics and clinical features from pre-vaccination candidates to derive predictors of immune responses to SARS-CoV-2 vaccination.

Discussion

Previous studies have explored omics perturbations in individuals receiving COVID-19 vaccines (21–23) but few have integrated transcriptomic, metabolomic, lipidomic, and cytokine profiling to examine host responses over time. Moreover, study on head-to-head comparison of host responses between different vaccines is limited. Our research delineates clear distinctions in the longitudinal omics profiles between two COVID-19 vaccines: the mRNA-based BNT162b2 and the inactivated virus vaccine CoronaVac. Specifically, BNT162b2 predominantly enhances adaptive immunity, while CoronaVac primarily activates innate immune pathways. Additionally, our analysis indicates a notable suppression of metabolomic functions after administering two doses of BNT162b2. Additionally, our results demonstrate significant omics crosstalk in response to BNT162b2, underscoring its potential for greater immunogenicity and durability. Crucially, baseline omics profiles can predict immune responses post-vaccination, providing a robust foundation for tailoring future personalized vaccine strategies.

During the prolonged period post-vaccination, host responses to the BNT162b2 vaccinee converged on adaptive immunity, while CoronaVac mainly activated innate immune pathways, some of which persisted up to 6 months post-vaccination. Concurrently, more cellular pathways were downregulated in the BNT162b2 group than in the CoronaVac group. Previous studies have shown that BNT162b2 also activates innate immunity during the first week after vaccination. The switching of immune functions can be attributed to the differences in the nature of the two vaccines. Compared to the inactivated virus vaccine, mRNA of the spike protein was still detected in the axillary lymph nodes within 30 days post-vaccination. The mRNA vaccine encodes a SARS-CoV-2 spike protein with conformationally stabilized pre-fusion mutations (24–26). The high stability of these modified spikes may lead to antigen persistence, resulting in prolonged stimulation of B and T cells, and higher antibody levels, as well as increased affinity and avidity, which are commonly observed in mRNA vaccines but not inactivated virus vaccines (27–29). The COVID-19 vaccine was also reported to mitigate disease severity by attenuating the inflammatory responses (30). *Watanabe et al.* also suggested the changes in genes associated with B cells, T-cell specific related genes, and both T- and B cell-related genes. Additionally, single-cell transcriptomics by *Zhou et. al* demonstrated that BNT162b2 vaccination activates antiviral and inflammatory responses in monocytes and dendritic cells (31). These transcriptional alterations, encompassing immune system regulation, metabolism, hemostasis, and platelet activation, further corroborate our findings, suggesting a more nuanced understanding of vaccine-induced immune dynamics (32).

Our analysis also revealed an enrichment of pathways annotated with neurological functions following BNT162b2 vaccination. For example, we identified significant changes in gene expressions, including the downregulation of FOSL1, AREG, LHX3, and EP300, which are associated with various neurological functions and clinical conditions. Additionally, changes in metabolites like D-Erythro-sphingosine 1-phosphate, choline, D-Sphingosine, and Oleoylethanolamide are involved in neural signaling and neuroprotection. Furthermore, our Co-Abundance Networks of Features further indicated intricate interactions between these molecular changes. These findings suggest a complex interaction of vaccine-induced effects on neural pathways not captured in the acute phase, which Watanabe Y. et al.'s research focused on, analyzing data only one day post-vaccination (32). Our longer timeline and different analytical methods, including distinct enrichment approaches and databases, may contribute to the disparities in findings. Interestingly, both patients with COVID-19 infection and individuals who have received COVID-19 vaccines may experience a range of neurologic consequences, such as headaches, fatigue, and dizziness (33–35). Recent evidence from plasma proteomics also suggests that SARS-CoV-2 infection may increase the risk of developing Alzheimer's disease in the future (36). In our cohort, changes in neurological-related were observed in 25–50% of participants, with some alterations persisting up to six months post-vaccination. However, current evidence does not

directly support a clinical risk of neurological diseases associated with receiving mRNA vaccine. We hypothesize that while the vaccine may trigger these molecular-level alterations, only a small proportion of vaccinees might further develop clinical symptoms, indicating that these molecular changes do not universally translate to clinical manifestations. Of note, the measurements were conducted in peripheral blood, which may not fully reflect changes within neural tissues and the central nervous system. Further research should explore the mechanisms via experimental studies validating these differences and assess the long-term clinical and subclinical neurological impacts of mRNA vaccines, to ensure safety and efficacy across diverse populations. Such insights are crucial for optimizing vaccine strategies and understanding the full scope of vaccine-induced physiological changes.

The complex interactions involving PCs, LPCs, and SMs observed in BNT162b2 recipients suggest that these lipids significantly enhance the vaccine's immunogenicity and durability. These molecules are crucial for activating humoral and cellular immune responses, which may impact the robustness of vaccine responses (37–39). Despite less complexity in CoronaVac responses, the consistent presence of SMs-cytokines crosstalk in networks associated with longer durability highlights their essential role in maintaining immune memory. The involvement of sphingolipids may be associated with trained immunity after vaccination (40). Understanding these interactions can help explain the variation of immune responses, thereby guiding the optimization of vaccine formulations for enhanced immunogenicity and durability.

Integrating multi-omics datasets provides a comprehensive view that not only helps predict vaccine efficacy but also enhances our understanding of why certain individuals exhibit stronger responses to vaccines than others, as suggested by our co-occurrence network and random forest model. Growing interest in the concept, baseline may predict outcome inspires researchers to investigate the role of baseline omics variation in diverse global populations, offering avenues for optimization of current vaccination strategies (41). *Feargal et al.* found that pre-vaccination immune state is predictive of antigen-specific responses to the BNT162b2 vaccine (22). Our earlier study also suggested that specific gut microbiota markers could predict immune responses, including immunogenicity and durability, following COVID-19 vaccines (17, 42). Moreover, studies have shown that integrating genomic and proteomic data can enhance predictions of vaccine-induced antibody responses. For instance, Moorlag et al. showed the predictive power of multi-omics analysis in understanding vaccine responses by demonstrating how epigenetic profiles can forecast trained immunity after BCG vaccination (43). Similarly, *Bakker et al.* utilized integrative multi-omics data to predict cytokine responses following vaccination. They found that a combination of genomic, transcriptomic, and other biological data could be used to predict immune responses, highlighting the value of a multi-omics approach to understand and predict vaccine efficacy (44). Furthermore, maintenance hemodialysis (HD) patients exhibit prolonged myeloid cell activity and a delayed return to baseline after vaccination compared to healthy controls (45). In terms of cellular immune responses, HD patients also demonstrate delayed T-cell activation but maintain strong cellular immune responses over time, secreting higher levels of IFN-γ and IL-2 after vaccinations, emphasizing the need for continuous monitoring and tailored vaccination strategies to optimize outcomes in populations with different vulnerabilities.

A major strength of our study is the longitudinal design and a large sample size with paired data from the same subjects before and after COVID-19 vaccines which allows inference of directionality. Also, we used a GSEA with BTM analysis and Hallmark analysis, rather than traditional enrichment methods. This approach enabled us to construct a more systematic map of gene expression changes, thereby strengthening the comparability and robustness of our results. However, several limitations should be noted. First, the interpretation of neural signatures is constrained by reliance on blood-derived omics data; many identified genes and metabolites have pleiotropic roles and are active in both immune and neural pathways, making precise attribution difficult. Secondly, the absence of early post-vaccination timepoints (e.g., day 1–14) limits our ability to capture acute innate and early adaptive responses. Moreover, the COVID-19 vaccines not only induce humoral response but also trigger cellular immunity especially the activation of CD4 and CD8 T cells (11). In this study, we mainly focused on analyzing the association between the results from multi-omic data and the antibody level. Thus, investigation into the interplay between the host responses and the T cell functions against SARS-CoV-2 is still not known and requires further examination. Moreover, the transcriptomic data was mainly obtained from the bulk analysis of PBMCs in which neutrophils were excluded due to the cell isolation procedure. Although transcriptional changes in specific populations such as T, B cells can be determined by single-cell sequencing, such approach will not be financially feasible at this stage due to

the relatively large sample size from our cohort. Finally, while we observed associations between omics profiles and immune responses, the establishment of causality is precluded by limitations including unmeasured confounders and potential selection bias (e.g., participant self-selection into groups). Future endeavors aim to validate and expand upon these findings within new, independent experimental studies to enhance our understanding of the interplay between omics and immune responses.

Previous studies focused on predicting disease progression or antibody expression in individuals (46). Our model extends this to evaluate individuals at a population level, assessing the intensity and durability of their immune responses over time and identifying associated baseline characteristics. The six models achieved an average AUC of 0.74, indicating acceptable performance (47). However, due to the complexity and acquisition challenges of multi-omics data, the current model is limited by the cohort scale and would benefit from further validation with larger datasets. Nevertheless, multi-omic data integration may help tailor vaccination schedules or booster compositions for elderly and immunocompromised individuals, whose immune landscapes differ from those of younger or healthy populations. In addition, by revealing signatures associated with robust cross-variant protection, these data can guide the design and optimization of next-generation vaccines that maintain efficacy against emerging SARS-CoV-2 variants and other evolving pathogens. Such precision-guided approaches will be instrumental in advancing equitable and adaptive vaccination programs in the post-pandemic era.

In summary, our research brings forth the concept how two distinct COVID-19 vaccines, mRNA and inactivated virus, induce divergent changes within the omics landscape, emphasizing that baseline omics variations significantly influence immune responses. Moving forward, the multiomics integration in vaccination strategies promises to maximize efficacy and minimize adverse reactions, thereby aligning modern vaccine science with the principles of personalized medicine.

Methods

Study design and participants.

Our previous study recruited 726 adults over 18 years old who received either two doses of BNT162b2 or CoronaVac in Hong Kong SAR, China, at vaccination centers at the Chinese University of Hong Kong Medical Centre, Prince of Wales Hospital and Kowloon Bay Vaccination Center between March 10 and August 31, 2021 as previously described (*12*). The participants received two doses of either BNT162b2 (21-day dose interval) or CoronaVac vaccine (28-day dose interval), according to the manufacturer's recommendation. Those with a confirmed history of COVID-19 at the day of sampling were excluded. Among all, samples from 553 (274 for BNT162b2 and 279 for CoronaVac group) who provided blood at before, 1- and 6-month after the second dose of vaccination were included for this study. This study was approved by the Joint Chinese University of Hong Kong- New Territories East Cluster Clinical Research Ethics Committee (reference number 2020.229) and conducted in compliance with the Declaration of Helsinki. All participants provided written informed consent.

Sample and epidemiological data collection. Epidemiological, demographic, and clinical presentation data were collected from each participant through a self-administered questionnaire during their first visit after vaccination. 10 ml of heparin blood samples were drawn at before, 1- and 6-month after vaccination. The blood was first centrifuged at 3000 × g for 10 minutes at room temperature to obtain plasma. Peripheral blood mononuclear cells (PBMCs) were isolated utilizing FicollPaque Plus medium (Cytiva, Amersham, UK), following the manufacturer's instructions. Subsequently, one million of PBMC were immediately mixed with 1ml of TRIzol®. Both plasma samples and the TRIzol® were stored at -80°C until further analysis. Sample size for each assay and timepoint was detailed in Supplementary Table 14.

Surrogate virus neutralization test (sVNT). SARS-CoV-2 surrogate virus neutralization test kits were obtained from GenScript, Inc., NJ, USA, and the tests were carried out according to the manufacturer's instructions. The test sera (10 µl) and positive and negative controls were diluted 1:10 or 1:200 and mixed with an equal volume of horseradish peroxidase (HRP) conjugated to wild wild-type strain of SARS-CoV-2 spike receptor binding domain (RBD) (6 ng) and incubated for 30 min at 37°C. A 100 µl volume of each mixture was added to each well on the microtiter plate coated with ACE-2 receptor. The plate was sealed and incubated at room temperature for 15 min at 37°C. Plates were then washed with wash solution and tapped

dry, and 100 µl of 3,3’,5,5’-tetramethylbenzidine (TMB) solution was added to each well and incubated in the dark at room temperature for 15 min. The reaction was stopped by addition of 50 µl of Stop Solution to each well and the absorbance read at 450 nm in an ELISA microplate reader. The assay validity was based on values representing optical density at 450 nm (OD450) for positive and negative results falling within the range of recommended values. On the basis of the assumption that the positive and negative controls gave the recommended OD450 values, percent inhibition of each serum was calculated as follows:

$$Percent\ inhibition = 1 - \frac{sample\ OD\ value}{negative - control\ OD\ value} \times 100$$

Percent inhibition values of 20% or more are regarded as positive results.

ELISA for spike RBD specific antibodies. A 96-well enzyme-linked immunosorbent assay (ELISA) plate (Nunc MaxiSorp, Thermo Fisher Scientific) was first coated overnight with 100 ng per well of purified recombinant protein in PBS buffer. The plates were then blocked with 100 µl of Chonblock blocking/sample dilution ELISA buffer (Chondrex Inc, Redmon, US) and incubated at room temperature for 1 hour. Each human plasma sample was serial diluted from 1:100 to 1:1000000 in Chonblock blocking/sample dilution ELISA buffer. Each diluted sample was then added into the ELISA plates for a two-hour incubation at 37°C. After extensive washing with PBS containing 0.1% Tween 20, each well in the plate was further incubated with the anti-human IgG secondary antibody (1:5000, Thermo Fisher Scientific) for 1 hour at 37°C. The ELISA plates were then washed five times with PBS containing 0.1% Tween 20. Subsequently, 100 µl of HRP substrate (Ncm TMB One; New Cell and Molecular Biotech Co. Ltd, Suzhou, China) was added into each well. After 15 min of incubation, the reaction was stopped by adding 50 µl of 2 M H2SO4 solution and analyzed on an absorbance microplate reader at 450 nm wavelength. The antibody level of each sample was determined as the area under curve (AUC) using the software of Prism 8.0.

Outcomes analysis. The immune outcomes, including both immunogenicity and durable immunogenicity, were categorized based on the sVNT levels. Those in the top quartile for sVNT titers post-vaccination were categorized under 'higher immunogenicity' due to their elevated neutralizing antibody levels, indicating a more potent immune response (*17*). For evaluating 'durability' of the immune response, an index was computed utilizing the sVNT levels at two time points: at one month (M1) and six months (M6) post-vaccination. The durability index was defined as the relative decline in sVNT levels over time, calculated by the formula: (sVNT at M1 - sVNT at M6) / sVNT at M1. This ratio reflects the percentage decrease in neutralizing antibody levels, serving as a surrogate for the longevity of the vaccine's efficacy (*42*). To classify individuals into 'longer durability' versus 'non- longer durability' groups, the top quantile of the durability index across the cohort served as the cut-off points.

Cytokine quantification. Inflammation cytokines including IL-1β, IL-17A, IL-18, IL-23, TNF-α, IFN-γ, IFN-α, IL-12, IL-8, IL-10, IL-33, CX3CL1, CCL11, MIP-1α, MIP-1β, CCL20, CCL2, RANTES, CXCL10, CXCL9, CCL17, CXCL5, and CXCL11 etc. from the plasma samples were quantitatively determined by flow cytometry-based immunoassay (LEGENDplexTM kits, Biolegend) according to manufacturer’s protocol. The amount of cytokine (pg/mL) in samples was quantitated using flow cytometry (AttuneNxT) and analyzed by FlowJo v10.

Metabolomics and Lipidomics analysis from plasma samples. For metabolomic analysis, deactivated plasma samples were extracted using a methanol extraction solution. The supernatants contained metabolites of diverse chemical natures. Four platforms were utilized for non-targeted metabolomics. The supernatant was divided into four fractions: two for analysis using two separate reverse-phase/ultra-performance liquid chromatography (RP/UPLC)-MS/MS methods with positive ion-mode electrospray ionization (ESI), one for analysis using RP/UPLC-MS/MS with negative-ion mode ESI, and one for analysis using hydrophilic interaction liquid chromatography (HILIC)/UPLC- MS/MS with negative-ion mode ESI. The QE mass spectrometer analysis alternated between MS and data-dependent MS2 scans using dynamic exclusion. The scan range was 70-1,000 m/z. After raw data pre-processing, peak finding/alignment, and peak annotation using in-house software, metabolites were identified by searching an in-house library containing more than 3,300 standards, with library data entries generated from running purified compound standards through the experimental platforms.

For lipidomic analysis, plasma samples were homogenized in a mixture of methanol, chloroform, and water to extract lipids. The resulting lipid extracts were dried under nitrogen gas and re-dissolved in a mixture of isopropanol, acetonitrile, and water. The lipid extracts were then analyzed using high-performance liquid chromatography coupled with tandem mass spectrometry (HPLC-MS/MS). Chromatographic separation was achieved using a C18 column, with a gradient elution of solvent A (60% acetonitrile in water with 10 mM ammonium formate and 0.1% formic acid) and solvent B (90% isopropanol with 10 mM ammonium formate and 0.1% formic acid). Mass spectrometry was conducted in both positive and negative ion modes to ensure broad coverage of lipid species. Data acquisition was performed using a TripleTOF 5600+ mass spectrometer (Sciex) with a data-dependent acquisition mode to maximize lipid identification and quantification. Lipid species were identified and quantified using LipidViewTM software, referencing established lipid databases. Quality control (QC) samples and internal standards were included to monitor instrument performance and ensure data accuracy.

Peripheral blood transcriptomics. One million of PBMC cells was solubilized in 1 ml of Trizol. RNA was then isolated from each sample using the miRNeasy Mini kit (Qiagen) and 10 ng of total RNA was used as input for cDNA synthesis using the Clontech SMART- Seq v4 Ultra Low Input RNA kit (Takara Bio) according to the manufacturer's instructions. Amplified cDNA was fragmented and appended with dual-indexed bar codes using the NexteraXT DNA Library Preparation kit (Illumina). Libraries were validated by capillary electrophoresis on an Agilent 4200 TapeStation, pooled at equimolar concentrations, and sequenced on an Illumina NovaSeq at 150SR, yielding at least 20 million reads per sample. Sequencing results were demultiplexed and converted to FASTQ format using Illumina bcl2fastq software (Illumina). The sequencing reads were adapter and quality trimmed and then aligned to the human genome (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.26/; version hg38) using the splice-aware STAR aligner50.

Statistical and computational analysis. All statical and computational analysis was performed in R version 4.3.0 (2023-04-21) and Python version 3.11.5. Continuous variables (ie, sVNT) were reported as median and IQR, and categorical data were presented as proportions. To establish whether there was an imbalance in participants' demographic characteristics between the two types of vaccines used for the initial doses, we used the Wilcoxon rank sum test for continuous variables and Fisher's exact test for the categorical variables. For the longitudinal outcome comparisons, we used the Friedman test for paired data followed by post-hoc analysis with pairwise Wilcoxon rank sum test for paired data for multiple comparisons.

Data preprocessing

Normalization and filtering of untargeted metabolomics data. Quality control samples, comprised of an equal aliquot of all test samples, were prepared and inserted in an interval of ten test samples throughout the assay. This procedure was performed to monitor the stability of instrument and to normalize the variations during the runs. Raw LC-MS/MS files were processed using Compound Discoverer 3.1 (Thermo Fisher Scientific) and metaZ. Features missing in more than 50% of QC samples or in more than 80% of experimental samples were excluded. KNN algorithm (K-nearest Neighbor algorithm) was used to fill in the missing value. Data underwent probabilistic quotient normalization (PQN), QC-based LOESS signal correction (QC-RLSC) for batch effects, and removal of metabolites with >30% CV in QCs. A principal component analysis plot was provided to evaluate the sample quality (Fig. S14). To enhance the confidence of potential metabolites, we used the Metabolomics Standards Initiative level of confidence for grading metabolite annotation (levels 1-4). For metabolomic annotation, we further used metaboliteIDmapping AnnotationHub package in R language for unique and unified ID (https://github.com/yigbt/metaboliteIDmapping). After aligning metabolites and lipids across three timepoints, 355 metabolites and 177 lipids were selected for downstream analyses or modeling. The abundance of lipidomic and metabolomics profiles was normalized using the median before the batch effect correction.

Filtering of RNAseq data. For the RNAseq gene expression counts in peripheral blood, we filtered out the low-count to increase the power of statistical tests and reduce false positives. The criteria for exclusion were set at a count per million (CPM) greater than 0.1 and a prevalence higher than 30%, as low-count genes are more susceptible to high variability and spurious differential expression. Finally, 20186 genes were left for subsequent differential analyses. Trimmed Mean of M-values (TMM) was applied for normalization.

Batch correction. To mitigate potential inter-batch variations, we employed the ComBat method was used for batch-effect removal in both metabolomics and lipidomics data (*48, 49*). To avoid overadjustment for biological variation in transcriptomic data, the sequencing batch was included as a covariate in the linear mixed model, where applicable.

Longitudinally differential analysis

Cytokines: To evaluate the differences in cytokine levels between groups, we employed the Wilcoxon rank-sum test. The significance of cytokine variations was determined using a Benjamini-Hochberg false discovery rate (FDR) adjusted p-value threshold of 0.05.

Metabolites and lipids: For longitudinal comparisons, to identify the significance of the influence of sex, age, and durable time post-vaccination on features, we utilized a linear mixed effect model. This model was defined as log2 (Abundance) ~ Time + Sex + Age + 1|ID. The significance of changes in metabolites and lipids affected by the vaccines were determined using a FDR adjusted p-value of 0.1 (*50-52*).

RNAseq: Peripheral blood RNAseq gene-expression counts were analyzed using the limma/voom pipeline (*53, 54*). This approach facilitated the linear mixed effect which models the counts of each gene. Genei counts ~ Time + Age + Sex + Batch + 1|ID. This model statement enables quantifying effects that are due to vaccination (pre-vaccination, 1-month post vaccination, 6-month post vaccination) within each cohort. A notable advantage of employing limma in this context was the ability to incorporate 1|subject as a random effect, which is critical for controlling baseline individual differences. For data visualization of heat maps, transcripts per million (TPM) was used and either TPM values or row Z-scores were shown.

Sensitivity analysis: To directly compare differential features (M0-M1) between the two vaccine groups, interaction terms between vaccine groups and timepoints were included in the limma models. log2 (Abundance)/ Genei counts ~ Time * Vaccines + Sex + Age + 1|ID (*53*).

Differential analysis between immune outcomes. To identify differentially expressed genes (DEGs) at baseline between various levels of immune outcomes (immunogenicity & durable immunogenicity), we applied limma/voom analysis, accounting for age, sex, and sequencing batch (specifically for RNAseq). For metabolites and lipids, the significance was determined using a generalized linear model (GLM) adjusted for age and gender.

Within sample GSEA analysis. We employed the single sample gene set enrichment analysis (ssGSEA) method (*53, 54*), using the GSVA package in R (*55*), to profile within-sample differences in pathways from the GO list. The GO pathway list is a well-validated set of general curated biological pathways that offer insights into specific biological, cellular, and molecular processes. Variance stabilized transformed (vst) counts derived from DESeq2 were used as input for the GSVA function, employing default parameters (kcdf = “Gaussian”). The scaled Normalized Enrichment Scores (NES) were visualized as heatmaps. Notably, unlike classical GSEA, this analysis is agnostic to sample phenotype. To identify differential pathways between dichotomous immune outcomes, a linear mixed- effect model was used where we modeled the normalized enrichment score (NES) of each pathway as NES ~ sex + batch + group + 1|ID. The significance of pathways influenced by the treatment was determined using a Benjamini–Hochberg FDR adjusted p-value of <0.05.

Machine learning model.

Random forest modeling was conducted using Python version 3.11.5, scikit-learn 1.2.2, and boruta_py 0.3. Six independent prediction tasks were defined based on vaccine types and immune outcomes: higher immunogenicity prediction (top quantile of antibody level at M1 vs non-top quantile), lower immunogenicity prediction (lowest quantile of antibody level at M1 vs non-lowest quantile) and longer durability prediction (top quantile of antibody maintenance rate at M6 vs non-top quantile). For each task, the dataset was stratified and split into training (80%) and hold-out test (20%) sets. Within the training set, 10-fold cross-validation was performed, and feature selection was conducted by the Boruta algorithm on the combined multi-omics data. The random forest model was trained on the selected features with default hyperparameters (n_estimators = 100, random_state = 42) and evaluated on the hold-out test set using AUC as the performance metric.

Correlation analysis of differential features. We conducted Spearman correlation analysis to determine the associations between differential metabolites, lipids, genes, and cytokines at baseline and at 1 month. For this analysis, we established criteria wherein correlations yielding an FDR < 0.05 and a correlation coefficient > 0.3 were deemed significant. These significant correlations were then visualized using Gephi, a network visualization software, allowing for a comprehensive representation of the intricate relationships among these biomolecules (https://gephi.org/)(*56*). Furthermore, we stratified our study cohort into two subgroups: higher immunogenicity and non-higher immunogenicity, based on their immunogenicity and durability. Spearman correlation analysis was again performed within each subgroup to elucidate the unique interaction networks prevalent in these distinct groups. The same criteria of FDR < 0.05 and correlation coefficient >0.3 were applied for this subgroup analysis. In our network visualization using Gephi, we utilized a specific layout algorithm, Fruchterman Reingold, to spatially arrange the nodes (representing metabolites, lipids, genes, and cytokines) and edges (representing their correlations) in a manner that reflects the strength and complexity of their associations. The nodes were color-coded and sized according to the number of connections, respectively. The network was compared between higher immunogenicity and non-higher immunogenicity using network.compare function of the EGAnet Package.

Pathway Enrichment

Gene set enrichment analysis (GSEA). Enriched gene sets were identified using the preranked gene-set enrichment analysis (GSEA) algorithm implemented in the fgsea R package. The GSEA was employed to assess pathway enrichment between pre- and post- vaccination. Pre-ranked analysis, based on Limma/voom output comparing pre-vaccination and 1-month post vaccination, was conducted. Gene sets for GSEA were derived from the Molecular Signatures Database (MSigDB, version 7.0) using the GO list and blood transcription modules (BTMs) which is a curated database comprising of an integrated large-scale network of publicly available human blood transcriptomes (https://github.com/shuzhao-li/BTM) (*14, 57*). Only gene sets with an FDR of 0.05 were considered significant. A significant positive NES value indicates that members of the gene set tend to appear at the top of the ranked transcriptome data, whereas a negative NES indicates the opposite.

Differentially expressed genes were also mapped to the Search Tool for Retrieval of Interacting Proteins/Genes (STRING) database (which integrates known and predicted protein/gene interactions) to predict functional gene interaction networks (https://string- db.org/). Then, functional local clusters in the interaction network were determined using a Markov Cluster algorithm (MCL) with an inflation parameter = 3.

KEGG pathway analysis for differential metabolites. KEGG metabolic pathways and involved metabolites were downloaded from KEGG API (https://www.kegg.jp/kegg/rest/keggapi.html). Significant enriched KEGG pathways were determined by R clusterProfiler (v 4.8.3) package with BH multiple comparison test as FDR threshold of < 0.1 (*58*). Enrichment criteria included at least two metabolites for significantly altered metabolites, with an FDR < 0.1.

Declarations

Code availability. All software used in this study are from publicly available sources. Codes used for the omics analyses or figures are available via GitHub at https://github.com/Zhaoshilin007/Multi-omics-Covid-19-vaccines.

Data availability. Raw metabolomics and lipidomics data are included in Supplementary data. Transcriptomic sequencing data in this study were deposited into the NCBI Sequence Read Archive database under access number BioProject no. PRJNA1359148 and are available at the following URL: https://dataview.ncbi.nlm.nih.gov/object/PRJNA1359148.

Acknowledgements. This study was funded by the Health and Medical Research Fund-Commissioned Research on the Novel Coronavirus Disease (COVID-19) [COVID1903003] from the Health Bureau, the Hong Kong Special Administrative Region, Emergency Key Program of Guangzhou Laboratory, China [Grant No. EKPG22-30-6], RGC theme-based research schemes (T11-712/19-N and T11-705/21-N). Hein M Tun is a visiting associate professor at Yong Loo Lin School of Medicine, National University of Singapore. We also thank Dr. Subhankar Mukhopadhyay for proofreading.

Author contributions. C.K.P.M, H.M.T, and D.S.H conceived the research idea and designed the study. K.Y. and K.K.P.C coordinated and carried out cohort recruitment. H.M.T., S.Z., Y.C., J.Z., Y.P., H.L., X.L., B.Y. I, K.C.C. and C.K.P.M. analyzed the data. Y.S.T., C.W.T., C.C. and Y.S. performed the experiments. H.M.T., S.Z., Y.P., Y.C., C.K.P.M. and D.S.H. wrote the manuscript.

Declaration of interests. The authors declare no competing interests.

References

1. Krammer F (2020) SARS-CoV-2 vaccines in development. Nature 586:516–527
2. (2024) (World Health Organization, https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports)
3. Jara A et al (2021) Effectiveness of an Inactivated SARS-CoV-2 Vaccine in Chile. N Engl J Med 385:875–884
4. Haas EJ et al (2021) Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. Lancet 397:1819–1829
5. Tanriover MD et al (2021) Efficacy and safety of an inactivated whole-virion SARS-CoV-2 vaccine (CoronaVac): interim results of a double-blind, randomised, placebo-controlled, phase 3 trial in Turkey. Lancet 398:213–222
6. Plotkin SA, Plotkin SL (2011) The development of vaccines: how the past led to the future. Nat Rev Microbiol 9:889–893
7. Grohskopf LA et al (2019) Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices - United States, 2019-20 Influenza Season. MMWR Recomm Rep 68:1–21
8. Bandyopadhyay AS, Garon J, Seib K, Orenstein WA (2015) Polio vaccination: past, present and future. Future Microbiol 10:791–808
9. Sahin U, Karikó K, Türeci Ö (2014) mRNA-based therapeutics–developing a new class of drugs. Nat Rev Drug Discov 13:759–780
10. Jackson NAC, Kester KE, Casimiro D, Gurunathan S, DeRosa F (2020) The promise of mRNA vaccines: a biotech and industrial perspective. NPJ Vaccines 5:11
11. Mok CKP et al (2023) Omicron BA.1-specific T-cell responses in adults vaccinated with CoronaVac or BNT162b2 in Hong Kong: an observational cohort study. Lancet Microbe 4:e418–e430
12. Mok CKP et al (2022) Comparison of the immunogenicity of BNT162b2 and CoronaVac COVID-19 vaccines in Hong Kong. Respirology 27:301–310
13. Pulendran B, Davis MM (2020) The science and medicine of human immunology. Science 369
14. Li S et al (2014) Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat Immunol 15:195–204
15. Franco LM et al (2013) Integrative genomic analysis of the human immune response to influenza vaccination. Elife 2:e00299
16. Lau JJ et al (2023) Real-world COVID-19 vaccine effectiveness against the Omicron BA.2 variant in a SARS-CoV-2 infection-naive population. Nat Med 29:348–357
17. Ng SC et al (2022) Gut microbiota composition is associated with SARS-CoV-2 vaccine immunogenicity and adverse events. Gut 71:1106–1116
18. Peng Q et al (2022) Waning immune responses against SARS-CoV-2 variants of concern among vaccinees in Hong Kong. EBioMedicine 77:103904
19. Pralhada Rao R et al (2013) Sphingolipid metabolic pathway: an overview of major roles played in human diseases. *J Lipids* 178910 (2013)

20. Yaribeygi H et al (2020) Neuromodulatory effects of anti-diabetes medications: A mechanistic review. Pharmacol Res 152:104611
21. He M et al (2022) Metabolomics-based investigation of SARS-CoV-2 vaccination (Sinovac) reveals an immune-dependent metabolite biomarker. Front Immunol 13:954801
22. Ryan FJ et al (2023) A systems immunology study comparing innate and adaptive immune responses in adults to COVID-19 mRNA and adenovirus vectored vaccines. Cell Rep Med 4:100971
23. Wang Y et al (2022) Single-cell transcriptomic atlas reveals distinct immunological responses between COVID-19 vaccine and natural SARS-CoV-2 infection. J Med Virol 94:5304–5324
24. Corbett KS et al (2020) SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. Nature 586:567–571
25. Hsieh CL et al (2020) Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. Science 369:1501–1505
26. Greaney AJ et al (2021) Antibodies elicited by mRNA-1273 vaccination bind more broadly to the receptor binding domain than do those from SARS-CoV-2 infection. Sci Transl Med 13
27. Doria-Rose NA, Joyce MG (2015) Strategies to guide the antibody affinity maturation process. Curr Opin Virol 11:137–147
28. Chan TD, Brink R (2012) Affinity-based selection and the germinal center response. Immunol Rev 247:11–23
29. Hwang JK et al (2017) Sequence intrinsic somatic mutation mechanisms contribute to affinity maturation of VRC01-class HIV-1 broadly neutralizing antibodies. Proc Natl Acad Sci U S A 114:8614–8619
30. Drury RE et al (2024) Multi-omics analysis reveals COVID-19 vaccine induced attenuation of inflammatory responses during breakthrough disease. Nat Commun 15:3402
31. Zhou C et al (2025) Systemic profiling of immune responses in healthy adults vaccinated with an RBD-targeting COVID-19 mRNA vaccine. *hLife* 3, 415–432
32. Watanabe Y et al (2023) Time-series transcriptome analysis of peripheral blood mononuclear cells obtained from individuals who received the SARS-CoV-2 mRNA vaccine. J Med Virol 95:e28884
33. Chatterjee A, Chakravarty A (2023) Neurological Complications Following COVID-19 Vaccination. Curr Neurol Neurosci Rep 23:1–14
34. Crook H, Raza S, Nowell J, Young M, Edison P (2021) Long covid-mechanisms, risk factors, and management. BMJ 374:n1648
35. Xu E, Xie Y, Al-Aly Z (2022) Long-term neurologic outcomes of COVID-19. Nat Med 28:2406–2415
36. Duff EP et al (2025) Plasma proteomic evidence for increased β-amyloid pathology after SARS-CoV-2 infection. Nat Med
37. Lee M, Lee SY, Bae YS (2023) Functional roles of sphingolipids in immunity and their implication in disease. Exp Mol Med 55:1110–1130
38. Zhang SS et al (2024) Lipidome is a valuable tool for the severity prediction of coronavirus disease 2019. Front Immunol 15:1337208
39. Hannun YA, Obeid LM (2018) Sphingolipids and their metabolism in physiology and disease. Nat Rev Mol Cell Biol 19:175–191
40. Koeken V et al (2022) Plasma metabolome predicts trained immunity responses after antituberculosis BCG vaccination. PLoS Biol 20:e3001765
41. Tsang JS et al (2020) Improving Vaccine-Induced Immunity: Can Baseline Predict Outcome? Trends Immunol 41:457–465
42. Peng Y et al (2023) Baseline gut microbiota and metabolome predict durable immunogenicity to SARS-CoV-2 vaccines. Signal Transduct Target Ther 8:373
43. Moorlag S et al (2024) Multi-omics analysis of innate and adaptive responses to BCG vaccination reveals epigenetic cell states that predict trained immunity. Immunity 57:171–187e114

44. Bakker OB et al (2018) Integration of multi-omics data and deep phenotyping enables prediction of cytokine responses. Nat Immunol 19:776–786
45. Chang YS et al (2024) Altered transcriptomic immune responses of maintenance hemodialysis patients to the COVID-19 mRNA vaccine. Elife 13
46. Hu Z et al (2022) Early immune markers of clinical, virological, and immunological outcomes in patients with COVID-19: a multi-omics study. Elife 11
47. Hosmer DW Jr, Lemeshow S, Sturdivant RX (2013) Applied logistic regression. Wiley
48. Zhang Y, Parmigiani G, Johnson WE (2020) ComBat-seq: batch effect adjustment for RNA-seq count data. NAR Genom Bioinform 2:lqaa078
49. Yu Y et al (2023) Correcting batch effects in large-scale multiomics studies using a reference-material-based ratio method. Genome Biol 24:201
50. Xiao N et al (2021) Integrated cytokine and metabolite analysis reveals immunometabolic reprogramming in COVID-19 patients with therapeutic implications. Nat Commun 12:1618
51. Cheng K et al (2021) Genome-scale metabolic modeling reveals SARS-CoV-2-induced metabolic changes and antiviral targets. Mol Syst Biol 17:e10260
52. Nie M et al (2021) Evolutionary metabolic landscape from preneoplasia to invasive lung adenocarcinoma. Nat Commun 12:6479
53. Wipperman MF et al (2021) Gastrointestinal microbiota composition predicts peripheral inflammatory state during treatment of human tuberculosis. Nat Commun 12:1141
54. Smyth GK (2005) In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S (eds) Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer New York, New York, NY, pp 397–420
55. Hänzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14:7
56. Bastian M, Heymann S, Jacomy M (2009) in *Proceedings of the international AAAI conference on web and social media*. vol. 3, pp. 361–362
57. Link VM et al (2024) Differential peripheral immune signatures elicited by vegan versus ketogenic diets in humans. Nat Med 30:560–572
58. Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. Omics 16:284–287

Tables

Table 1. Baseline characteristics of study population

	BNT162b2 (N=274)	CoronaVac (N=279)	P-value
Age (years)			<0.001
Mean (SD)	45.77 (13.45)	52.34 (9.68)	
Median (IQR)	46.00 (35.00, 56.00)	54.00 (46.50, 59.00)	
Gender			0.0771
Female	162 (59.6%)	187 (67.0%)	
Male	110 (40.4%)	92 (33.0%)	
Lifestyle habits			
Smoker	9 (3.3%)	17 (6.1%)	0.159
Alcohol consumption	121 (44.5%)	111 (39.8)	0.494
Regular exercise	144 (52.9%)	139 (49.8%)	0.496
Comorbidities			
Cardiovascular diseases	21 (7.7%)	16 (5.7%)	0.397
Diabetes	12 (4.4%)	19 (6.8%)	0.268
Chronic respiratory diseases	3 (1.1%)	5 (1.8%)	0.725
Vaccine history			
Influenza	170 (62.5%)	167 (59.9%)	0.541
Never received	102 (37.5%)	112 (40.1%)	0.254
Received once but not within previous 2 years	18 (6.6%)	15 (5.4%)	
Received once within these two years	58 (21.3%)	74 (26.5%)	
Received every year	94 (34.6%)	78 (28.0%)	
Hepatitis A/B	146 (53.7%)	82 (29.4%)	<0.001
Mumps	56 (20.6%)	17 (6.1%)	<0.001
Pneumococcal conjugate vaccine	21 (7.7%)	9 (3.2%)	0.0238
Rabies	15 (5.5%)	8 (2.9%)	0.139
Typhoid	14 (5.1%)	7 (2.5%)	0.122
Hemorrhagic fever	5 (1.8%)	0 (0%)	0.0288
COVID-19 history in household	2 (0.7%)	1 (0.4%)	0.62
Convalescent from SARS-CoV infection	2 (0.7%)	2 (0.7%)	1
SARS-CoV-2 antibody response[£]			
AUC of spike RBD IgG level (median (IQR))	2.46 (2.25, 2.68)	1.01 (0.72, 1.37)	<0.001
sVNT (inhibition %) at one-month post 2nd dose* (Median (IQR))	96.08 (93.63, 96.92)	55.02 (39.73, 69.34)	<0.001

sVNT (inhibition %) at six-month post 2nd dose (Median (IQR))	73.11 (55.00, 84.75)	16.18 (7.44, 25.87)	<0.001

Data are n (%), unless otherwise stated. [£]The antibody level and host responses between the two vaccine groups were compared based on multivariable regression analysis, adjusting by age, gender, and vaccine history (hepatitis A/B, mumps, pneumococcal conjugate, and hemorrhagic fever). *One participant excluded due to insufficient peripheral blood mononuclear cells for analysis.

Figures

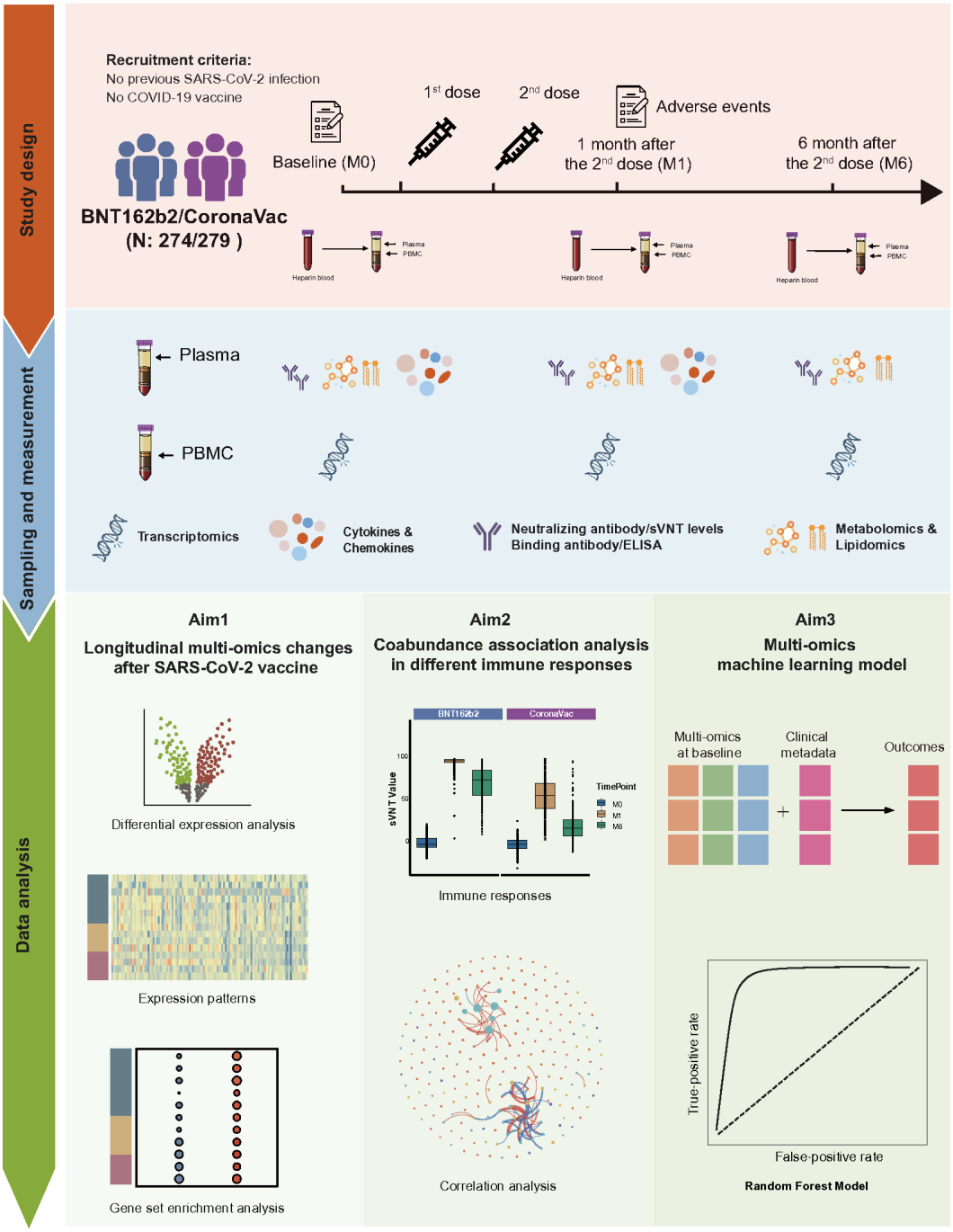


Figure 1

Study overview. Overview of the clinical protocol with three-timepoint analysis, that is, before, 1 month, and 6 months following BNT162b2 or CoronaVac vaccine. At each timepoint, heparin blood samples were obtained. Experimental procedures, analysis, and integration were carried out as described in Methods.

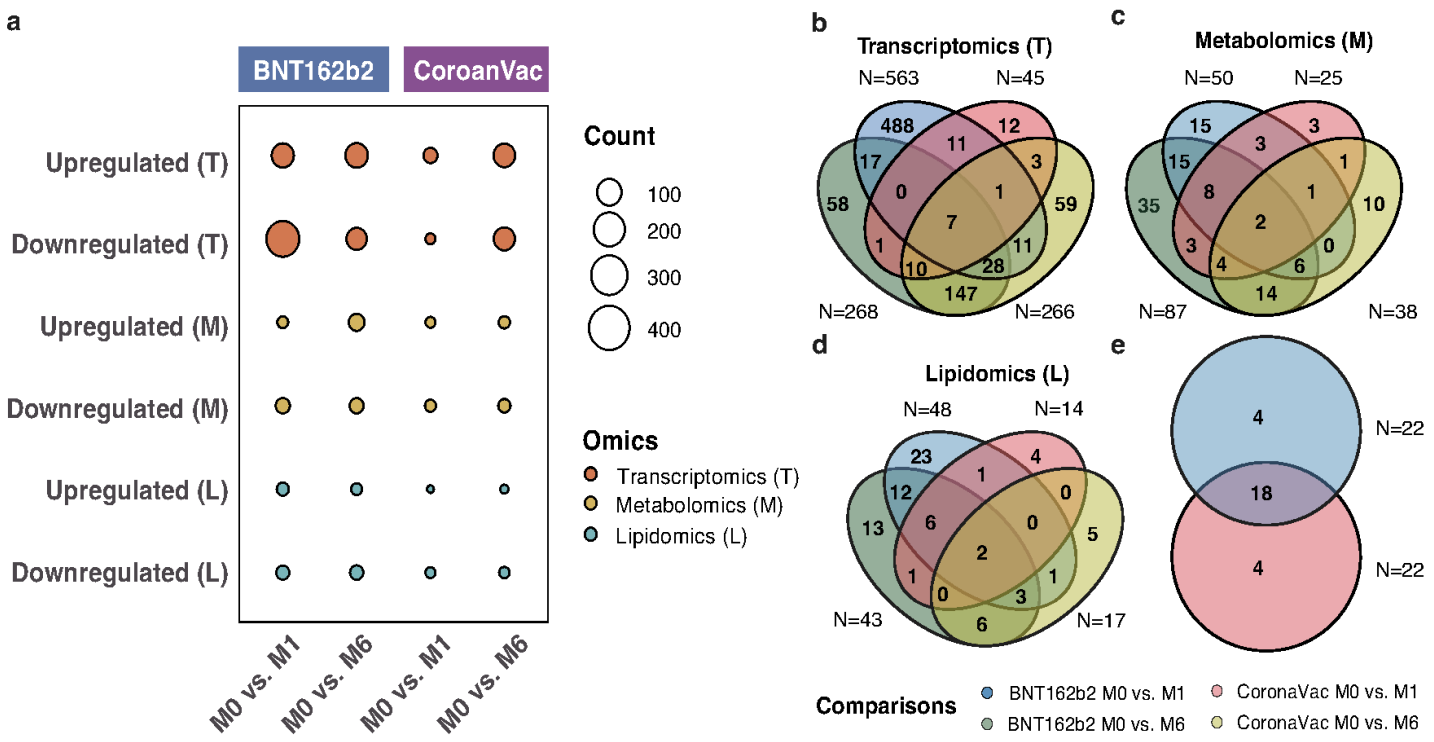


Figure 2

A wider range of features are regulated after BNT162b2 than CoronaVac. a, Scatter Plot Depicting Distribution of Differential Features. This plot illustrates the count of differential features within BNT162b2 or CoronaVac group, at baseline vs. 1-month post-vaccination (M1) and baseline vs. 6-month post-vaccination (M6). Each omics is positioned according to the direction of regulation, omics, and vaccine groups. The size of the circle represents the number of differential features in each respective comparison. The colors represent transcriptomics, metabolomics, and lipidomics, respectively. **b-e, Comparative Analysis of Vaccine-Induced Changes in genes, metabolites, lipids, and cytokines.** The Venn diagrams depict the number of shared and unique genes (b), metabolites (c), lipids (d), and cytokines (e) identified between pre-vaccination and post-vaccination timepoints for BNT162b2 and CoronaVac vaccines. The intersections indicate the count of features overlapping between different comparison sets. Metabolites, lipids, and cytokines were considered significant if the FDR was less than 0.1. For gene expression, significant differential features were defined by an FDR < 0.05 and an absolute log2 fold change (|log2(FC)|) > 1.5. A comprehensive dataset, including the complete number of features and the detailed statistical results, is available in the supplementary Source Data file (Supplementary Table 3-4).



Figure 3

BNT162b2 is associated with heightened adaptive immunity and decreased neural functions and CoronaVac with heightened innate immunity. a. Bar plot for different Gene Ontology (GO) Biological Process. Each row in the chart corresponds to a different Gene Ontology (GO) Biological Process. The x-axis represents the count of occurrences or measurements associated with each GO term. For each GO term, there are two bars extending in opposite directions from a central axis at 0. The left panel represents a negative normalization enrichment score (suppression), and the right panel indicates a positive normalization enrichment score (activation). Normalized enrichment scores calculated for the GO

Pathway list for one-month post-vaccination vs. Baseline. The colors of the bars distinguish between two different vaccine groups: BNT162b2 (blue) and CoronaVac (red). **b-c. Over-representation Analysis by GO list and Blood transcription module (BTM) analysis.** These panels illustrate the comparative normalized enrichment scores (NES) derived from genes associated with Neural Function (b), and Immunological Functions (c) for pre-vaccination versus 1-month post-vaccination samples from BNT162b2 and CoronaVac recipients. For Cellular Processes, shared pathways between the two vaccines and the top 25 distinct pathways, as ranked by NES, are presented. The input data comprised variance-stabilized counts from DESeq2, which were processed using the GSVA function to calculate scaled NES. Dot size is proportional to the significance level ($-\log_{10}$ p-value), with the color indicating the NES. The significance of enrichment was determined using the fgsea package and adjusted for multiple testing. Complete source data are available as supplementary material. **d. Distribution of Pathways with Positive Enrichment Post-Vaccination.** The cumulative distribution function (CDF) plots illustrate the frequency of pathways exhibiting positive normalized enrichment scores (NES) when comparing pre-vaccination (M0) to 1-month post-vaccination (M1) samples for the BNT162b2 vaccine (left) and the CoronaVac vaccine (right). Pathways are categorized by their associated functions: Cellular Process (blue), Immunological Function (yellow), and Neural Function (red). The x-axis indicates the number of pathways showing positive effect sizes, while the y-axis displays the cumulative frequency of these pathways within the dataset. The dashed horizontal line at 0.5 marks the median of the distribution. **e. Over-representation Analysis of Significantly Altered Metabolites by Vaccination (FDR <0.1).** The bar plot displays enriched longitudinal differences in metabolite levels between baseline and 1-month post-vaccination. Bar plots are categorized by vaccine type, with BNT162b2 (left) and CoronaVac (right). The color intensity is proportional to the significance level.

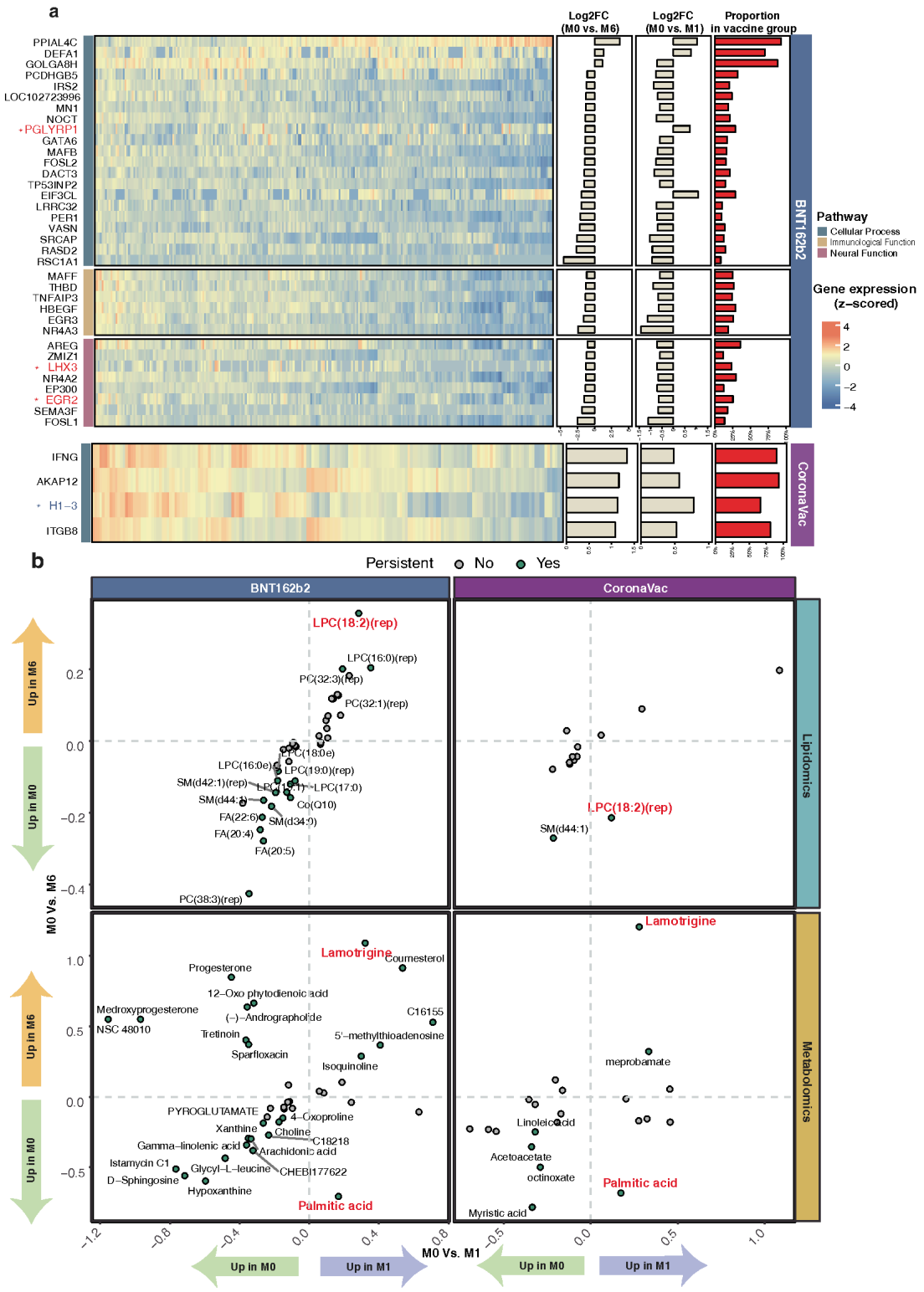


Figure 4

Longitudinal Profiling of Vaccine-induced Omics Features and Pathways Alterations. a. Heat map showing significance of differences in transcripts vs. fold change from baseline at 6m post-BNT162b2 and 6m post-CoronaVac. Adjacent to the heatmaps, the log2 fold changes (6-month post-vaccination/pre-vaccination and 1-month post-vaccination/pre-vaccination) are depicted in yellow, alongside the proportion of the cohort exhibiting an increase in gene expression, indicated in red. Only genes that were significantly differentially expressed and overlapped DEGs in the pre-vaccination and 1-month post-vaccination comparison after adjusting for a 5% FDR correction are included. The genes are categorized by their functions:

Cellular Process, Immunological Function, and Neural Function. The color code reflects the magnitude of expression, with the intensity of each color correlating to the expression level in each subject. The y-axis in the heatmap is highlighted when genes whose expression levels at M6 are significantly correlated with the relative decrease in sVNTs between 1 and 6 months post-vaccination. Red indicates a positive correlation, while blue indicates a negative correlation. *p value < 0.05. **b. Scatter plots showing significantly differential metabolites/ lipids (FDR<0.05) between baseline and M1 and M6.** Only the significantly differential features between pre-vaccination and 1-month post vaccination were depicted to present the expression levels at 6-month post-vaccination. The overlapping markers between the two vaccine groups were highlighted in red. P value was given by $\log_2$(Abundance)= Time + Sex + Age + 1|ID. C16155: UDP−beta−L−threo−pentopyranos−4−ulose, C18218: 16−Hydroxyhexadecanoic acid, CHEBI177622: Oleoyl-L-α-lysophosphatidic acid.

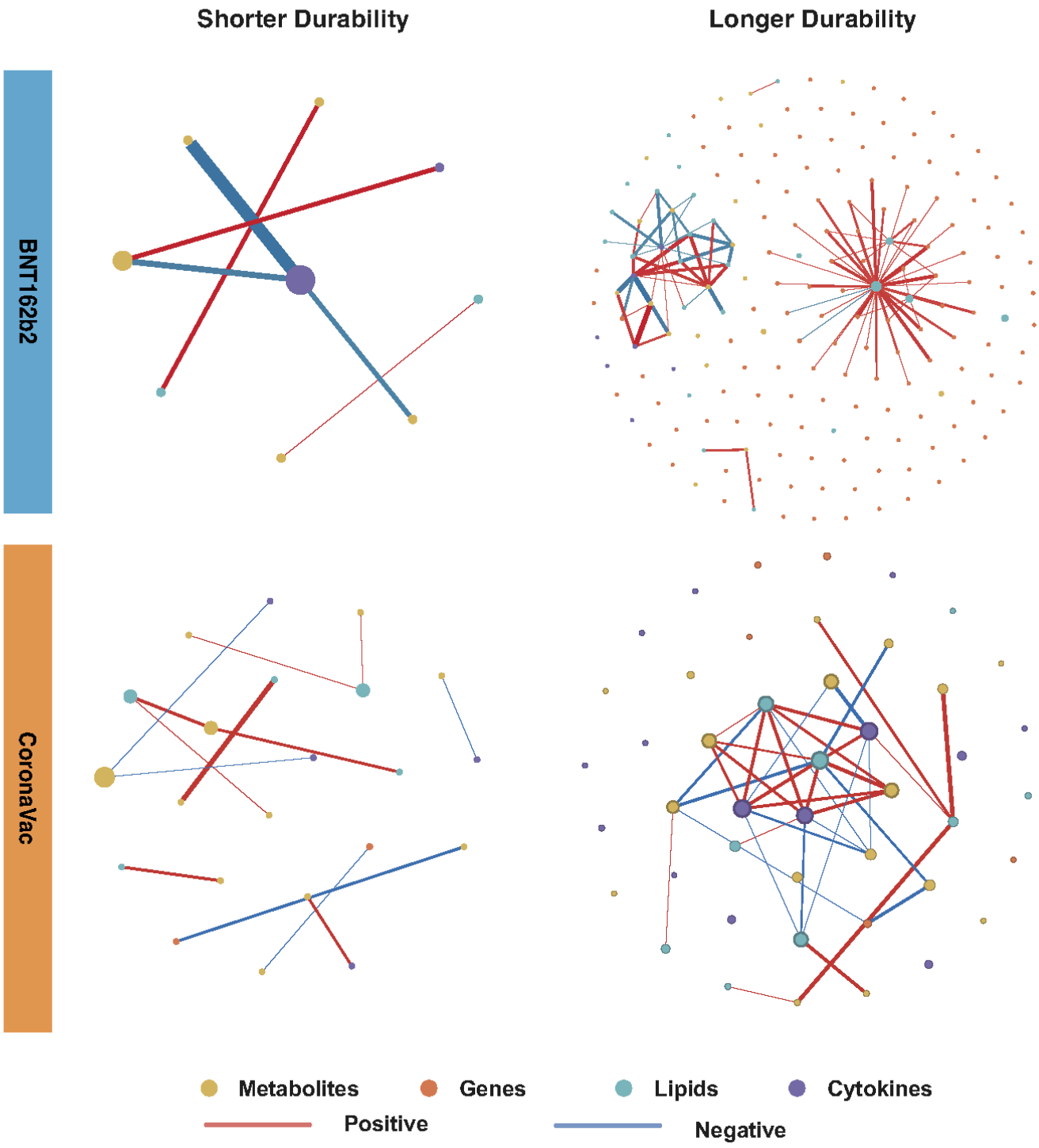


Figure 5

Co-Abundance Networks of Omics Features between Longer and Shorter Durability. Network plots illustrate the co-abundance networks that emerged from integrating all differentially expressed omics markers between baseline and 1-month post-vaccination. The features abundance at M1 were used for the correlation calculation. Nodes in these networks are color-coded by omics categories: lipidomics (green), metabolomics (yellow), transcriptomics (orange), and cytokines (purple). Significance is assigned to fair absolute correlations greater than 0.3 (FDR < 0.05).

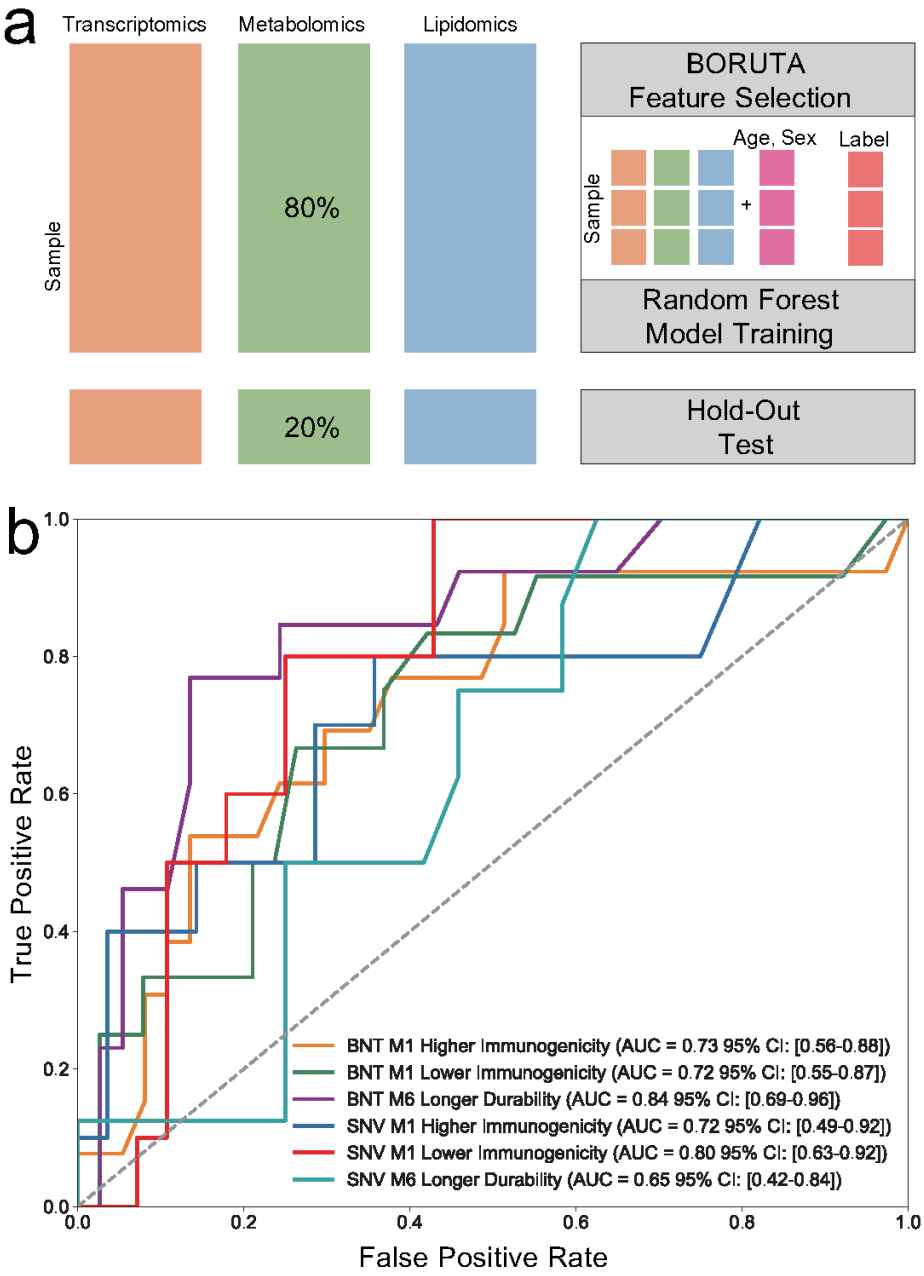


Figure 6

Multi-omics-constrained mathematical modeling for immune responses. The prediction model was carried out as described in Methods. a. Workflow of the prediction model. b. Model performance in classifying subjects with higher immunogenicity or non-higher, lower immunogenicity or non-lower and differentiating between those with longer or non-longer durability across the two vaccine groups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S11.jpg
- Fig.S12.jpg
- Fig.S13.jpg
- Fig.S1.jpg
- Fig.S10.jpg
- Fig.S2.jpg
- Fig.S3.jpg
- Fig.S4.jpg
- Fig.S5a.jpg
- Fig.S5b.jpg
- Fig.S6.jpg
- Fig.S7.jpg
- Fig.S8.jpg
- Fig.S9.jpg
- SupplementaryTable1.docx
- SupplementaryTable11.xlsx
- SupplementaryTable12.xlsx
- SupplementaryTable13.xlsx
- SupplementaryTable14.docx
- SupplementaryTable2.docx
- SupplementaryTable3.xlsx
- SupplementaryTable4.xlsx
- SupplementaryTable5.xlsx
- SupplementaryTable6.xlsx
- SupplementaryTable7.xlsx
- SupplementaryTable8.xls
- SupplementaryTable10.xlsx
- SupplmentaryTable9.xls