Case report

Utility of metagenomic Next-Generation Sequencing for simultaneously detecting pathogens and neoplasms

Feng Qin [a,1], Xuejiao Hu [b,1], Xiaojia Wang [c,1], Weijiang Liu [b], Qianyun Deng [b], Yunhu Zhao [b], Caiyun Li [c], Chao Liu [c,*], Zhenchao Huang [a,**], Bing Gu [b,***]

[a] Department of Neurosurgery, Lingnan Hospital, Branch of the Third Affiliated Hospital of Sun Yat-sen University, 2693 Kaichuang Avenue, Guangzhou, Guangdong Province, 510530, China
[b] Department of Laboratory Medicine, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, 106 Zhongshan 2nd Rd, Guangzhou, Guangdong Province, 510000, China
[c] Medical Department, Hangzhou Matridx Biotechnology Co., Ltd., Bd 2-4, 2073 Jinchang Rd, Hangzhou, Zhejiang Province, 310030, China

A B S T R A C T

Objectives: Clinicians often face the challenge of differentially diagnosing febrile patients who are suspected of infectious diseases, since the clinical manifestations of infection and cancer may overlap. A single test that can detect both pathogens and tumor could provide timely and accurate diagnostic clues to aid the treatment and management of these patients.
Methods: We enrolled eight patients to evaluate the utility of metagenomic Next-Generation Sequencing for simultaneously detecting pathogens and neoplasms using body fluids and tissue samples. Patients were selected by the following criteria: 1) Tumor was not considered upon hospitalization, but mNGS testing indicated neoplasm; 2) Tumor was not excluded, but microbial infection was primarily suspected according to initial clinical assessment.
Results: We detected potential pathogens in five patients, three of whom had progressed into critical infections. Moreover, abnormal chromosomal copy numbers were identified in all patients that indicated presence of neoplasms, which were pathologically confirmed.
Conclusions: Although copy number variations do not render a definitive cancer diagnosis, it can prompt clinicians to conduct more focused diagnostic testing for cancer, potentially saving time and cost. As a result, integrating copy number analysis with pathogen detection in mNGS may help establish rapid and accurate diagnosis for febrile patients.

1. Introduction

Metagenomic Next-Generation Sequencing (mNGS) represents a revolutionary molecular technique for detecting pathogens. This method has gained significant traction for its utility in diagnosing infectious diseases, particularly when conventional diagnostic

List of abbreviations

mNGS	metagenomic Next-Generation Sequencing
FUO	fever of unknown origin
CNV	copy number variation
PET	positron emission tomography
CT	computed tomography
DLBCL	diffuse large B-Cell lymphoma
MRI	cranial magnetic resonance imaging
SWI	susceptibility weighted imaging
PCR	polymerase chain reaction

methods, such as microbiological culture, fall short [1]. Uniquely, mNGS operates without any preconceived hypotheses and does not rely on culture methods, making it versatile across a range of sample types including blood, body fluids, and biopsy tissues [2–5]. Furthermore, the scope of mNGS extends beyond mere pathogen detection, offering substantial clinical applications [6]. For instance, two recent studies have used mNGS to identify neoplasms from body fluids and cerebrospinal fluid with an overall sensitivity of 87 % and specificity of 100 % [7,8]. Indeed, many patients present with symptoms such as fever, pain, and fatigue are empirically treated for infectious diseases, but eventually diagnosed with cancer. Notably, the most common causes for fever of unknown origin (FUO) are infection, malignancy, and collagen vascular diseases, while malignancy alone accounts for 14 % of adult FUO cases [9].

Biopsy and histological staining are the gold standards for cancer diagnosis. In contrast, physical exam and serology tests usually lack specificity in differentiating cancer from other diseases [10]. Positron emission tomography/computed tomography (PET/CT) scan is routinely performed for diagnosis and staging of cancer. It has an overall sensitivity of 64 % in identifying malignant lung nodules and 88.7 % in diffuse large B-Cell lymphoma (DLBCL) [11,12]. However, when infectious diseases are mainly considered, oncology tests may be overlooked and thus cause a delay of proper diagnosis. Provided that pathogens and neoplasms can be detected simultaneously in a single test, it would be useful for diagnosing febrile patients when neither infection nor malignancy can be confidently ruled out.

Genomic instability is a recognized hallmark of tumor cells, as evidenced in multiple studies [13]. These studies highlight the significance of copy number variation (CNV) as a crucial aspect of genetic variation that contributes to the development and progression of tumors. Recently, chromosomal CNVs have gained attention for their role in the early diagnosis of various cancers, including colorectal, breast, and hematologic malignancies [14–16]. In light of this, mNGS presents a dual diagnostic potential for detecting both infections and cancers. It achieves this by unbiasedly sequencing DNA from human cells and microorganisms alike. In our study, mNGS was employed on eight patients with febrile illnesses, leading to the detection of CNVs indicative of neoplasms. Subsequent pathological staining validated these findings, paving the way for the timely administration of appropriate anticancer treatments.

2. Methods

2.1. Patient selection

We enrolled eight patients from Nov. 2020 to May 2022, who were hospitalized at Guangdong Provincial People's Hospital or Lingnan Hospital, Guangzhou, China. The patients were selected if they met either of the following criteria: 1. Tumor was not considered but CNV analysis by mNGS indicated tumor (patient 1/2/3/4/5/6); 2. Tumor was not ruled out while infection was mainly considered without etiological evidence (patient 7/8).

2.2. Nucleic acid extraction and library preparation

For DNA extraction, the Matridx kit (Cat# MAR002) was used, adhering to the manufacturer's Standard Operating Procedures (SOPs). Peripheral blood underwent centrifugation at 16,000g for 10 minutes to extract cell-free DNA (cfDNA) from plasma. In the case of other sample types, genomic DNA was extracted from 1 mL of specimens. The extracted DNA was then subjected to library preparation, which included enzymatic fragmentation of genomic DNA, end repairing, terminal adenylation and adaptor ligation, followed by purification. These procedures and methodologies were outlined in a previous study [17]. For plasma cfDNA, no enzymatic fragmentation was included since cell-free DNA is intrinsically fragmented. Library concentrations were quantified by real-time PCR (KAPA) and subsequently sequenced on an Illumina Nextseq platform.

2.3. Metagenomic sequencing and bioinformatics

Shotgun sequencing was performed on the Illumina Nextseq, generating approximately 20 million 50bp single-end reads per library, utilizing dual barcode sequencing to minimize index hopping. Bioinformatic analysis was conducted in accordance with methodologies described in a previous report [18]. This involved filtering out human-origin sequences (GRCh38. p13) and aligning the

remaining reads against a comprehensive reference database (NCBI nt, GenBank, and an in-house curated genomic database) to ascertain microbial species and read count. Each sequencing run included a negative control comprising a culture medium with 10^4 Jurkat cells/mL.

2.4. Pathogen reporting criteria

Microbial reads from a library were reported if: 1) the sequencing data met specific quality control thresholds (library concentration >50 pM, Q20 > 85 %, Q30 > 80 %); 2) the negative control (NC) in the same sequencing run either did not contain the species or the RPM ratio (sample to NC) was ≥5, a cutoff established from previous studies for differentiating true positives from background contaminations [2,17,19].

2.5. CNV analysis

DNA sequences were aligned to the human hg19 (GRCh37) reference genome. Guanine-cytosine (GC) content bias was corrected using LOESS regression, with the criteria that the GC ratio should be lower than 0.44, and the count of clean reads should surpass 3 million for quality assurance. The standard deviation of the read fold change for each bin of data (bin size 100k) and normalized read counts were determined. Copy number variations were identified using XHMM and Canoes, referencing normal chromosomal copy numbers [20,21]. Due to the mixture of tumor and normal DNA in the specimens, detecting overall copy number changes against the background signal could be challenging. For this, a computational algorithm was utilized to predict neoplasm likelihood, as mentioned in our previous study [22]. Larger CNVs (>10 M), covering multiple chromosomes, were indicative of tumor cells rather than inherited disorders [7].

3. Results

3.1. Clinical characteristics and mNGS findings

Diagnosis, results of routine clinical tests and patient age and gender were listed in Table 1. The results of microbiological culture、 mNGS findings and follow-up/outcomes were shown in Table 2. Patients 1,4, 7 and 8 were discharged, and patients 2, 3, 5 and 6 were transferred to other hospitals or departments. The normalized copy number results for chromosomes in each patient were presented in Fig. 1. The Hematoxylin and Eosin (H&E) staining and the immunohistochemical staining were shown in Fig. 2 A-I.

3.2. Diagnosis of lymphoma in patients suspected of pneumonia

Patients 1, 2 and 3 were initially diagnosed with pneumonia and later confirmed to have lymphoma **(Figure A-D)**, which can be

Table 1
Patient information, clinical test results and diagnosis.

Patient#/ age/gender	Serum CRP	Serum PCT	Blood WBC	Blood NEU	Symptoms	Initial diagnosis	Final diagnosis
	mg/L	ng/mL	x109/L	×109			
1/77/male	91.12	0.98	4.6	2.6	Fever, cough	CAP	lymphoma
2/57/female	202.49	1.56	11.45	10.55	Weakness, fever	Pneumonia	B-cell lymphoma
3/67/male	171.49	0.43	6.09	4.63	Difficulty breathing, cough, expectoration	Pneumonia	Diffused B-cell lymphoma
4/53/male	29.26	<0.05	14.18	9.28	Cough, expectoration, back pain	Lung abscess	Lung cancer
5/71/male	156	0.7	7.7	6.34	Cough, expectoration, hemoptysis	Pneumonia	Lung cancer Pneumonia
Patient#/ age/ gender	CSF Protein **g/L**	CSF chloride **mmol/L**	CSF WBC **x10^6/L**	CSF RBC **x10^6/L**	Symptoms	Initial diagnosis	Final diagnosis
6/18/male	0.26	146.5	7	0	Minimally conscious state	Intracranial infection	Purulent meningitis and Plurihormonal PIT-1 positive pituitary adenoma
Patient#/ age/ gender	Blood Hb **g/L**	Blood PLT **x10^9/L**	Blood WBC **x10^9/L**	Blood NEU **x10^9**	Symptoms	Initial diagnosis	Final diagnosis
7/44/male	86	11	15.97	2.52	Fever, headache	Pancreatic lesion	Malignant tumor with secondary bone marrow neoplasm
8/72/male	51	24	0.42	0.16	Fever, cough	AML septicemia	AML septicemia

CRP, C-reactive protein; NEU, neutrophil; PCT, procalcitonin; WBC, white blood cell; CSF, cerebrospinal fluid; RBC, red blood cell; Hb, hemoglobin; PLT, platelet; MDS, myelodysplastic syndrome; AML, acute myeloid leukemia.

Table 2
Microbiological culture and mNGS results.

Patient#	Microbial Culture	Sample type	Detected Pathogens By mNGS	CNVs Detected	Follow-up/Outcome
1	N/A	Blood	*Human Herpesvirus 5*	Yes	Deceased
2	Negative	Blood	*Human Herpesvirus 5*	Yes	Improved and transferred to another
			Human Herpesvirus 7		hospital
		Pleural effusion	*Human Herpesvirus 5*	No	
		CSF	None	No	
3	Negative	Blood	None	Yes	Transferred to EICU
4	Negative	Lung biopsy	None	Yes	Discharged
5	Negative	BALF	*Candida tropicalis*	Yes	Improved and transferred for further
		Lung biopsy	*Human Herpesvirus 7*	No	treatment
6	*Acinetobacter*	Blood	None	No	Transferred to a specialized hospital
	baumannii	CSF	None	No	
		Brain biopsy	*Escherichia coli*	Yes	
			Stenotrophomonas maltophilia		
7	N/A	Blood	None	Yes	Discharged
8	N/A	Blood	*Legionella pneumophila*	Yes	Discharged
			Bacteroides fragilis		
			Enterococcus faecium Human polyoma		
			virus		

misdiagnosed since it can be present in multiple organs with a diverse range of symptoms. Patient 3 (67-year-old male) was hospitalized due to difficulties in breathing, cough, expectoration for more than 30 days and fever for more than 15 days. He had hyperuricemia for 5 years and was allergic to tetracycline and oxytetracycline. Chest CT showed cavities in the right upper anterior mediastinum, and nodular shadow in the lower part of both lungs. Serum tumor markers were normal (AFP 1.22ng/mL, CEA 3.13ng/mL) except for a slightly elevated level of CA125 (40.81 U/mL). PET/CT scan reported several findings: 1) multiple enlarged lymph nodes, right ethmoid sinus lesions, bilateral adrenal masses, multiple lesions in both lungs, bone destruction of right clavicle and thoracic 11 right pedicle. The glucose metabolism of the above lesions was elevated; 2) emphysema; 3) hepatic cyst; 4) multiple cysts on both kidneys. Ultrasound imaging showed liver and kidney cysts and lymphadenopathy in the neck and abdomen regions. Lesions in the right lobe of the thyroid gland were seen. The patient was empirically treated with moxifloxacin sodium chloride (250mg qd) and methylprednisolone (40 mg bid). The peripheral blood was sent for mNGS, which reported no pathogens but significant chromosomal CNVs (Fig. 1). Right supraclavicular lymph node biopsy was performed for pathological examinations. Proliferative lymphoid tissues were seen. The tissue was composed of singular and large heterotypic proliferative lymphoid cells (Fig. 2D). Cell nuclei were round or slightly irregular shape (Fig. 2D). Nuclear division and apoptotic cells were seen. Infiltration of blood vessels and adipose tissue was seen (Fig. 2D). These observations were consistent with diffuse large B-cell lymphoma and the patient was transferred for further treatment.

3.3. Diagnosis of lung cancer in a patient suspected of lung abscess

The CT findings of lung cancer can be similar to those of lung abscesses that may lead to misdiagnosis. Patient 4 (53-year-old male) was hospitalized due to cough, expectoration, and back pain. Chest CT showed emphysema, and right lower lobe abscess, but no cancerous lesions were found. The patient was empirically treated with moxifloxacin and cefazoxime sodium. The symptoms were relieved after treatment. However, the level of white blood cell (14.68×10^9/L) and Eosinophil ratio (0.103) remained elevated. The possibility of parasitic infection and tumor was considered. Bronchoalveolar lavage fluid was sent for mNGS with no pathogen detected and CNV analysis revealed abnormal chromosomal copy number (Fig. 1). Ciliated columnar epithelial cells and scattered neutrophils and lymphocytes were found in bronchoscope brush samples with no tumor cells. Further biopsy had confirmed the presence of tumor cells, which were epithelioid, moderately heterotypic with large nuclei and interstitial fibrosis (Fig. 2E). The patient was diagnosed with non-small cell lung cancer (IIIB) but refused anti-cancer treatment and discharged. Patient 5 had a similar clinical course, initially diagnosed with pneumonia upon admission, and subsequent mNGS analysis of bronchoalveolar lavage fluid sample suggested chromosomal number variations (Table 1). A later pathological biopsy confirmed the diagnosis of a malignant lung tumor (Fig. 2F), and the patient was transferred to another hospital for further treatment.

3.4. Diagnosis of pituitary tumor in addition to purulent meningitis

Malignancies can be accompanied by microbial infection that complicates diagnosis and treatment. Patient 6 (18-year-old male) was admitted due to headache, fever, and unconsciousness. The patient was unresponsive to verbal command but sensitive to light. One month prior to admission, the patient started to experience headache, fever, and limb convulsions. Two weeks prior to admission, the symptoms progressed to tachycardia and unconsciousness (blood pressure 100/53 mmHg). The patient was sent to ICU since septic shock was suspected. The biochemical tests of cerebrospinal fluid showed protein 1463.7mg/L, glucose 2.27mmol/L, chloride 126.9mmol/L. The levels of WBC and neutrophil were elevated (Table 1). MRI scan revealed multiple lesions occupying pituitary space
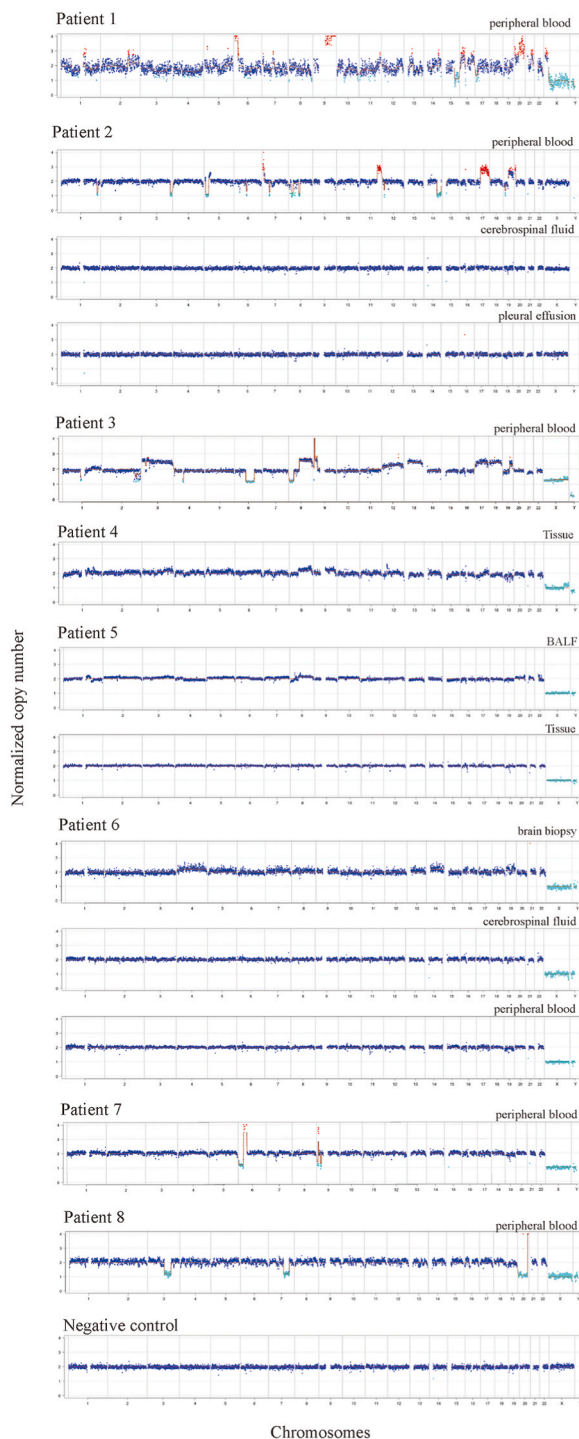
Fig. 1. Results of CNV analysis. The normalized copy number of every chromosome (chr1-22, X,Y) was plotted. The Y axis indicated copy number while X axis shows the chromosome number. The negative control was peripheral blood obtained from a healthy individual with no known inherited diseases or cancer.

and bilateral frontal straight gyrus. The signal of the adjacent meninges was enhanced, indicating invasive fungal infection and purulent meningitis. Meropenem was given and the patient's body temperature declined. Resection of the lesions was performed and sent for culture, which reported growth of *Acinetobacter baumannii*.

The tissue samples sent for pathological examination were mostly necrotic. Fibrous hyperplasia with inflammatory cell infiltration
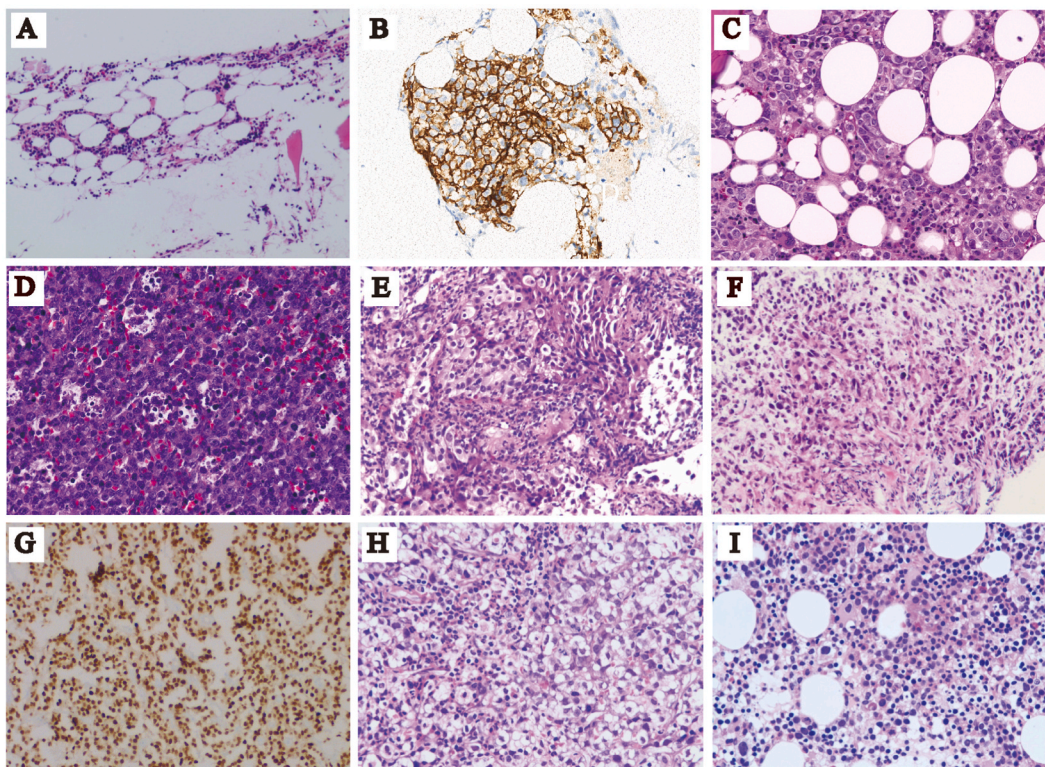
Fig. 2. H&E staining and immunohistochemical staining. H&E staining of bone marrow biopsy of patient 1 was shown in (A). CD20 staining (B) and H&E staining (C) of bone marrow biopsy of patient 2. H&E staining of lymph biopsy of patient 3 was shown in (D). H&E staining of lung biopsy of patient 4 (E) and 5 (F). Pituitary transcription factor 1 (Pit-1) staining of brain biopsy of patient 6 was shown in (G). H&E staining of bone marrow biopsy of patient 7 (H) and 8 (I).

surrounding the necrotic tissue was observed and tumor cell was not found. Proliferative osteoblasts were seen on the surface of bone trabecular with fibrous tissue hyperplasia. Ceftriaxone sodium, mannitol, euthyrox and hydrocortisone were administered, but the patient remained unconscious. Peripheral blood, biopsy and cerebrospinal fluid were sent for mNGS (Fig. 1). No pathogens were detected. Instead, chromosomal CNVs were discovered from biopsy. Immunohistochemical staining for growth hormone, PIT-1 (Fig. 2G) and thyroid-stimulating hormone were positive in the cells of brain biopsy. Purulent meningitis and plurihormonal PIT-1 positive pituitary adenoma were finally confirmed, and the patient was transferred for further treatment.

3.5. Diagnosis of acute myeloid leukemia (AML) in septic patients

For immunosuppressed cancer patients, infection can be fatal. Patient 8 (72-year-old male) was hospitalized due to dizziness and fatigue for more than 7 months. Blood routine test reported low blood cell count (WBC 3.22×10^9/L; RBC 3.07×10^{12}/L; HGB 94g/L; PLT 103×10^9/L). The patient developed fever, cough, and had difficulties in breathing. Pneumonia and sepsis were suspected, but no improvement of symptoms was seen using Tylenol 1g tid. PCT remained higher than 100ng/mL. Peripheral blood was sent for mNGS. Both pathogens (*Legionella pneumophila*, *Bacteroides fragilis*, *Enterococcus faecium*, human Polyomavirus) and chromosomal CNVs were detected (Fig. 1). The treatment was adjusted to linezolid (0.6g q12h intravenous injection), voriconazole (200mg q12h intravenous injection), and polymyxin B (first dose of 1 million units, followed by 500000 units q12h, intravenous injection) for anti-infection treatment. Furthermore, the bone marrow smear showed hyperplasia, decreased granulation, erythrocyte proliferation, and megaloblastic transformation (Fig. 2I). The proportion of bone marrow primordial cells was 22 %. The bone marrow biopsy had finally confirmed acute myeloid leukemia, and the patient was discharged.

4. Discussion

Metagenomic Next-Generation Sequencing is primarily employed in the diagnosis of infectious diseases. However, sequencing reads of host origin offers the potential for diagnosis of neoplasms. Deletion or duplication of chromosomes can be attributed to hereditary disorders or tumor-induced genomic instability. However, large CNVs (>10 million base pairs) involving multiple chromosomes are almost exclusively due to presence of tumor cells. Several techniques have been developed to analyze genetic copy number, including PCR-based, hybridization-based and microarray-based methods [23,24]. Among these, hybridization-based

methods are complex to perform with low resolution. PCR and microarray-based approaches are more commonly used. On the other hand, mNGS has the distinctive advantage over these methods since it can simultaneously detect pathogens and copy number of all chromosomes with a resolution of 100 kilo base pairs under typical sequencing depth (around 20 million reads).

Although CNVs does not necessarily equal to cancer diagnosis, they nevertheless serve as valuable clues to direct confirmatory tests for cancer and potentially save time, money, and efforts on finding an infectious etiology, especially when microbiological tests are negative and empirical antibiotics have little effect. For instance, both infectious diseases and cancer can trigger inflammatory responses, resulting in increased levels of markers such as C-Reactive Protein (CRP) and Procalcitonin (PCT). The process of differential diagnosis can be intricate and prolonged, as the clinical symptoms of various diseases often exhibit significant overlap. This was evident for patients 1, 2 and 3. In these cases, the clinical signs (symptoms and CT scans) were highly suggestive of pneumonia but instead, eventually diagnosed with lymphoma.

mNGS does not require intact cells while cytology and pathological tests rely on the quantity and integrity of cells for proper analysis. This enables mNGS to detect potential neoplasms while cytology fails to identify tumor cells. The clinical manifestations of patient 4 suggested lung abscess (cough, expectoration, and CT scan revealed necrotic lung cavities). Moreover, serum tumor markers were normal and imaging tests failed to identify cancerous lesions. These observations were more consistent with infectious diseases than cancer, despite negative microbiological tests. Under such circumstances, CNVs reported by mNGS have led to a more detailed clinical investigation on cancer.

Determining the sample type and site of sampling is essential for CNV analysis. For patient 2 (B-cell lymphoma), the CNVs were detected in peripheral blood, but not cerebrospinal fluid or pleural effusion. Indeed, B-cell lymphoma is a hematological malignancy that affects B lymphocytes, and thus CNVs may be more likely to be identified in blood. As a result, a normal copy number cannot rule out the possibility of neoplasms.

The cases of patients 5、6 and 8 highlighted the utility of mNGS in detecting both pathogens and neoplasms. Hematological patients are prone to microbial infections due to neutropenia, especially after chemotherapy, bone marrow transplantation and immunosuppressive therapy. For instance, *Legionella pneumophila* was identified by mNGS in patient 8 when conventional microbiological tests were negative. On the other hand, for patients like patient 7, who presented with fever of unknown origin and were admitted with a diagnosis of pancreatic lesions, mNGS can be utilized to differentiate between tumors and infections. In Patient 7's case, mNGS detected CNVs and did not identify any pathogens, aiding clinicians in ruling out infection. Subsequent pathological staining confirmed the diagnosis of a malignant tumor with a secondary bone marrow malignancy (Fig. 2H).

mNGS offers distinct advantages and limitations in cost-effectiveness compared to traditional cancer and pathogen detection methods. The advantages of mNGS include its ability of simultaneously detecting a wide range of pathogens or genetic mutations, with higher sensitivity and accuracy, particularly for complex samples or rare variants. It also offers quantitative data such as "reads", useful for assessing pathogen load or mutation abundance. Moreover, rapid diagnosis in identifying rare or novel pathogens; and future potential, as technological advances and cost reductions are expected to enhance its applicability and cost-effectiveness. On the other hand, the limitations of mNGS involve higher equipment and operational costs than some traditional methods and the requirement of specialized skills for operating and interpreting mNGS data.

As for CNV analysis, the biggest limitation is that CNV analysis using mNGS data can only detect gain or loss of chromosomes, but not gene mutations (point mutations, frameshift mutations etc.) due to limited depth. In addition, unlike karyotyping, structural variations that do not involve numerical copy number changes cannot be detected, such as balanced translocation and inversion. As a result, tumor cannot be ruled out with a negative CNV analysis. Another limitation is that experimental and quality control standards are currently lacking for mNGS, which are technically challenging to perform. It is conditionally recommended for febrile or immunocompromised patients when conventional tests fail to elucidate the etiology and empirical treatment has no effect. Moreover, the CNVs detected by mNGS can only suggest the presence of tumor DNA but cannot pinpoint the location or whether the tumor is benign or malignant. Therefore, a positive CNV result should be corroborated by additional diagnostic tests.

5. Conclusions

The established gold standard for diagnosing cancer is biopsy and subsequent pathological staining, which are procedures that are invasive and usually necessitate several days for completion. In contrast, mNGS can be carried out on a variety of clinical samples and yield results in hours. Our research indicates that for patients presenting with fever and suspected infections, mNGS emerges as a valuable technique. It offers timely differential diagnosis by integrating host Copy Number Variation analysis, simultaneously detecting both pathogens and neoplastic cells within a single test.

Ethics approval and consent to participate

We received approval by the board of Ethics at the Lingnan Hospital, Branch of the Third Affiliated Hospital of Sun Yat-sen University and Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences. The approval number of this study is **KY-N-2022-003-03**.

Consent for publication

Written informed consent was obtained from the patient for publication of this study and any accompanying images. A copy of the written consent is available for review by the Editor-in-Chief of this journal.

Funding

Data availability statement

The mNGS sequencing data supporting the conclusion of this article is available in the [NCBI SRA] database, [Accession number: **PRJNA776914;** and hyperlink to the database: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA776914]. The intermediate CNV data were deposited to [Zenodo]as txt files [DOI: 10.5281/zenodo.6578839; and hyperlink to the database: https://zenodo.org/record/6578839#.YpAyv8iP_GM].

CRediT authorship contribution statement

Feng Qin: Investigation, Conceptualization. **Xuejiao Hu:** Investigation, Conceptualization. **Xiaojia Wang:** Writing – original draft, Investigation, Conceptualization. **Weijiang Liu:** Formal analysis. **Qianyun Deng:** Formal analysis. **Yunhu Zhao:** Formal analysis. **Caiyun Li:** Data curation. **Chao Liu:** Writing – review & editing, Validation, Resources. **Zhenchao Huang:** Validation, Resources. **Bing Gu:** Validation, Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

[1] C.Y. Chiu, S.A. Miller, Clinical metagenomics, Nat. Rev. Genet. 20 (6) (2019) 341–355.
[2] R. Schlaberg, C.Y. Chiu, S. Miller, G.W. Procop, G. Weinstock, Professional practice C, committee on Laboratory Practices of the American Society for M, Microbiology Resource Committee of the College of American P: validation of metagenomic next-generation sequencing tests for universal pathogen detection, Arch. Pathol. Lab Med. 141 (6) (2017) 776–786.
[3] Y. Li, B. Sun, X. Tang, Y.L. Liu, H.Y. He, X.Y. Li, R. Wang, F. Guo, Z.H. Tong, Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in critically ill patients, Eur. J. Clin. Microbiol. Infect. Dis. 39 (2) (2020) 369–374.
[4] H. Li, H. Gao, H. Meng, Q. Wang, S. Li, H. Chen, Y. Li, H. Wang, Detection of pulmonary infectious pathogens from lung biopsy tissues by metagenomic next-generation sequencing, Front. Cell. Infect. Microbiol. 8 (2018) 205.
[5] T.A. Blauwkamp, S. Thair, M.J. Rosen, L. Blair, M.S. Lindner, I.D. Vilfan, T. Kawli, F.C. Christians, S. Venkatasubrahmanyam, G.D. Wall, et al., Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease, Nat Microbiol 4 (4) (2019) 663–674.
[6] R.J. Leary, M. Sausen, I. Kinde, N. Papadopoulos, J.D. Carpten, D. Craig, J. O'Shaughnessy, K.W. Kinzler, G. Parmigiani, B. Vogelstein, et al., Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing, Sci. Transl. Med. 4 (162) (2012) 162ra154.
[7] W. Gu, E. Talevich, E. Hsu, Z. Qi, A. Urisman, S. Federman, A. Gopez, S. Arevalo, M. Gottschall, L. Liao, et al., Detection of cryptogenic malignancies from metagenomic whole genome sequencing of body fluids, Genome Med. 13 (1) (2021) 98.
[8] W. Gu, A.M. Rauschecker, E. Hsu, K.C. Zorn, Y. Sucu, S. Federman, A. Gopez, S. Arevalo, H.A. Sample, E. Talevich, et al., Detection of neoplasms by metagenomic next-generation sequencing of cerebrospinal fluid, JAMA Neurol. 78 (11) (2021) 1355–1366.
[9] W.F. Wright, P.G. Auwaerter, Fever and fever of unknown origin: review, recent advances, and lingering dogma, Open Forum Infect. Dis. 7 (5) (2020) ofaa132.
[10] S. Sharma, Tumor markers in clinical practice: general principles and guidelines, Indian J. Med. Paediatr. Oncol. 30 (1) (2009) 1–8.
[11] G. Veronesi, L.L. Travaini, P. Maisonneuve, C. Rampinelli, R. Bertolotti, L. Spaggiari, M. Bellomi, G. Paganelli, Positron emission tomography in the diagnostic work-up of screening-detected lung nodules, Eur. Respir. J. 45 (2) (2015) 501–510.
[12] L. Valls, C. Badve, S. Avril, K. Herrmann, P. Faulhaber, J. O'Donnell, N. Avril, FDG-PET imaging in hematological malignancies, Blood Rev. 30 (4) (2016) 317–331.
[13] A. Shlien, D. Malkin, Copy number variations and cancer, Genome Med. 1 (6) (2009) 62.
[14] J.F. Xu, Q. Kang, X.Y. Ma, Y.M. Pan, L. Yang, P. Jin, X. Wang, C.G. Li, X.C. Chen, C. Wu, et al., A novel method to detect early colorectal cancer based on chromosome copy number variation in plasma, Cell. Physiol. Biochem. 45 (4) (2018) 1444–1454.
[15] M. Kumaran, C.E. Cass, K. Graham, J.R. Mackey, R. Hubaux, W. Lam, Y. Yasui, S. Damaraju, Germline copy number variations are associated with breast cancer risk and prognosis, Sci. Rep. 7 (1) (2017) 14621.
[16] L. Lenaerts, P. Vandenberghe, N. Brison, H. Che, M. Neofytou, M. Verheecke, L. Leemans, C. Maggen, B. Dewaele, L. Dehaspe, et al., Genomewide copy number alteration screening of circulating plasma DNA: potential for the detection of incipient tumors, Ann. Oncol. 30 (1) (2019) 85–95.
[17] Y. Luan, H. Hu, C. Liu, B. Chen, X. Liu, Y. Xu, X. Luo, J. Chen, B. Ye, F. Huang, et al., A proof-of-concept study of an automated solution for clinical metagenomic next-generation sequencing, J. Appl. Microbiol. 131 (2) (2021) 1007–1016.
[18] H. Shen, D. Shen, H. Song, X. Wu, C. Xu, G. Su, C. Liu, J. Zhang, Clinical assessment of the utility of metagenomic next-generation sequencing in pediatric patients of hematology department, Int. J. Lab Hematol. 43 (2) (2020) 244–249.
[19] M.R. Wilson, H.A. Sample, K.C. Zorn, S. Arevalo, G. Yu, J. Neuhaus, S. Federman, D. Stryke, B. Briggs, C. Langelier, et al., Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis, N. Engl. J. Med. 380 (24) (2019) 2327–2340.
[20] M. Fromer, J.L. Moran, K. Chambert, E. Banks, S.E. Bergen, D.M. Ruderfer, R.E. Handsaker, S.A. McCarroll, M.C. O'Donovan, M.J. Owen, et al., Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth, Am. J. Hum. Genet. 91 (4) (2012) 597–607.

[21] D. Backenroth, J. Homsy, L.R. Murillo, J. Glessner, E. Lin, M. Brueckner, R. Lifton, E. Goldmuntz, W.K. Chung, Y. Shen, CANOES: detecting rare copy number variants from whole exome sequencing data, Nucleic Acids Res. 42 (12) (2014) e97.

[22] Y. Guo, H. Li, H. Chen, Z. Li, W. Ding, J. Wang, Y. Yin, L. Jin, S. Sun, C. Jing, et al., Metagenomic next-generation sequencing to identify pathogens and cancer in lung biopsy tissue, EBioMedicine 73 (2021) 103639.

[23] C.P. Gonzalez-Hunt, J.P. Rooney, I.T. Ryde, C. Anbalagan, R. Joglekar, J.N. Meyer, PCR-based analysis of mitochondrial DNA copy number, mitochondrial DNA damage, and nuclear DNA damage, Curr Protoc Toxicol 67 (2016), 20 11 21-20 11 25.

[24] N.P. Carter, Methods and strategies for analyzing copy number variation using DNA microarrays, Nat. Genet. 39 (7 Suppl) (2007) S16–S21.