



OPEN ACCESS

EDITED BY

Ana Cristina Sampaio,
Universidade de Trás-os-Montes e Alto
Douro, Portugal

REVIEWED BY

Tim Sandle,
The University of Manchester,
United Kingdom
Filip Boyen,
Ghent University, Belgium
Mona T. Kashef,
Cairo University, Egypt

*CORRESPONDENCE

Yiling Fan
✉ tcfyl@163.com
Hong Shao
✉ shaohong@smda.sh.cn
Meicheng Yang
✉ yangmeicheng@vip.sina.com

RECEIVED 24 October 2023

ACCEPTED 30 January 2024

PUBLISHED 12 February 2024

CITATION

Song MH, Li QQ, Liu CZ, Wang PE, Qin F,
Zhang LC, Fan YL, Shao H, Chen GL and
Yang MC (2024) A comprehensive technology
strategy for microbial identification and
contamination investigation in the sterile drug
manufacturing facility—a case study.
Front. Microbiol. 15:1327175.
doi: 10.3389/fmicb.2024.1327175

COPYRIGHT

© 2024 Song, Li, Liu, Wang, Qin, Zhang, Fan,
Shao, Chen and Yang. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

A comprehensive technology strategy for microbial identification and contamination investigation in the sterile drug manufacturing facility—a case study

Minghui Song^{1,2}, Qiongqiong Li^{1,2}, Chengzhi Liu³, Peien Wang^{1,2},
Feng Qin¹, Lichun Zhang⁴, Yiling Fan^{1,2,5*}, Hong Shao^{1,2*},
Guiliang Chen^{5,6} and Meicheng Yang^{1,5*}

¹NMPA Key Laboratory for Testing Technology of Pharmaceutical Microbiology, Shanghai Institute for Food and Drug Control, Shanghai, China, ²Shanghai Quality Inspection and Testing Center for Innovative Biological Products, Shanghai, China, ³Hangzhou Digital-Micro Biotech Co., Ltd., Hangzhou, China, ⁴Shanghai SPH New Asia Pharmaceutical Co., Ltd., Shanghai, China, ⁵China State Institute of Pharmaceutical Industry Co., Ltd., Shanghai, China, ⁶Shanghai Center for Drug Evaluation and Inspection, Shanghai, China

Objective: A comprehensive strategy for microbial identification and contamination investigation during sterile drug manufacturing was innovatively established in this study, mainly based on MALDI-TOF MS for the identification and complemented by sequencing technology on strain typing.

Methods: It was implemented to monitor the bacterial contamination of a sterile drug manufacturing facility, including its bacterial distribution features and patterns. In three months, two hundred ninety-two samples were collected covering multiple critical components of raw materials, personnel, environment, and production water.

Results: Based on our strategy, the bacterial profile across the production process was determined: 241/292 bacterial identities were obtained, and *Staphylococcus* spp. (40.25%), *Micrococcus* spp. (11.20%), *Bacillus* spp. (8.30%), *Actinobacteria* (5.81%), and *Paenibacillus* spp. (4.56%) are shown to be the most dominant microbial contaminants. With 75.8% species-level and 95.4% genus-level identification capability, MALDI-TOF MS was promising to be a first-line tool for environmental monitoring routine. Furthermore, to determine the source of the most frequently occurring *Staphylococcus cohnii*, which evidenced a widespread presence in the entire process, a more discriminating *S. cohnii* whole-genome SNP typing method was developed to track the transmission routes. Phylogenetic analysis based on SNP results indicated critical environment contamination is highly relevant to personnel flow in this case. The strain typing results provide robust and accurate information for the following risk assessment step and support effective preventive and corrective measures.

Conclusion: In general, the strategy presented in this research will facilitate the development of improved production and environmental control processes for the pharmaceutical industry, and give insights about how to provide more sound and reliable evidence for the optimization of its control program.

KEYWORDS

microbial identification, pharmaceutical process, contamination investigation, MALDI-TOF MS, sequencing technology

Introduction

Microbial contamination of pharmaceutical products continues to be a severe challenge to drug manufacturing. Drug products, especially high-risk sterile products such as injection and ophthalmic products, can cause irreversible damage once contamination occurs during the production and storage process, in severe cases, detrimental to the patient's health and even their lives. It was shown that between 2012 and 2019, more than 50% of all the recalls related to drug products registered by the US Food and Drug Administration (FDA) were associated with microbiological incidents (Santos et al., 2018; Jimenez, 2019). The high frequency of recalls indicates inadequate awareness of contamination risk and poor implementation of relevant control programs. Substantial evidence such as warning letters, alert notifications, and failures suggests a direct relationship between the level of environmental control and the product's final quality (Jimenez, 2007; Sandle, 2012; Jain and Jain, 2018).

Understanding the types and distribution of environmental microflora during drug manufacturing is essential to the company's in-process surveillance program (Hamdy et al., 2018). Once a relevant microflora database is established, it can be used for investigations of product contamination, environmental excursions, etc. The trend changes in the type and number of predominant species may also signal a weakness in contamination control strategy (CCS) and trigger some preventive and corrective actions (Akers, 1997; Halls, 2004). In addition, the database can also provide further insight into other elements of the CCS, such as the design and evaluation of an effective disinfection scheme (Wu and Liu, 2007). It is crucial to implement a microbial identification and characterization framework as part of an effective microbial contamination control strategy. While the consensus has been established and higher expectations have been placed in the latest regulations and guidance, relatively few studies are associated with the microbial profile in clean pharmaceutical environments (Wu and Liu, 2007; Sandle, 2011; Vijayakumar et al., 2012, 2016; Park et al., 2014). Methodologies employed in those studies range from biochemical methods to genomics, which are either complicated, less accurate, or expensive to promote in industry. A fast, accurate, and cost-effective solution, along with a more rapid and efficient tool to speed up and improve the quality of the analytical results, is highly desirable for implementing a comprehensive monitoring system.

Several methods are available for rapid and automated microorganism identification and typing, generally classified as phenotypic or genotypic techniques (Ferone et al., 2020). Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS), a methodology based on the extraction of protein fingerprint signatures from the whole bacterial cells, is raising interest. MALDI-TOF MS has gained extensive clinical use due to its rapid and accurate bacterial identification capabilities (Bizzini and Greub, 2010; Patel, 2015; Singhal et al., 2015). Previous studies showed that bacterial identification using MALDI-TOF MS was significantly better than conventional biochemical systems. It was concluded that MALDI-TOF MS could be implemented quickly to routinely identify microorganisms in a clinical microbiology laboratory (Schubert and Kozrzewa, 2017). However, the progression in using this technique for bacterial identification in other industries like the pharmaceutical area is very slow, this is due to the limited availability of relevant mass

spectral libraries, and the bias of existing libraries more focused on clinical isolates (Andrade et al., 2018).

In addition to understanding the bacterial species profiles of the production environment, early detection of cross-transmissions is also critically important in blocking the spread of microbial contamination and quickly implementing control measures to minimize the damages. This introduced a demand for more discriminating methods of bacterial characterization-strain typing, which can characterize and confirm epidemiological linkage in an area and provide insights into bacterial population dynamics (Scott et al., 2002). Multiple methods based on different phenotypic and genotypic principles have been adopted to decipher epidemiological links to identify reservoirs and routes of transmission (Simar et al., 2021); among these, molecular biology methods with a relatively high discriminatory power such as multi-locus sequence typing (MLST) and pulsed-field gel electrophoresis (PFGE) have been proved to be beneficial for surveillance and contamination investigation (Simar et al., 2021). Strain typing methodologies have recently undergone a paradigm shift as whole-genome sequencing (WGS) has become more accessible to clinical and industry laboratories. WGS provides higher discriminatory power for highly related strains, and it has significant potential for outbreak detection, epidemiological surveillance, and infection control strategies (Miro et al., 2020; Uelze et al., 2020; Simar et al., 2021; Menghwar et al., 2022). The identification of single-nucleotide polymorphisms (SNPs) between bacterial isolates is one of the most commonly used analyses of WGS data and has been widely used as an alternative to MLST and PFGE on several bacterial species, facilitating its use in outbreak investigations (McDonald et al., 2006; Roe et al., 2016; Hoefler et al., 2021).

The objectives of this study were to establish a comprehensive strategy for microbial identification and investigation during the drug manufacturing process. The recovered isolates from a monitoring system covering the entire production were classified using a Bruker MALDI-TOF MS system and 16S rRNA gene sequencing combination approach to determine the predominant bacterial flora distributed in the cleanroom environment, the capability of the current MALDI-TOF MS system and its commercial database to identify pharmaceutical industry-relevant contaminants was evaluated. The bacterial distribution features and patterns were further reviewed, and a more discriminative WGS typing method was used to investigate the transmission routes of the predominant and critical contaminating *Staphylococcus cohnii*. This study intends to provide the pharmaceutical industry with a solution on how to understand the microbial risks of the entire manufacturing process.

Materials and methods

Environmental monitoring plan and sampling

From June to August 2018, the environmental monitoring program was carried out in a sterile powder for injection drug manufacturing facility in Shanghai, China. Samples were taken from its two primary production areas: formulation manufacture (MF) and substance manufacture (MS), both areas maintained the 4 grades of cleanrooms. Samples were taken from multiple elements of the production process including finished products, raw materials,

personnel, production environment, and water. Different pharmacopeial techniques were used: (a) active air sampling using a volumetric air sampler; (b) passive air sampling using settle plates; (c) swabs for irregular surface sampling; (d) contact plates for flat surfaces; (e) 100 mL membrane filtration onto Reasoner's 2A Agar medium (R2A, bioMérieux, France) for water for injection. Trypticase Soy Agar medium (TSA, bioMérieux, France) was used for the sampling except for the water samples, and a neutralizer to neutralize the effects of disinfectant residues on surfaces was added when necessary. An intensive sampling strategy was implemented, and the sampling frequencies were set according to the potential microbiological contamination risk assessed for different monitoring items and scopes. In general, settle plates were exposed for the duration of operations and changed after 4 h for grades A and B; active air sampling was conducted for each shift. Both settle plates and active air samples were taken for the grade C area for each batch. For hard surfaces like the product contact area, filling machine, and relevant components, sampling was conducted every shift for grades A and B, and each week for grade C. The collected plates were under 30°C–35°C aerobic incubation for 3–5 days for bacterial growth. The specific sample information is listed in [Supplementary Table S1](#).

Cultivation and isolation

The sample information was recorded in detail according to the rules interpreted in [Supplementary Table S1](#). ChP chapter <9205> ([Chinese Pharmacopoeia Commission, 2020](#) edition) were referred to detect and isolate microorganisms from the samples. In general, after incubation, the plates with microorganism growth were reserved for determination of the next steps: any growth on the plates collected from the A/B grade areas were subcultured and identified; for bacterial colonies on the plates collected from the C/D grade areas, only representative colonies from the same plate, as determined by colony morphology and microscopic examination, were selected for subsequent subculture and identification. The selected colonies were transferred to TSA medium (bioMérieux, France), incubated under appropriate conditions (30°C–35°C aerobic incubation for 3–5 days), and then examined. The type of colonies on the plate was assessed primarily based on the morphological characteristics and recorded. Distinctive colonies were selected and subcultured onto TSA plates and incubated under 30°C–35°C for 1–3 days to get fresh and purified culture for subsequent preservation, identification, and investigation. The pure cultures were picked and fully suspended into 0.5 mL tryptic soy broth (TSB, bioMérieux, France) vials, and then added same volume of 30% sterile glycerol solution. The cryovials were preserved at –20°C.

MALDI-TOF MS identification

Preparations of bacterial isolates for MALDI-TOF MS measurement were done as previously described ([Croxatto et al., 2012](#)). The on-target extraction preparation method was used to identify all the isolates to establish a standardized rapid routine identification process. Briefly, one fresh colony material was picked and smeared onto a 384 polished steel MSP target (Bruker Daltonik GmbH, Germany) using a toothpick, then one μL of 70% formic acid was added onto the bacterial spot and allowed to air dry afterward add

one μL of 10 mg/mL *a*-cyano-4-hydroxy-cinnamic acid (HCCA, Bruker Daltonik GmbH, Germany) matrix solution (50% acetonitrile; 2.5% trifluoroacetic acid; 47.5% distilled water) and allowed to air dry. The acquisition and analysis of mass spectra were then performed by an Autoflex TOF/TOF mass spectrometer (Bruker Daltonik GmbH, Germany) under the linear positive mode, MALDI Biotyper software package (version Compass) with the reference 7,854 database entries and default parameter settings was used. The Bruker Bacterial Test Standard (BTS) was used for calibration according to the manufacturer's instructions. Two colony/sample material preparations were analyzed for each isolate, and the higher-scoring identification result was taken.

16S rRNA gene sequencing analysis

All isolates preserved after purification from the environmental monitoring plan were sub-cultured on TSA medium at 35°C for 24 h during identification. The whole DNA was extracted using a TAKARA DNA extraction kit (Takara Bio Inc., Shiga, Japan) following the manufacturer's instructions. Primer 27F (5'-AGTTTGATCMTGGCTC AG-3') and 1492R (5'-GGTAC CTTGTTACGACTT-3') were used for PCR amplification. PCR products were resolved and visualized by electrophoresis in gels containing 1.2% agarose and 0.5 $\mu\text{g}/\text{mL}$ ethidium bromide. After that, all the amplified DNA fragments were sequenced commercially in both directions (Sangon Co., Ltd., Shanghai), and the sequence reads were analyzed using Lasergene 7.0 (DNASTar, Madison, WI). A query sequence with 98.65% identity to a reference in GenBank was acceptable for species identification ([Kim et al., 2014](#)).

Whole-genome sequencing of *Staphylococcus cohnii*

Genomic DNAs of all 33 *S. cohnii* isolates were extracted from overnight cultures by using a Wizard genomic DNA purification kit (Promega, Beijing, China) and were subjected to whole-genome sequencing using 150 bp pair-end strategies using the Illumina Miseq platform (Illumina Inc., San Diego, CA, United States). All operations followed standard Illumina protocols. Raw reads were trimmed and assembled to contigs using SPAdes version 3.11. Using FDAARGOS_538(GCA_003956025.1) as a reference genome, single-nucleotide polymorphisms (SNPs) were called by the Gngnr online software of Harvestsuite.¹ The phylogenetic trees were analyzed by parsnp software based on the whole genomes, and the trees were visualized with FigTree software (Version 1.4.3) and Adobe Illustrator software (Version 22.1). Homologous recombination events were detected by Gubbins v3.1.3² and related sites were filtered in the subsequent analysis. After filtering, 172 wgSNPs were obtained ([Croucher et al., 2014](#)). Simpson's index of diversity (SID), a discriminatory index of a microbial typing method, was calculated as previously described ([Hunter and Gaston, 1988](#)).

1 <http://github.com/marbl/harvest>

2 <https://github.com/nickjcroucher/gubbins>

Results

The abundance of microorganisms at different processing locations

From June to August 2018, 292 isolates were recovered and purified from the production process of a sterile drug manufacturing facility, including the materials, personnel, environment, and the production water system, respectively. The bacterial contamination distribution is shown in Table 1. The data showed that the largest contamination source was the production environment (212 isolates), occupying more than 70% of all isolates, followed by the water system (39 isolates). Further examination of the contamination distribution in different grade areas indicated that the highest rate (52%) was found in grade C clean areas, with a decreasing tendency to the higher-grade cleanroom. The increased quantity and diversity of contaminants suggested a wide range of contamination sources. Notably, 30 isolates were recovered from the grade A production area, demonstrating the existence of severe contamination within this production environment, and relevant action should be taken urgently.

Characterization of microflora in the pharmaceutical production environment

The 292 isolates were identified using a multiplex method of MALDI-TOF MS and 16S rRNA gene sequencing. Then, the potential duplicates of the 292 isolates were excluded, combining with the isolation context, growth features, colony morphology and species information. The identification results of the 241 selected isolates using MALDI-TOF MS method are given in Supplementary Table S2. Of these samples, 189 confident results were obtained, including 147 (61.0%) with species-level results (score ≥ 2.0) and 42 (17.4%) with genus-level results ($1.7 \leq \text{score} < 2.0$) according to the manufacturer's threshold interpretation. Non-reliable identification results (score < 1.7) were given to the remaining 52 isolates (21.6%).

The total of 94 isolates with low confidence ($1.7 \leq \text{score} < 2.0$) and no identification results were further subjected to 16S rRNA gene sequencing as the flow illustrated in Figure 1 as complementation. The sequencing details are shown in Supplementary Table S2. Among the 94 unreliable results obtained by MALDI-TOF MS, 79 isolates were identified as species-level by the matching criteria ($\geq 98.65\%$ similarity), and 15 isolates were identified as genus-level. Integrating the results of the whole identification module, it was finally

determined that the 241 isolates were from 44 genera and 94 species. With the high-throughput capability, all mass spectrometry identifications can be accomplished within a single batch on the same day.

As shown in Table 2; Supplementary Table S2, comparing the 16S rRNA gene sequencing identification results with the MALDI-TOF MS commercial database revealed that of the 94 unreliable MALDI-identifies, 46 isolates from 31 species were not covered by the commercial MALDI database. If we remove those out-of-coverages, the identification capacity of MALDI-TOF MS method could reach 75.8% species-level and 95.4% genus-level. The most significant discrepancy impacted by database coverage was with *Bacillus*, where the accuracy of species-level identification improved from 30% to 85.7% by removing non-coverage species from the identification set.

The microflora in the pharmaceutical production environment

Of the 241 identified isolates, which represent 44 genera and 94 species, the most predominant species/genera (≥ 3 isolations) are shown in Table 2. As demonstrated, the most significant proportion of genus-level contamination was associated with *Staphylococcus* spp. (40.25%), *Micrococcus* spp. (11.20%), *Bacillus* spp. (8.30%), *Actinobacteria* (5.81%), and *Paenibacillus* spp. (4.56%), the top 5 genera contributing to 70% of all the isolates. The 19 predominant species accounted for 64% of all 94 species isolated, with a distinct predominance of *Staphylococcus* spp., which occupied 9 of the 19 species that comprise most of the microflora, of which *S. cohnii*, *S. epidermidis*, and *S. hominis* were detected more than 10 times. *Micrococcus luteus*, was also shown to be a frequently occurring contaminant (Eissa, 2016).

The investigation plan further classified the monitoring objective into airborne (settle plates and volumetric air sampling) and surface flora (contact plates and swabs) for a more specific analysis. As illustrated in Figure 2, the most diverse microflora in quantity and type was airborne flora, covering 26 different genera, with the majority recovered of *Staphylococcus* spp. (69%) and *Micrococcus* spp. (74%). Within the predominant genus listed in Table 3 (\geq three isolations), four genera were only detected from airborne sampling. Unexpectedly, no spore-producing *Bacillus* species was identified from the air samples of this study, which deviates from the conventional understanding that most spores could be transmissible through the air, which could be related to the selection of sampling locations and

TABLE 1 The distribution of microorganisms in different locations categorized by number of isolates, genus, and species.

Environmental monitoring objectives		Numbers of isolates		Percentage (%)	Genus	Species
Raw materials		30		10.27	9	19
Personnel		11		3.77	8	8
Production environment	A	212	30	72.60	8	16
	B		71		15	28
	C		111		16	31
Water		39		13.36	14	24
Total		292		100.00	44	94

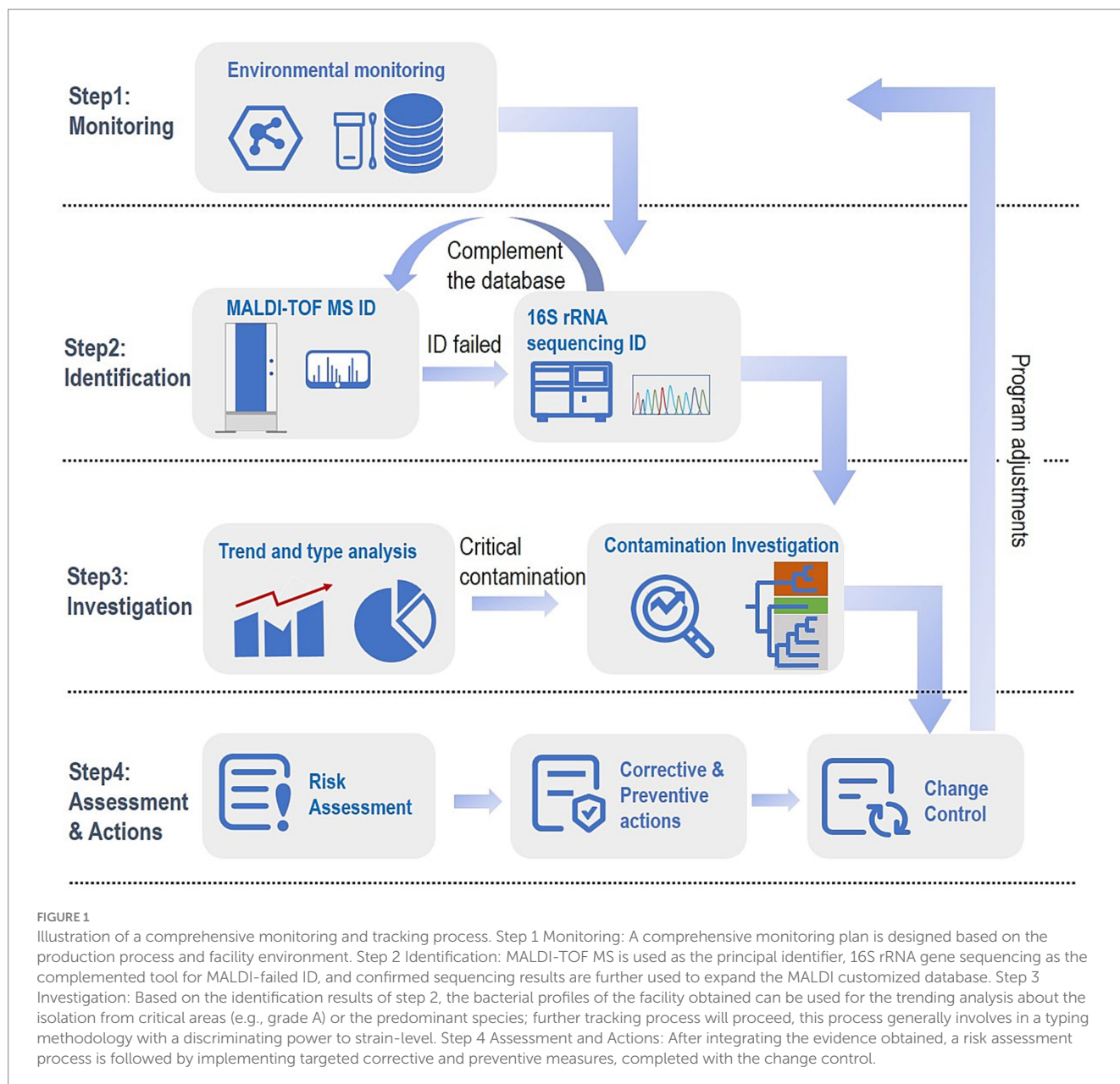
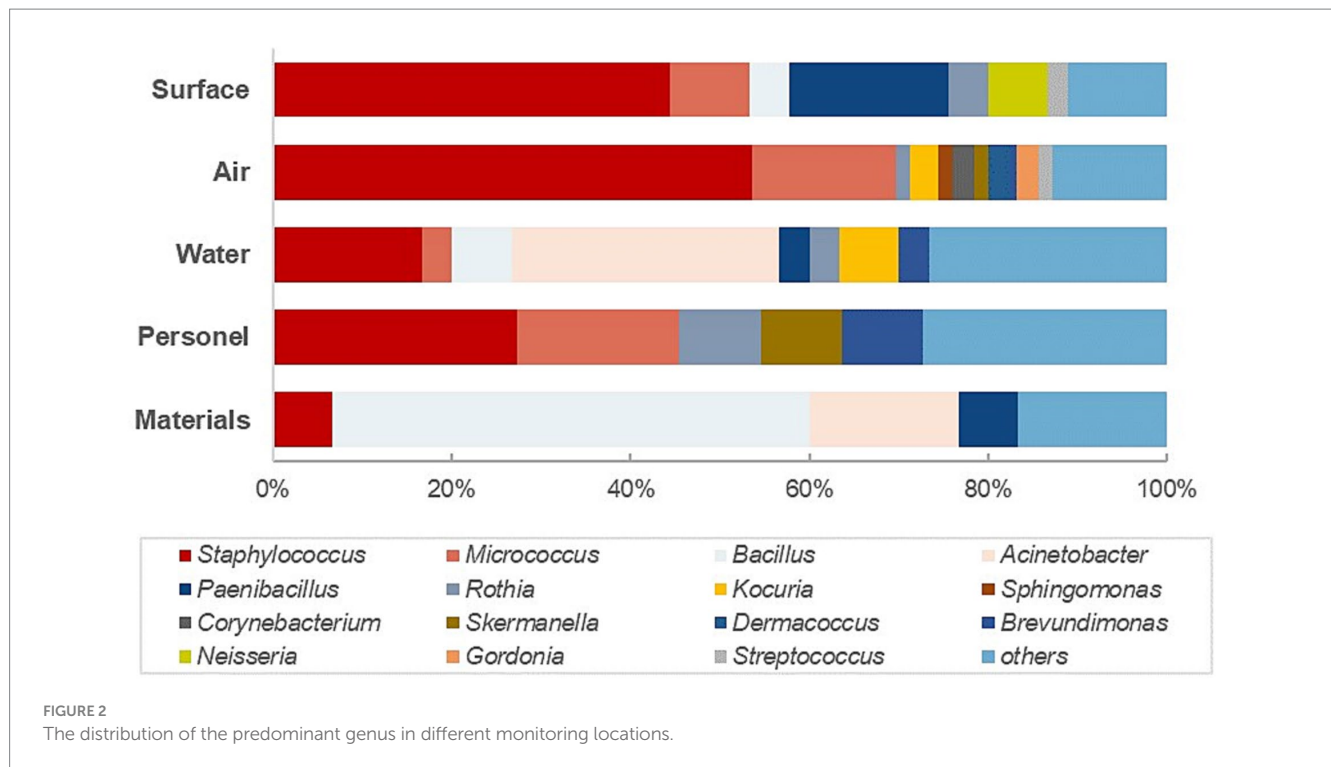


TABLE 2 Evaluation of the identification capability of MALDI-TOF MS for all isolates collected and the dominant contaminated genus in this study.

Genus	MALDI ^a claimed covered species	Isolates in this study	MALDI not covered	Identification results for all isolates		Identification results only MALDI covered	
		Isolates (species)	Isolates (species)	Species level	Genus level	Species level	Genus level
<i>Staphylococcus</i>	45	97 (10)	3 (1)	80.4% (78)	13.4% (13)	83%	13.8%
<i>Micrococcus</i>	6	27 (4)	0 (0)	85.2% (23)	14.8% (4)	85.2%	14.8%
<i>Bacillus</i>	108	20 (10)	13 (6)	30% (6)	5% (1)	85.7%	14.3%
<i>Actinobacteria</i>	39	14 (4)	1 (1)	57.1% (8)	21.4% (3)	61.5%	23.1%
<i>Paenibacillus</i>	85	11 (3)	0 (0)	72.7% (8)	27.3% (3)	72.7%	27.3%
In total	2,969	241 (92)	46 (31)	60% (147)	17.4% (42)	75.4%	21.5%

^aMALDI here specifically refers to the MALDI system and database version used in this study, as stated in Materials and Methods section.



sampling protocols. The same with the highly abundant detection of Gram-positive cocci in the water system which might be related to inappropriate sampling practices. These are the aspects that need to be improved in the subsequent design and execution of the environmental monitoring protocols. In contrast, 58% of the bacterial spore-forming species were isolated from raw materials and 32% from environmental surfaces, highlighting the critical importance of raw material quality control and transfer disinfection.

Contamination source tracking of *Staphylococcus cohnii* using the WGS typing method

As the present results revealed, *Staphylococcus* spp. is the primary microbial contaminant during this drug production process, accounting for 40.25% of the viable contaminant, of which species *S. cohnii* was the most critical one, contributing 33 isolates recovered in this study, the information of the isolates is given in [Supplementary Tables S1, S2](#). It was shown that *S. cohnii* were recovered from grade A (3), grade B (7), grade C (20) environments, personnel (2), as well as raw materials (1), evidencing its widespread presence in the entire process. In the two key areas where the environmental monitoring was conducted, the majority (28/33 isolates) of the recoveries of *S. cohnii* were from the formulation manufacture (MF), and only five were from the substance manufacture (MS). As shown in [Figure 3](#), there were two critical areas (grade A) of contamination presence among the 28 isolates recovered from MF, and these isolates were further subjected to a retrospective investigation.

The contamination investigation of the 28 *S. cohnii* strains was carried out with a *Staphylococcus* whole genome SNP typing method established by our laboratory. The analysis results are shown in [Figure 3](#).

Twenty-eight isolates were classified into 11 genotypes (SID=0.765), with three clusters, A, B, and C containing more than two isolates. The strain H71 which recovered from MF's critical area grade A—laminar flow cabinet, was classified into clustered A with 13 nearly identical isolates (between 20 and 172 SNPs differences). Strain H71 was closely related to strains recovered from several sampling points located in its grade C background area, i.e., airlocks (G2, G1, G32, G39, F28), capping rooms (H28, G30 F30), corridor (F8, G19), etc., indicating a high probability that the contamination of the grade A surface might originate from the lower-level grade C background area where the transmission occurred either through the material, air or personnel flow, considering its process and isolation level, the latter is more probably. The cluster B suggested a high potential for introducing contamination through poor personnel practice or movement, with strain K6 found in the grade A area being closely related to strain H66 (38 SNPs), which was recovered from the right wrist of work staff.

Unlike the MF area, where a more significant transmission route was observed, the 5 *S. cohnii* strains recovered from the MS production area did not show substantial relevancy to each other, indicating the presence of *S. cohnii* in the area was incidental and associated transmission has not yet occurred (data not shown).

Discussion

Based on the understanding and assessment of a sterile drug manufacturer's process and facility, a microbial surveillance scheme in the present study, including monitoring, identification, contamination tracking, and risk assessment, was established. The types and distribution of microflora across the entire production process were examined comprehensively. A total of 292 bacterial isolates of bacteria were recovered within 3 months, with a wide range

TABLE 3 Most commonly occurring species/genera in the pharmaceutical production environment.

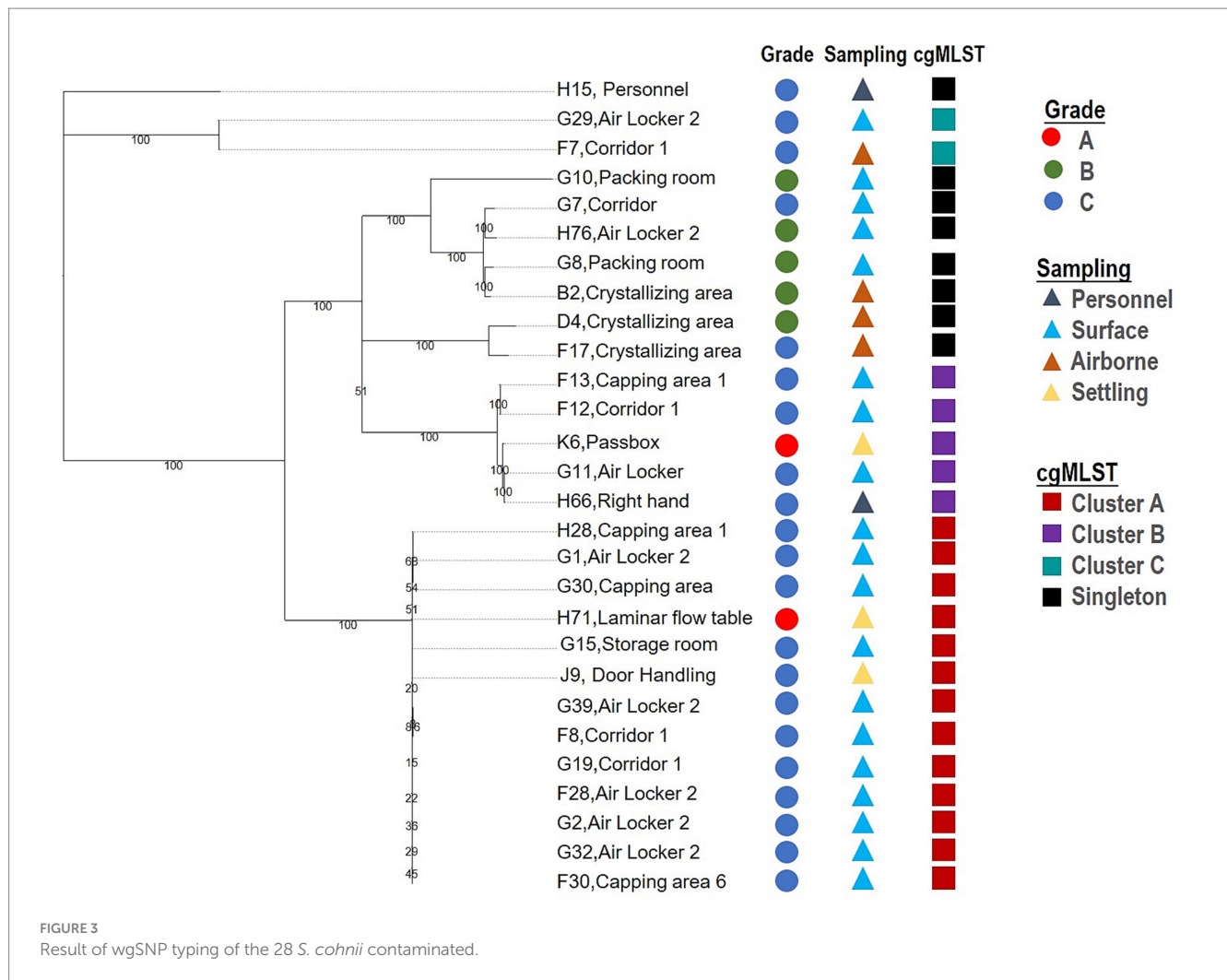
Species	Isolation	Percentage
<i>Staphylococcus cohnii</i>	33	13.69%
<i>Staphylococcus epidermidis</i>	27	11.20%
<i>Micrococcus luteus</i>	17	7.05%
<i>Staphylococcus hominis</i>	10	4.15%
<i>Paenibacillus lautus</i>	9	3.73%
<i>Staphylococcus haemolyticus</i>	6	2.49%
<i>Acinetobacter pittii</i>	6	2.49%
<i>Staphylococcus nepalensis</i>	5	2.07%
<i>Bacillus aryabhatai</i>	5	2.07%
<i>Staphylococcus pasteurii</i>	4	1.66%
<i>Staphylococcus capitis</i>	4	1.66%
<i>Micrococcus cohnii</i>	4	1.66%
<i>Micrococcus antarcticus</i>	4	1.66%
<i>Staphylococcus saccharolyticus</i>	3	1.24%
<i>Staphylococcus argensis</i>	3	1.24%
<i>Rothia amarae</i>	3	1.24%
<i>Bacillus paralicheniformis</i>	3	1.24%
<i>Bacillus megaterium</i>	3	1.24%
<i>Acinetobacter nosocomialis</i>	3	1.24%
Others	89	36.93%

Genus	Isolation	Percentage
<i>Staphylococcus</i>	97	40.25%
<i>Micrococcus</i>	27	11.20%
<i>Bacillus</i>	20	8.30%
<i>Acinetobacter</i>	14	5.81%
<i>Paenibacillus</i>	11	4.56%
<i>Rothia</i>	6	2.49%
<i>Kocuria</i>	4	1.66%
<i>Sphingomonas</i>	4	1.66%
<i>Corynebacterium</i>	3	1.24%
<i>Skermanella</i>	3	1.24%
<i>Dermacoccus</i>	3	1.24%
<i>Brevundimonas</i>	3	1.24%
<i>Neisseria</i>	3	1.24%
<i>Gordonia</i>	3	1.24%
<i>Streptococcus</i>	3	1.24%
Others	37	15.36%

of sources covering all parts of the production process and different grades of clean areas, prompting a high contamination risk and the urgency to optimize its microbial surveillance program and related prevention measures, such as an adequate cleaning and disinfection program.

In this study, we used MALDI-TOF MS as the first-line identification method, complemented by sequencing, enabling a quick characterization of production environmental flora, and the

development of a rapid and effective surveillance scheme. As numerous studies have confirmed that MALDI-TOF MS is a fast, simple, and high-throughput proteomic technique for the identification of a variety of bacterial species, in the present study, where database coverage was taken into consideration, the identification accuracy could be improved to 75.8% species-level and 95.4% genus-level, this also adequately confirms MALDI-TOF MS can be utilized as a first-line tool for routine environmental monitoring.



But based on the commercial database, only 77.4% of genus-level identities and 60.0% of species-level identities were obtained; when it comes to the case of *Bacillus* spp., the results are even worse (30% species-level and 35% genus-level), though the commercial database contains over 100 species of *Bacillus*. This result is lower than some evaluation studies from clinics, which can typically achieve 90% or higher species-level identification (Croxatto et al., 2012; Patel, 2015; Schubert and Kostrzewa, 2017), This discrepancy observed may arise from the lack of sufficient reference spectrums in the commercial database of MALDI-TOF MS systems. Even though the species-level may be covered by the database, the diversity of strains may be incomplete, as these isolates are collected from the highly controlled and disinfected clean environments, the strains are more stressed than the clinical collected isolates. Our findings are concordant with Seuylemezian's et al. (2018) studies on clean rooms of spacecraft, which under more rigorous microbial reduction techniques and the manufacturer-provided database yielded successful identification of only 8% of bacterial isolates. These results demonstrated that commercial MALD-TOF databases need to be enhanced by the inclusion of more isolations from pharmaceutical cleanroom facilities, which requires the developers of MALDI-TOF Mass

systems to work more closely with pharmaceutical manufacturers. Meanwhile, because of the differences of MALDI users existing in geographic areas, product types, and production processes, some unique strains may exist in their manufacturing environments, using these strains to establish their own customized reference database is also an effective way to improve the identification capability of the MALDI system.

The increased availability of faster and more accurate analytical methods facilitates a more in-depth understanding of microflora distribution and contamination sources in the pharmaceutical manufacturing environment. For this case, 44 genera and 92 species were classified in 2 months across various locations and grades, indicating relatively high risk and wide-ranging sources of contamination in the production process and environment. In line with previous studies, *Staphylococcus* spp. and *Micrococcus* spp. remained the dominant contaminants, reconfirming that the personnel could be the most significant contributor to cleanroom contamination. The distinct predominance of *Staphylococcus* spp., confirms our concern for CoNS contamination in the pharmaceutical industry (Song et al., 2019). At the same time, the presence of resistant bacterial spores present a considerable challenge to the CCS, particularly the validity of disinfection programs during the

pharmaceutical process, further highlighting the imperative for the periodic use of sporicide as regulations are expected.

Whole Genome Sequencing (WGS) is a strain-typing technique under development but is already being increasingly exploited in tracing transmission routes and identifying contamination events that occur in the production-to-patient (Balloux et al., 2018; Simar et al., 2021). With the more comprehensive properties, WGS analysis has the potential to be integrated into and strengthen microbiological risk assessment. In this research, we used whole genome SNP typing method to track the source of the most frequently contaminated species, 28 *S. cohnii* isolates distributed widely across grades A to C. The results demonstrate that the contamination of high-grade areas is most likely introduced by lower-grade air or personnel activities. This case also indicates that species-level identification is inadequate to give definitive tracing evidence under certain scenarios; more discriminatory typing methods at the strain-level are required to confirm or rule out the correlation.

In contrast to the conventional practice when a microbiological environmental monitoring excursion occurs, the unavailability of precise contamination risk information leaves only the option of non-targeted and extensive cleaning and disinfection procedures, which are costly, time-consuming, and compromise production efficiency. The bacterial contamination profiling scheme we established as illustrated in Figure 1, based on substantial species distribution and tracking information, will enable more accurate location of the contamination, integrate more evidence for risk assessment, facilitate the formulation and implementation of more targeted corrective and preventive measures, and provide a more scientific and substantiated basis for the development and refinement of contamination control strategies for the whole manufacturing process and associated environment.

The present study developed a multi-method solution mainly based on MALDI-TOF MS and complemented by sequencing to monitor and track pharmaceutical bacterial contamination. Due to the limitations of the MALDI database and the challenges of sample collection and cultivation, we did not address fungi, especially mold detection in this study. As illustrated in Figure 1, a dynamic, life-cycle approach will be our next research blueprint: leveraging accurate sequencing results for constant complementation of the MALDI-TOF MS database, gradually enhancing MALDI's identifying capability by creating a custom database, establishing an effective way to reduce the microbial burden and improve the level of aseptic assurance.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

MHS: Writing – original draft, Writing – review & editing, Funding acquisition, Methodology, Project administration. QQL:

Data curation, Methodology, Writing – review & editing. CZL: Data curation, Methodology, Writing – review & editing, Software. PEW: Data curation, Methodology, Writing – review & editing. FQ: Writing – review & editing. LCZ: Writing – review & editing, Investigation. YLF: Writing – review & editing, Funding acquisition. HS: Writing – review & editing. GLC: Writing – review & editing. MCY: Writing – review & editing, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from the Science and Technology Commission of Shanghai Municipality (22142201600), the Standard improvement project of the Chinese Pharmacopoeia Commission (2022Y21 and 2023Y36), and the Major Science and Technology Medicine and Healthcare in Zhejiang (WKJ-ZJ-2309).

Acknowledgments

The authors thank Dr. Kun Zhu from East China University of Science and Technology for her technical support and advice on the manuscript.

Conflict of interest

CZL was employed by Hangzhou Digital-Micro Biotech Co., Ltd. LCZ was employed by Shanghai SPH New Asia Pharmaceutical Co., Ltd. YLF, GLC, and MCY were employed by China State Institute of Pharmaceutical Industry, Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2024.1327175/full#supplementary-material>

References

- Akers, J. E. (1997). Environmental monitoring and control: proposed standards, current practices, and future directions. *PDA J. Pharm. Sci. Technol.* 51, 36–47.
- Andrade, L. D. O., Awasthi, R., Dua, K., and de Jesus Andreoli Pinto, T. (2018). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of bacteria isolated from pharmaceutical clean rooms. *Interv. Med. Appl. Sci.* 10, 45–53. doi: 10.1556/1646.9.2017.40
- Balloux, F., Brynildsrud, O. B., Van Dorp, L., Shaw, L. P., Chen, H., Harris, K. A., et al. (2018). From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends Microbiol.* 26, 1035–1048. doi: 10.1016/j.tim.2018.08.004
- Bizzini, A., and Greub, G. (2010). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clin. Microbiol. Infect.* 16, 1614–1619. doi: 10.1111/j.1469-0691.2010.03311.x
- Chinese Pharmacopoeia Commission (2020). *Pharmacopoeia of the People's Republic of China*. Beijing: The Medicine Science and Technology Press of China.
- Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2014). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 3:e15. doi: 10.1093/nar/gku1196
- Croxatto, A., Prod'homme, G., and Greub, G. (2012). Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol. Rev.* 36, 380–407. doi: 10.1111/j.1574-6976.2011.00298.x
- Eissa, M. E. (2016). Distribution of bacterial contamination in non-sterile pharmaceutical materials and assessment of its risk to the health of the final consumers quantitatively. *Beni-Suef Univ. J. Basic Appl. Sci.* 5, 217–230. doi: 10.1016/j.bjbas.2016.08.005
- Ferone, M., Gowen, A., Fanning, S., and Scannell, A. G. M. (2020). Microbial detection and identification methods: bench top assays to omics approaches. *Compr. Rev. Food Sci. Food Saf.* 19, 3106–3129. doi: 10.1111/1541-4337.12618
- Halls, N. (Ed.). (2004). *Microbiological contamination control in pharmaceutical clean rooms* (1st ed.). Boca Raton: CRC Press. doi: 10.1201/9781420025804
- Hamdy, A. M., El-Massry, M., Kashef, M. T., Amin, M. A., and Aziz, R. K. (2018). Toward the drug factory microbiome: microbial community variations in antibiotic-producing clean rooms. *Omic* 22, 133–144. doi: 10.1089/omi.2017.0091
- Hoefler, A., Boyen, F., Beierschmitt, A., Moodley, A., Roberts, M. C., and Butaye, P. (2021). Methicillin-resistant and methicillin-susceptible *Staphylococcus* from Vervet monkeys (*Chlorocebus sabaeus*) in Saint Kitts. *Antibiotics* 10:290. doi: 10.3390/antibiotics10030290
- Hunter, P. R., and Gaston, M. A. (1988). Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J. Clin. Microbiol.* 26, 2465–2466. doi: 10.1128/jcm.26.11.2465-2466.1988
- Jain, S. K., and Jain, R. K. (2018). Review of FDA warning letters to pharmaceuticals: cause and effect analysis. *Res. J. Pharm. Technol.* 11, 3219–3226. doi: 10.5958/0974-360X.2018.00592.9
- Jimenez, L. (2007). Microbial diversity in pharmaceutical product recalls and environments. *J. Am. Pharm. Rev.* 61, 383–399.
- Jimenez, L. (2019). Analysis of FDA enforcement reports (2012–2019) to determine the microbial diversity in contaminated non-sterile and sterile drugs. *J. Am. Pharm. Rev.* 4, 1–21.
- Kim, M., Oh, H. S., Park, S. C., and Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 64, 346–351. doi: 10.1099/ijs.0.059774-0
- McDonald, M., Dougall, A., Holt, D., Huygens, F., Oppedisano, F., Giffard, P. M., et al. (2006). Use of a single-nucleotide polymorphism genotyping system to demonstrate the unique epidemiology of methicillin-resistant *Staphylococcus aureus* in remote aboriginal communities. *J. Clin. Microbiol.* 44, 3720–3727. doi: 10.1128/JCM.00836-06
- Menghwar, H., Guo, A., Chen, Y., Lysnyansky, I., Parker, A. M., Prysliak, T., et al. (2022). A core genome multilocus sequence typing (cgMLST) analysis of *Mycoplasma bovis* isolates. *Vet. Microbiol.* 273:109532. doi: 10.1016/j.vetmic.2022.109532
- Miro, E., Rossen, J. W. A., Chlebowicz, M. A., Harmsen, D., Brisse, S., Passet, V., et al. (2020). Core/whole genome multilocus sequence typing and core genome SNP-based typing of OXA-48-producing *Klebsiella pneumoniae* clinical isolates from Spain. *Front. Microbiol.* 10:2961. doi: 10.3389/fmicb.2019.02961
- Park, H., Han, J. H., Joung, Y., Cho, S. H., Kim, S. A., and Kim, S. J. (2014). Bacterial diversity in the indoor air of pharmaceutical environment. *J. Appl. Microbiol.* 116, 718–727. doi: 10.1111/jam.12416
- Patel, R. (2015). MALDI-TOF MS for the diagnosis of infectious diseases. *Clin. Chem.* 61, 100–111. doi: 10.1373/clinchem.2014.221770
- Roe, C. C., Horn, K. S., Driebe, E. M., Bowers, J., Terriquez, J. A., Keim, P., et al. (2016). Whole genome SNP typing to investigate methicillin-resistant *Staphylococcus aureus* carriage in a health-care provider as the source of multiple surgical site infections. *Hereditas* 153, 1–7. doi: 10.1186/s41065-016-0017-x
- Sandle, T. (2011). A review of cleanroom microflora: types, trends, and patterns. *PDA J. Pharm. Sci. Technol.* 65, 392–403. doi: 10.5731/pdajpst.2011.00765
- Sandle, T. (2012). Review of FDA warning letters for microbial bioburden issues (2001–2011). *Pharma Times* 44, 29–30.
- Santos, A. M. C., Doria, M. S., Meirinhos-Soares, L., Almeida, A. J., and Menezes, J. C. (2018). A QRM discussion of microbial contamination of non-sterile drug products, using FDA and EMA warning letters recorded between 2008 and 2016. *PDA J. Pharm. Sci. Technol.* 72, 62–72. doi: 10.5731/pdajpst.2016.007252
- Schubert, S., and Kostrzewa, M. (2017). MALDI-TOF MS in the microbiology laboratory: current trends. *Curr. Issues Mol. Biol.* 23, 17–20. doi: 10.21775/cimb.023.017
- Scott, T. M., Rose, J. B., Jenkins, T. M., Farrah, S. R., and Lukasik, J. (2002). Microbial source tracking: current methodology and future directions. *Appl. Environ. Microbiol.* 68, 5796–5803. doi: 10.1128/AEM.68.12.5796-5803.2002
- Seuylemezian, A., Aronson, H. S., Tan, J., Lin, M., Schubert, W., and Vaishampayan, P. (2018). Development of a custom MALDI-TOF MS database for species-level identification of bacterial isolates collected from spacecraft and associated surfaces. *Front. Microbiol.* 9:780. doi: 10.3389/fmicb.2018.00780
- Simar, S. R., Hanson, B. M., and Arias, C. A. (2021). Techniques in bacterial strain typing: past, present, and future. *Curr. Opin. Infect. Dis.* 34, 339–345. doi: 10.1097/QCO.0000000000000743
- Singhal, N., Kumar, M., Kanaujia, P. K., and Virdi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front. Microbiol.* 6:791. doi: 10.3389/fmicb.2015.00791
- Song, M., Li, Q., He, Y., Lan, L., Feng, Z., Fan, Y., et al. (2019). A comprehensive multilocus sequence typing scheme for identification and genotyping of *Staphylococcus* strains. *Foodborne Pathog. Dis.* 16, 331–338. doi: 10.1089/fpd.2018.2565
- Uelze, L., Grütze, J., Borowiak, M., Hammerl, J. A., Juraschek, K., Deneke, C., et al. (2020). Typing methods based on whole genome sequencing data. *One Health Outlook* 2, 1–19. doi: 10.1186/s42522-020-0010-1
- Vijayakumar, R., Saleh Al-Aboody, M., and Sandle, T. (2016). A review of melanized (black) fungal contamination in pharmaceutical products—incidence, drug recall and control measures. *J. Appl. Microbiol.* 120, 831–841. doi: 10.1111/jam.12888
- Vijayakumar, R., Sandle, T., Manoharan, C., and Sciences, P. (2012). A review on fungal contamination in pharmaceutical products and phenotypic identification of contaminants by conventional. *Eur. J. Parenter. Pharm. Sci.* 17, 4–18.
- Wu, G.-F., and Liu, X.-H. (2007). Characterization of predominant bacteria isolates from clean rooms in a pharmaceutical production unit. *J. Zhejiang Univ. Sci. B* 8, 666–672. doi: 10.1631/jzus.2007.B0666