
Molekylär diagnostik av epidemier

JARI JALAVA OCH LAURA LINDHOLM

Målet för epidemiologisk utredning av infektionssjukdomar är att få en lägesbild av den infektiösa strändande mikrobens utbredning, smittvägar och av exponerade patienter för att bekämpningsåtgärderna ska kunna riktas in rätt. Traditionell epidemiologisk analys är i nyckelställning, men molekylär diagnostik finns nu för tiden ofta med i utredningen av en epidemi. Molekylär diagnostiken har utvecklats mycket snabbt, och till exempel metoder som grundar sig på sekvensering av hela genomet håller på att ersätta traditionella metoder för typning av mikrober. Fördelen med helgenomsekvensering jämfört med traditionella typningsmetoder är att den ger större mängd data och är bättre reproducerbar. Att få fram sekvensdata är numera inte längre ett problem. Däremot finns det fortfarande utmaningar med att analysera de data som produceras och att förstå resultaten.

SKRIBENTERNA

Jari Jalava, FD, är docent och ledande expert vid Institutet för hälsa och välfärd, avdelningen för hälsosäkerhet och enheten för smittskydd och vaccinationer. Hans ansvarsområde är resistensövervakning.

Laura Lindholm, FM, är specialforskare vid Institutet för hälsa och välfärd, avdelningen för hälsosäkerhet, enheten för mikrobiologisk expertis. Hennes ansvarsområde är typbestämning av MRSA- och VRE-stammen.

Utredning av epidemier är en del av infektionsbekämpningen. Målet är att få en så noggrann bild som möjligt av epidemins förlopp och utbredning för att bekämpningsåtgärderna ska kunna inriktas rätt och kostnadseffektivt. Vid sidan av traditionell epidemiologisk analys är typning av mikrober som hör till samma art ett viktigt verktyg vid epidemiutredning, det vill säga att särskilja olika stammar, eftersom patienterna på samma gång har både stammar som orsakat epidemin och stammar som inte hör till den. Genom att särskilja stammarna kan man ofta avgränsa bilden av epidemins omfattning, vilket minskar bekämpningskostnaderna.

För att typa mikrober, särskilt bakterier, har man traditionellt använt metoder som

grundar sig på fenotypen, såsom serotyper, biotyper eller fagtyper. Under de senaste åren har molekylär diagnostiska metoder, som alltså grundar sig på att mångfaldiga nukleinsyror (DNA och RNA) och bestämma deras bassekvens, utvecklats mycket snabbt. Molekylärmetoderna har i själva verket ersatt de traditionella typningsmetoderna.

Molekylär diagnostik används vid sidan av epidemiutredningar också för att producera epidemiologisk uppföljningsinformation.

I det följande koncentrerar vi oss på typning av bakterier med molekylära metoder. Förutom epidemiutredning går vi också in på hur uppföljningsinformation tas fram. Tyngdpunkten ligger på nya metoder baserade på helgenomsekvensering.

Uppföljning av infektionssjukdomar och betydelsen av typning

Många bakterier som ger upphov till epidemier är vanliga patogener på vårdinrättningar och inom öppenvården. Till exempel hittas meticillinresistent *Staphylococcus aureus* (MRSA) överallt i Finland på vårdinrättningar, men också i samband med öppenvårdsinfektioner. Att MRSA upptäcks innebär alltså ännu inte en epidemi. Om man ser att antalet fall ökar jämfört med normalnivån kan man däremot misstänka en epidemi. För att veta vilken den så kallade normalnivån är måste vi ha uppföljning av infektionssjukdomar och

patogener. En epidemi kan upptäckas om vi har en uppfattning om den sedvanliga nivån av bakteriefynd i kliniska prover. Det första tecknet på en epidemi eller ett kluster är vanligen att den sedvanliga nivån överskrids. Den sedvanliga nivån kan variera mellan vårdinrättningar beroende exempelvis på vilken typ av patienter som vårdas där. Epidemier kan lättare upptäckas om vi vet vilka bakterietyper som normalt förekommer, till exempel vilka MRSA-typer vi har och vilka deras relativa andelar är. En ovanlig MRSA-typ som upptäcks på samma avdelning under en kort tidsrymd är ett starkare tecken på en epidemi än en vanlig MRSA-typ som upptäcks hos flera patienter. Upptäckten av epidemier underlättas alltså av att vi vet vilka bakterietyper som förekommer under normala förhållanden. Därför görs fortlöpande uppföljning av de viktigaste mikroberna också genom att typa dem. Efter som typningsmetoderna är relativt dyra följer man inte upp ens tillnärmelsevis alla mikrober. I Finland upprätthåller Institutet för hälsa och välfärd (THL) Registret för smittsamma sjukdomar, där man samlar uppföljningsdata om olika infektionssjukdomar och patogener samt typningsdata om dem.

Epidemiutredningar

Målet för epidemiologisk utredning av infektionssjukdomar är att få en lägesbild av den infektionssjukdomens utbredning, smittvägar och exponerade patienter. Traditionell epidemiologisk analys har en nyckelställning i epidemiutredningarna och typningsdata för bakterier behövs som stöd för analysen. Vanligen säkerställer man med typning att det verkligen är fråga om en epidemi, det vill säga att en viss bakterietyp förekommer mer än normalt. Alla upptäckta bakterier har alltså inte nödvändigtvis samband med epidemin, trots att de med avseende på tid och plats kan passa in. Detta gäller särskilt relativt vanliga bakterier som MRSA. Att typa bakteriestammarna är till hjälp för att skilja bakteriestammar som hör till epidemin från sådana som inte gör det. Detta har betydelse förutom för att säkerställa epidemin också för att identifiera patienter som hör till epidemin och för att klarlägga smittvägar.

Regelbunden uppföljning av infektionssjukdomar och typning av bakteriestammar gör det ibland möjligt att upptäcka epidemier som inte har upptäckts av de lokala aktörer som svarar för infektionsbekämpning. Det är då fråga om en epidemi som har framskridit

långsamt, till exempel under flera år, eller en epidemi som omfattar flera olika vårdinrättningar. Då går det inte att upptäcka epidemin enbart genom att undersöka bakteriens förekomst i tid och rum.

Metoder för att typa bakterier

Det finns flera olika typningsmetoder och de delas vanligen in i fenotypiska metoder eller molekylära metoder. De fenotypiska metoderna, såsom sero- och fagtypning, har varit viktiga verktyg för att förstå infektionssjukdomarnas epidemiologi. Serotypning är fortfarande en viktig metod som är i användning. Data från serotypning behövs till exempel för att följa upp effekten av pneumokockvaccin. Fagtypning har länge använts för att typa bland annat tarmbakterier, till exempel Salmonella. De fenotypiska metoderna har dock ersatts eller håller på att ersättas med molekylära metoder. Orsaken är vanligen att de molekylära metoderna har bättre upplösningsförmåga och är mer reproducerbara. Dessutom kan tillgången till de serum, antikroppar och fager som behövs för de fenotypiska metoderna vara ett problem. Vanliga och mycket använda molekylära typningsmetoder är metoder som grundar sig på analys av DNA-fragment, exempelvis pulsfältselektrofores (PFGE), och metoder med DNA-fragment som har mångfaldigats med PCR, exempelvis ribotypning, amplified fragment length polymorphism (AFLP) och repetitive-element PCR (rep-PCR). Dessutom är metoder som utgår från sekvensering av DNA, exempelvis spa-typning (sekvensering av genen för protein A i *Staphylococcus aureus*) och MLST (Multi Locus Sequence Typing) fortfarande viktiga typningsmetoder, särskilt för att producera molekylär uppföljningsinformation (1). Metoder med den nya generationens sekvenseringstekniker (next generation sequencing, NGS) håller snabbt på att anammas av den mikrobiologiska diagnostiken och bli en del av uppföljningen av infektionssjukdomar. Det finns flera olika NGS-tekniker och utvecklingen på detta område har varit mycket snabb (2). NGS är inte i sig en typningsmetod utan en teknik för att snabbt och effektivt sekvensera vilket DNA som helst. Den används mycket för helgenomsekvensering av bakterier (Whole Genome Sequencing, WGS), som är den bästa av de typningsmetoder som för närvarande finns att tillgå (3).

Helgenomsekvensering

Med helgenomsekvensering (WGS) kan man bestämma bassekvensen i en organisms hela arvsmassa vid en enda sekvenseringssession. Trots att sekvensering av DNA i och med NGS-metoderna har blivit betydligt effektivare, är det fortfarande svårt att bestämma bassekvensen i en bakteries hela genom, det vill säga kromosomerna och eventuella plasmider, under en sekvenseringssession. För närvarande nöjer man sig därför oftast med en partiell genomsekvens som omfattar cirka 95 procent av genomet. Denna genomsekvens kallas draft genome. Det är dock en betydande förbättring jämfört med den tidigare situationen. För tio år sedan tog sekvenseringen av ett bakteriegenom flera månader och kostade tiotusentals euro. Med moderna NGS-apparater kan man sekvensera 20 MRSA-stammar med en sekvenseringskörning som tar 23 timmar. Att sekvensera en stam kostar ett tiotal eller högst några hundra euro beroende på kvalitet och metod. I praktiken kan WGS användas vid utredning av epidemier och för vissa bakterier också för att producera molekylär uppföljningsinformation. Det är svårt att tänka sig att utvecklingen av tekniken skulle stanna upp. Utvecklingen kan bli långsammare, men sannolikt kommer NGS-teknikerna att ytterligare utvecklas. Det i sin tur innebär att kostnaderna minskar och kvaliteten blir bättre. Redan nu används så kallade tredje generationens sekvenseringsapparater och tekniker som gör det möjligt att producera slutna bakteriegenom, som alltså är sekvenserade till 100 procent. Dessa tekniker är dock ännu inte i omfattande användning vid epidemiutredningar, utan utredningarna görs i huvudsak med draft genome-tekniker.

Helgenomsekvensering vid epidemiutredningar

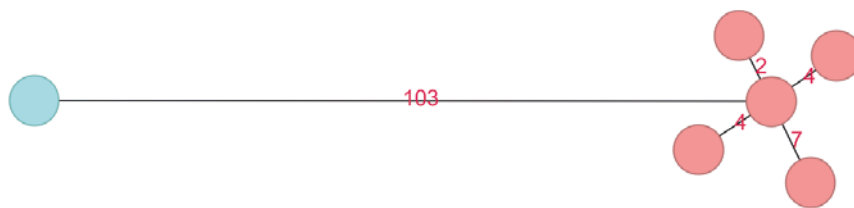
Fördelen med WGS jämfört med traditionella typningsmetoder är att den ger en större mängd data och är bättre reproducerbar. Eftersom WGS ger tillgång till nästan hela den genetiska informationen i bakterien är dess upplösningsförmåga klart bättre än med någon annan metod. I princip kan det inte finnas någon typningsmetod med bättre upplösning än WGS, eftersom all eller nästan all genominformation redan används. God upplösningsförmåga har betydelse särskilt vid epidemiutredningar. Vi kan använda skillnader var som helst i genomet för att typa

bakteriestammar. WGS används också i synnerhet för släktskapsanalyser. Med metoden kan vi beräkna vilka bakteriestammar som är nära släkt med varandra och vilka som har särutvecklats. Den informationen används när man försöker sluta sig till om bakteriestammar hör till samma epidemi eller inte. WGS ger så mycket sekvensdata att sådana släktskapsanalyser inte låter sig göras utan utvecklade tolkningsalgoritmer och program. Core genome (kärn genom) MLST (cgMLST) är en av de metoder som har utvecklats för släktskapsanalyser. Där undersöks alla gener som hittas i alla bakteriestammar som hör till epidemin och de jämförs sinsemellan. En jämförelse kan omfatta tusentals gener, till exempel för MRSA cirka 2 500 stycken. Om två stammar uppvisar skillnader i färre än tio gener anses de vara så nära släkt med varandra att de hör till den lokala epidemin. Om skillnader finns i fler än tjugo gener hör stammarna inte längre till samma lokala epidemi. Däremellan finns en gräzon där vi inte med säkerhet kan skilja epidemistammar från stammar som inte hör till epidemin. Tolkningsreglerna varierar för olika bakterier och de kan också förändras när kunskapen ökar. WGS har redan visat att den har en upplösningsförmåga som är synnerligen bra jämfört med alla andra typningsmetoder. Dessutom är reproducerbarheten utmärkt jämfört med till exempel analys av DNA-fragment med gelelektrofores med metoden PFGE. Reproducerbarheten vid WGS beror på att metoden i praktiken grundar sig på jämförelse av teckensträngar, vilket åtminstone i princip är mycket exakt.

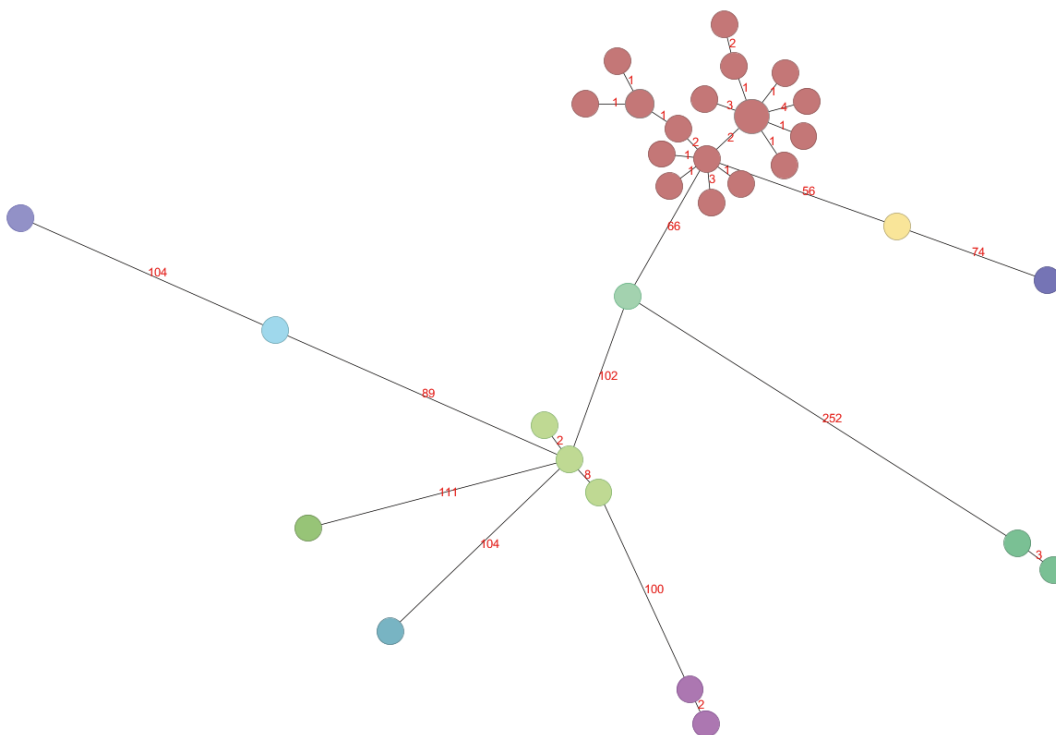
Figur 1 visar ett exempel på användning av cgMLST-analys vid utredning av ett kluster på en sjukhusavdelning. I exempelfallet hittades MRSA först i ett infektionsprov, varpå man tog screeningprover på övriga exponerade patienter på avdelningen. Av dessa prover var fyra positiva. Alla fem MRSA-isolat hade samma spa-typ och MLST-typ. Vid WGS-baserad cgMLST-analys användes data från sammanlagt 2 216 gener, varvid skillnader mellan infektionsprovet och screeningproverna påvisades i 2–7 gener; det var alltså klart och tydligt en klonal utbredning av samma stam på avdelningen. På samma avdelning hade det tre månader tidigare upptäckts ett enstaka MRSA-isolat med samma spa- och MLST-typ, som alltså med traditionella typningsmetoder skulle höra till samma kluster. En helgenomsekvensering gjordes på isolatet och den lades till samma cgMLST-analys. Isolatet uppvisade skillnader från klustrets stammar i 103 gener

Isolat som inte hör till
klustret (t223, ST22)

Isolat som hör till
klustret (t223, ST22)



Figur 1. WGS-baserad cgMLST-släktskapsanalys av isolat från ett MRSA-kluster på en avdelning (röda bollar). Den blå bubblan representerar ett enskilt isolat lite tidigare på samma avdelning, som hade samma spa-typ (t223) och sekvenstyp (ST22) som klusterstammen. Siffrorna vid strecken som förenar isolaten visar skillnaderna mellan isolaten (antalet gener med skillnader). Det skulle inte gå att skilja isolaten från varandra med traditionell spa- och MLST-typning.



Figur 2. WGS-baserad cgMLST-släktskapsanalys av en epidemi med utbredning i ett sjukvårdsdistrikt. Analysen avslöjade att det inom sjukvårdsdistriktet fanns ett stort kluster (bruna bubblor) och tre mindre kluster. Utöver dessa var en del av isolaten helt fristående och hörde inte till epidemin. Siffrorna vid strecken som förenar isolaten visar skillnaderna mellan isolaten (antalet gener med skillnader). Det skulle inte gå att skilja isolaten från varandra med traditionell spa- och MLST-typning eftersom alla isolat i analysen hade samma spa-typ och MLST-typ.

och det hörde alltså inte till samma kluster. Analysen var till betydande ekonomisk nytta: före helgenomsekvenseringsepoken hade isolatet ansetts höra till samma kluster och man hade inlett omfattande screening av patienter som hade vårdats på avdelningen under de tre månaderna mellan det enstaka fallet och klustret.

Figur 2 visar ett exempel på en omfattande epidemiutredning inom ett sjukvårdsdistrikt,

där en cgMLST-analys visade sin överlägsna upplösningsförmåga jämfört med traditionella metoder. När THL fick en begäran om helgenomsekvensering hade MRSA-stammar av samma spa- och MLST-typ hittats under ett drygt år, främst på två av sjukvårdsdistriktets sjukhus hos fler än tjugo patienter, och nya fall upptäcktes fortlöpande. Uppföljningen av epidemin var besvärlig och det fanns felande länkar i smittkedjan. Var det fråga om ett

större eller eventuellt om flera kluster? Alla de isolat som hade samma spa-typ som vid epidemin inom sjukvårdsdistriktet sekvenserades. Den WGS-baserade cgMLST-analysen visade att det inom distriktet fanns ett stort kluster och tre små kluster som inte hängde ihop sinsemellan. Alla sekvenserade isolat hade alltså samma spa-typ och MLST-typ; dessa traditionella metoder visade alltså att isolaten inte alls skiljde sig från varandra. Den noggranna differentiering av stammarna som möjliggjordes med cgMLST-analysen preciserade och avgränsade bilden av epidemin betydligt.

Helgenomsekvensering för att ta fram uppföljningsdata

Under flera decennier har vi samlat in rikligt med data som har tagits fram med fenotypiska och molekylära typningsmetoder och vi kan tolka resultaten av dessa data. Man måste kunna jämföra data från WGS med existerande data och man måste kunna tolka WGS-producerade data rätt. Problemet är för närvarande inte att producera sekvenser utan att förstå den erhållna informationen. Framsteg har dock gjorts inom denna sektor och vi kan redan nu dra många slutsatser av draft genome-sekvenser. Det lyckas nästan alltid att kalkylera sekvenstyp (MLST) och för MRSA:s del spa-typ ur draft genomen. Däremot är det betydligt svårare att beräkna serotyp, men också inom denna sektor har det skett framsteg. Pneumokockers och salmonellabakteriers serotyper kan redan nu beräknas ur draft-genomsekvenser i mer än 95 procent av fallen (5). Däremot går det inte att beräkna pulsältstyper förrän vi har tillgång till hundraprocentiga genomsekvenser.

Övriga WGS-tillämpningar

WGS-data kan utnyttjas också till annat än bara för att jämföra bakteriestammar sinsemellan eller för att få epidemiologisk typningsinformation för uppföljning av infektionsjukdomar. En viktig sektor är att bestämma antibiotikakänsligheten. I princip bestäms alla fenotypiska egenskaper hos en bakteriestam av stammens genom. Det är således möjligt att sluta sig till resistensprofilen för antibiotika enbart utgående från genomsekvensen. *S. aureus* är ett exempel på en bakterie vars resistensprofil relativt väl kan fastställas enbart ur genomsekvensen. Det går i själva verket bättre att bestämma resistensen ur genomsekvensen

för *S. aureus* än med traditionella metoder för känslighetsbestämning. WGS är dock än så länge en för dyr metod. För många bakteriers del har vi dessutom inte tillräckliga kunskaper om de gener eller mutationer som orsakar resistens för att enbart på basis av WGS kunna bestämma bakteriens resistensegenskaper tillräckligt omfattande och exakt. Epidemiologisk uppföljning av resistensgener kan dock göras uteslutande med WGS, och metoden håller också på att ersätta traditionella PCR-baserade metoder för att påvisa resistensgener. Dessutom kommer användningen av WGS att öka vid känslighetsbestämning för bakterier som är svåra att odla, bland annat *Mycobacterium tuberculosis*. WGS används sannolikt för att identifiera resistenta stammar av *M. tuberculosis*, och den egentliga resistensen säkerställs dessutom med fenotypisk känslighetsbestämning. I något skede kommer fenotypisk känslighetsbestämning av *M. tuberculosis*-stammar kanske att helt överges. Vad beträffar mykobakterier är det också viktigt att påvisa bakterien och dess resistens direkt i provet för att spara tid. Direkt påvisning av rifampicinresistens i sputumprov med PCR-metodik är för närvarande ett användbart sätt att snabbt påvisa multiresistenta *M. tuberculosis*-stammar, och metoden kan än så länge inte ersättas med WGS. Också inom denna sektor kommer WGS att spela en roll i framtiden.

Vid sidan av bestämning av läkemedelsresistens kan möjligheten att sluta sig till bakteriestammens virulens, det vill säga dess förmåga att orsaka sjukdom, vara en ny viktig tillämpning baserad på genomsekvensen. Man känner till en mängd gener som inverkar på virulensen och det är inte svårt att leta fram dem i gensekvensen. Problemet är att förstå betydelsen av olika slags virulensgener och kombinationer av dem. Det behövs mycket större kunskaper för att virulensgenomdata ska vara till praktisk nytta för bekämpningsåtgärder. Detsamma gäller gener som har samband med bakteriernas spridning. Vi vet av erfarenhet att en del bakteriestammar upprätthålls och sprids lätt till exempel på sjukhus. De är till sina egenskaper sådana att de har anpassat sig till sjukhusmiljön. Om vi skulle känna till de gener som inverkar på virulens och spridningsförmåga och deras olika kombinationer, kunde vi med hjälp av WGS dra slutsatser om varje bakteriestams egenskaper och utgående från det rikta in bekämpningsåtgärderna på de stammar som är anpassade till vårdinrättningarna och som är mer virulenta än normalt

eller som har båda egenskaperna. Än så länge är detta inte möjligt. Vi behöver ytterligare grundforskning om infektionssjukdomar för att bättre kunna förstå vilka faktorer som inverkar på virulensen och på spridningen av bakterierna.

Egenskaper som har samband med antibiotikaresistens och virulens finns ofta utanför bakteriernas kromosomer i plasmiderna. Nuvarande NGS-tekniker kan oftast inte ge sekvenser för hela plasmider och först tredje generationens tekniker erbjuder en lösning på detta problem. Det största problemet också inom denna sektor är dock att tolka resultaten och förstå deras betydelse.

Laboratorieteknikerna utvecklas synnerligen snabbt och molekylärdiagnostiken ersätter i snabb takt traditionella metoder. Trots det behövs mikrobstammar också i framtiden bland annat för att ge data om antibiotikakänslighet och för att bedöma mikrobstammarnas genetiska variation. Med molekylärdiagnostik hittar vi bara det som vi förstår att söka efter. WGS är en viktig del av molekylär epidemiutredning och den har ersatt många traditionella fenotypiska eller molekylära metoder för att typa bakterier. Att få fram sekvensdata är inte längre ett problem, utan problemet är att analysera framtagna data och framför allt att förstå resultaten. På vissa delområden såsom epidemiutredningar har man framskridit mycket bra med de viktiga släktskapsanalyserna. Det går bra att hitta de gener som orsakar antibiotikaresistens hos bakterier i genomsekvensen, men att tolka

resultaten är fortfarande en stor fråga. I framtiden kan vi använda WGS för att identifiera bakteriestammar som har utbredningsförmåga och är mer virulenta än normalt, vilket gör att bekämpningsåtgärderna kan riktas in mer effektivt.

Jari Jalava
jari.jalava@thl.fi

Laura Lindholm
laura.lindholm@thl.fi

Inga bindningar

Referenser

1. Sabat AJ, Budimir A, Nashev D, Sá-Leão R et al.; ESCMID Study Group of Epidemiological Markers (ESGEM). Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill.* 2013; 18:20380.
2. Margulies M, Egholm M, Altman WE, ym. Genome sequencing in microfabricated high-density picolitre reactors. *Nature.* 2005;437:376–380.
3. Holden MT, Hsu LY, Kurt K, ym. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Res.* 2013;23:653–664.
4. Bletz S, Mellmann A, Rothgänger J, Harmsen D. Ensuring backwards compatibility: traditional genotyping efforts in the era of whole genome sequencing. *Clin Microbiol Infect.* 2015;21:347.e1–4.
5. Kapatai G, Sheppard CL, Al-Shahib A, Litt DJ et al. Whole genome sequencing of *Streptococcus pneumoniae*: development, evaluation and verification of targets for serogroup and serotype prediction using an automated pipeline. *PeerJ.* 2016;4:e2477.
6. Aanensen DM, Feil EJ, Holden MT, Dordel J et al. European SRL Working Group. Whole-Genome Sequencing for Routine Pathogen Surveillance in Public Health: a Population Snapshot of Invasive *Staphylococcus aureus* in Europe. *MBio.* 2016;7:444–416.

Summary

Molecular diagnostics of epidemics

Epidemiological investigation of an outbreak is part of infection prevention and control. Typing methods for differentiating among various bacterial isolates are essential tools of modern outbreak investigation. Traditional methods such as serotype-, biotype-, or phage-typing have been in use for many decades. However, the development of molecular biology techniques is rapidly displacing traditional typing methods. Whole-genome sequencing is a new typing technique that has revolutionized bacterial typing. It can utilize almost all of the genetic information of the bacterial chromosome and can differentiate very efficiently among different bacterial isolates.