

### 細菌基因分型技術在食媒疾病分子流行病學上的應用

邱乾順、劉儼毅、廖盈淑\*

#### 摘要

細菌基因分型是現今公衛實驗室進行疾病監測與疾病爆發事件調查經常使用之方法。脈衝電泳(pulsed-field gel electrophoresis, PFGE)與多點可變異重覆序列分析法(multilocus variable number tandem repeat analysis, MLVA)因對大多數細菌種類具有高分型解析力，是最常用的細菌基因分型技術。細菌基因分型結果與疾病爆發事件的解讀有三種主要模式：H1，一個疾病群聚感染事件是由同一基因型菌株所引發；H2，一個疾病群聚感染事件是由多種基因型菌株或多個不同病原菌種所引發；H3，多個疾病群聚感染事件是由同一基因型菌株所引發。隨著次世代定序技術的進展，全基因體定序(whole genome sequencing, WGS)將成為公衛實驗室常規分析菌株的方法。菌株全基因體序列提供豐富的訊息，可用於擷取菌株血清型、致病因子、抗藥基因與可展示菌株經歷長短時距演化之基因圖譜資料。全基因體定序將在數年內取代脈衝電泳與多點可變異重覆序列分析法，成為公衛實驗室進行細菌菌株基因分型的最終方法。疾病管制署需儘速建立實驗室全基因體定序之量能與序列資料分析能力，充分準備以面對次世代定序技術的來臨。

**關鍵字：**基因分型、分子流行病學、脈衝電泳法、次世代定序技術、全基因體定序

#### 前言

細菌分型(typing)可供建構菌株間之遺傳關聯性(genetic relatedness)，提供流行病學研究調查之用。細菌分型可分表現型分型(phenotyping)與基因分型(genotyping)，分型方法的選擇視目的而定，例如為了疫苗研發與疫苗接種政策所需，進行

衛生福利部疾病管制署檢驗及疫苗研製中心

通訊作者：廖盈淑\*

E-mail : yingshu0439@cdc.gov.tw

投稿日期：2016年05月25日

接受日期：2016年07月05日

DOI : 10.6524/EB.20170124.33(2).001

*Streptococcus pneumoniae* 的多醣莢膜血清型流行趨勢調查；為了探討病原經歷長時間(timescale)的流行傳播，使用演化速率較慢的多點基因定序方法(multilocus sequence typing, MLST)分析菌株；為了疾病監測(disease surveillance)與疾病爆發事件調查(disease outbreak investigation)，使用高分型解析力(discriminatory power)的脈衝電泳(pulsed-field gel electrophoresis, PFGE)與多點可變異重覆序列分析(multilocus variable number tandem repeat analysis, MLVA)偵測菌株經歷短時距的變異。本文介紹現行食媒細菌病原主要的基因分型方法，舉例說明分型方法在細菌性食媒疾病監測與疾病爆發事件流病調查之應用，提出分型結果與疾病爆發事件的解讀模式與未來細菌分型技術的發展趨勢，供公衛實驗室與傳染病流行病學調查人員參考。

## 食媒細菌病原基因分型方法與應用

細菌基因分型方法眾多，其中 MLST、PFGE 與 MLVA 是目前公衛實驗室最常使用的分型方法，預估數年內全基因體定序(whole genome sequencing, WGS)會成為細菌最終基因分型技術。

### 一、多點基因定序方法(MLST)

英國 Wellcome Trust Center 的 Maiden 等人[1]於 1998 年提出分析 6 個 *Neisseria meningitidis* 的持家基因(house-keeping genes)(後來增為 7 個基因)，約 470 bp 片段序列，以鹼基序列差異建立菌株之親緣關係。由於研究者可透過網路進入 MLST 資料庫取得菌株基因序列，容易與各地之菌株比對親緣關係，因此 MLST 曾廣為學界使用，也成功分析數個菌種的族群結構(population structure)[2]。由於 DNA 鹼基的變異速度較慢[3]，MLST 適合用於探討菌株數年至數十年或更長時間的演化關係，監控特定病原菌 strains 或 clones 在國際間的傳播流行與消長。MLST 也曾被應用於人類學的研究；Moodley 等人曾分析各地原住民族胃部分離之幽門桿菌(*Helicobacter pylori*)基因序列，依據菌株 MLST 序列資料描繪兩波史前人類在太平洋遷移的歷程：第一波在 3 萬多年前遷移到達新幾內亞與澳洲，第二波遷移在 5,000 年前以臺灣為起點，向大太平洋的美拉尼西亞(Melanesia)和波利尼西亞(Polynesia)區域遷徙[4]。

### 二、脈衝電泳(PFGE)

PFGE 對許多菌種具高分型解析力，是目前細菌基因分型的黃金標準。PFGE 可用於評估菌株經歷短時距演化之親緣關係，適合做為疾病主動監測與疾病爆發事件進行流病調查時的菌株分型工具。美國疾病防治暨預防中心(US Centers for Disease Control and Prevention)使用以 PFGE 為標準的菌株基因分型工具，建立了一個全國性的食媒疾病主動監測網—PulseNet [5]，該監測系統後來被擴展為有 80 多個國家參與的全球性食媒疾病分子分型監測網—PulsNet International [6]。臺灣為 PulseNet Asia Pacific 區域監測網成員，疾病管制署(以下簡稱疾管署)亦於 2006 年成立 PulseNet Taiwan，採用 PulseNet

標準化的 PFGE 技術，並建立多種細菌病原的 DNA 指紋圖譜資料庫。PFGE 對許多細菌具有高分型解析力，但對一些單形性(monomorphic)菌種如 *Salmonella enterica* serovar Enteritidis (*S. Enteritidis*)、*S. sonnei*、*E. coli* O157:H7 等菌種之分型解析力時有侷限[7]，加上 PFGE 相當費時費力且每次只能分析少數菌株，操作上需訓練有素的技術員，只能在參考實驗室實施，這些缺點驅動研究人員發展其它分型技術，例如 MLVA 以取代 PFGE。

### 三、多點可變異重覆序列分析(MLVA)

細菌基因體存在數百到上千個由短 DNA 序列單元頭尾相接的重覆序列(tandem repeat, TR)位點，其中有非常少數 TR 之重覆單元(repeat unit)的數目會發生變化，稱為可變異重覆序列(variable number tandem repeat, VNTR)，分析多個 VNTR 的重覆單元數目，可得到菌株多個 VNTR 之重覆單元數目的排列組合，即為 MLVA 基因型別。VNTR 位點之重覆單元數目變化速率不同，少數幾個變異速率快的 VNTR 組合，其分型解析力即可高於 PFGE [8]。MLVA 適合用於檢測菌株經歷短時距的變化，而使用變異速度低的 VNTR，則可建立菌株經歷較長時距的演化關係[7]。疾管署曾和數個國家合作，利用自行研發的 MLVA 方法分析 50 個國家來源的 *S. sonnei* 菌株，分析結果指出自 1943 至 2000 年間，全世界流行之 *S. sonnei* 有 3 個主要的 clonal groups [9]；該結論一年後再次被 Holt 等人確認[10]，該研究分析 132 株世界各地 *S. sonnei* 分離株之全基體序列，結果指出 *S. sonnei* 是 500 年內才出現的新生病原菌，目前全世界有 3 個主要的係譜(lineages)。這 3 個 lineages 與使用 VNTR 建立之 3 個 clonal groups 可相互對應[11]。

MLVA 曾被認為具有取代 PFGE 做為 PulseNet 監測網標準分型工具的潛力，然而 MLVA 菌種專一性太高[12]，無法成為取代 PFGE 的選擇，而隨著次世代定序技術(next generation sequencing, NGS)的發展，WGS 將取代 PFGE 成為細菌之標準分型工具。

### 四、全基因體定序(WGS)

隨著次世代定序技術的進步，WGS 的成本已降低至可供例行性分析細菌株的實用階段。WGS 比其它分型方法提供更豐富訊息，可依基因演化速度選擇特定基因組合，用於探討不同演化時距的菌株親緣關係。理論上，在建立菌株經長時間演化後之親緣關係時，選用演化速率較慢的基因組合，當要呈現菌株經歷短時間演化後之親緣關係時，會選用演化速率較快的基因組合；因此探討菌株經長時距的變化時，會選用較少數量的核心基因(core genes)組合，當要探討菌株經歷較短時距的微演化(microevolution)時，需增加比對附屬基因(accessory genes)上的變異。現階段應用 WGS 技術最大的困難已不在定序的成本，而是在 DNA 序列的分析方法；一株菌株 WGS 分析會產生百萬條短序列，如何從這些巨量序列資料中擷取所要的基因資訊，是 WGS 方法在實際應用上最核心的關鍵。

## 基因分型資料的解讀

每種基因分型結果與疾病爆發事件流病關聯性的解讀標準，應依實證經驗而定。1995 年 Tenover 等人[13]曾提出解讀 PFGE 分型結果的準則，該準則用於解釋 PFGE 圖譜與流病的關係，認為 PFGE 圖譜差異等於或小於 3 條 DNA 片段之菌株，才可能源於同一疾病爆發事件，然而實證經驗指出該準則並非全然符合[14,15]。另外，在 MLVA 圖譜的應用上，學者根據研究調查的經驗，認為可接受同一爆發流行事件之菌株有一個 VNTR 的差異[16]，這個準則也非一體適用，至於 WGS 基因圖譜資料仍未有一個公認的解讀標準。

基因分型結果決定的是菌株間的親緣關係，在流行病學的應用是利用菌株的親緣關係來推想病例間的流病關聯性，因此解讀基因分型資料需考慮感染源之污染模式。基因分型結果與疾病爆發事件的關係有三種可能模式：

### 一、H1: $G_1 \rightarrow O_1$

一件疾病爆發事件是由同一基因型病原菌種所引起；即共同感染源為單一基因型病原菌株(strain)所污染。

H1 是食媒疾病主動監測（例如 PulseNet）與疾病爆發事件流病調查的假設基礎。PulseNet 食媒疾病分子分型監測網即設定在某個時間（例如 1 個月）內出現 2 株或以上具相同基因型菌株，即認定為一起群聚感染(cluster)，因而啟動流行病學調查追查病例是否有共同感染來源。在疾病爆發事件或院內群聚感染事件的調查上，也是假定具相同基因型分離株來自同一感染源。實務上此類模式之疾病爆發事件占多數，但也很有可能因為前提的設定而排除了有共同感染來源但不同基因型菌株感染之病例。

### 二、H2: $G_m \rightarrow O_1$

一件疾病爆發事件是由同一菌種的多種基因型菌株（或由多種病原菌種）所引起；即共同感染源為同一菌種的多重基因型菌株或多重菌種所污染。

H2 是具共同感染來源的疾病爆發事件，由多種基因型別菌株或多種病原菌種所引發。這種情況在傳染病爆發事件中並不罕見，但許多人包括從事細菌分型研究的學者其觀念仍停留在 H1 模式。國內曾發生的 3 起傷寒感染事件[14]可說明 H2 模式，此 3 個案例指出傷寒長期帶菌者在同一時間可排出不同基因型別的菌株，這 3 個實例改變解讀基因分型結果的傳統觀念(H1 模式)。

#### (一) 案例一：2005 年桃園傷寒群聚

在 2005 年，當時的桃園縣在一個月期間內出現 14 例傷寒病例，臺灣傷寒病例相當少，一個縣一個月內出現多達 14 例病例，屬相當不尋常，推測應該有共同感染來源。然而來自病例的 14 株菌株有 5 種不同 PFGE 圖譜，最大差距高達 8 條 DNA 片段，MLVA8 則有 4 種圖譜，在測試的 8 個 VNTR 中有 2 個 VNTR 點的差距，因此這個共同感染來源的推論受到同儕學者的質疑。

## (二) 案例二：2007 年新竹傷寒群聚

新竹地區從 2007 年開始，即有零星傷寒病例，菌株擁有相同 PFGE 圖譜（圖譜編號：SIX.022），但 MLVA8 圖譜則不盡相同，也找不到病例的共同流病關聯。2010 年新竹某高科技電子廠有 4 名員工出現傷寒群聚感染，菌株擁有相同的 PFGE 圖譜(SIX.022)與 MLVA8 圖譜（圖譜編號：TY8.040），此 PFGE 基因型和 2007 年以來出現在新竹地區幾個傷寒菌株相同，但仍未能找到感染來源。2011 年，該高科技電子廠又出現 4 例傷寒病例，菌株具 SIX.022 圖譜。疾管署經詳細疫調查出這些病例皆在下午茶時外叫印尼風味餐，該麵攤在數公里外由一原印尼籍歸化的 60 歲婦人經營；該婦人連續 2 日的糞便檢體分離出傷寒桿菌，2 件檢體各分離 1 株菌株，和電子廠員工病例的分離株之 PFGE 圖譜相同，但第一天檢體的菌株 MLVA8 圖譜(TY8.281)則有 3 個 VNTR 點的差異，第二天檢體的分離菌株 MLVA8 圖譜(TY8.040)和電子廠病例菌株又相同，據此一結果可推論病原在長期帶原者體內演化變異，帶原者可同時排出不同基因型別的菌株。可惜原糞便檢體已銷毀，無法再分離更多菌株以確認此一推論。

## (三) 案例三：2012 年新北市傷寒群聚

2012 年 4 月新北市出現 1 名 92 歲傷寒病例，其印尼籍看護為傷寒帶原者，而最初帶原者的 1 株分離株與病例分離株 PFGE 與 MLVA8 基因型不同，因此再由帶原者同一糞便檢體分離 100 個菌株，帶原者的 101 株分離株共有 5 種不同 PFGE 圖譜與 23 種 MLVA8 圖譜，其中只有 3 株分離株與病例分離株有相同 PFGE 與 MLVA8 圖譜(TY8.300)。PFGE 圖譜間最多有 8 條 DNA 差異，MLVA8 圖譜最多相差 4 個 VNTR 點。PFGE 結果打破了 Tenover 準則[13]，該準則認為 PFGE 圖譜差異等於或大於 7 條 DNA 片段之菌株，應非源於同一疾病爆發事件。另外，在菌株 MLVA 圖譜資料的解讀上，實證經驗指出同一爆發流行事件之菌株可有一個 VNTR 的差異[16]，但此長期帶原者卻可同時排出有具有 4 個 VNTR 點差異的菌株。長期帶原者可同時排出基因型差異大之菌株，這種情形也可能發生在桿菌性痢疾長期帶原者[15,17]。

在美國也常發生由不同基因型別病原菌株所引發的食媒疾病爆發流行事件，例如 2011 年在美國發生一起導致 33 人死亡的哈密瓜李斯特菌爆發流行事件，污染的哈密瓜、環境與病人檢體即檢出 2 種血清型、5 種 PFGE 基因型的李斯特菌[18]。國內食品中毒事件的調查經驗，也發生同一疾病爆發事件牽涉多種病原菌種的案例；例如 2010 年臺中市發生的一起三明治食品中毒事件[19]，就從病人與其食用污染的三明治分離出 2 種不同血清型的沙門氏菌，三明治亦同時分離出沙門氏菌之外的其它 4 種病原細菌。

### 三、H3: G<sub>1</sub> → O<sub>m</sub>

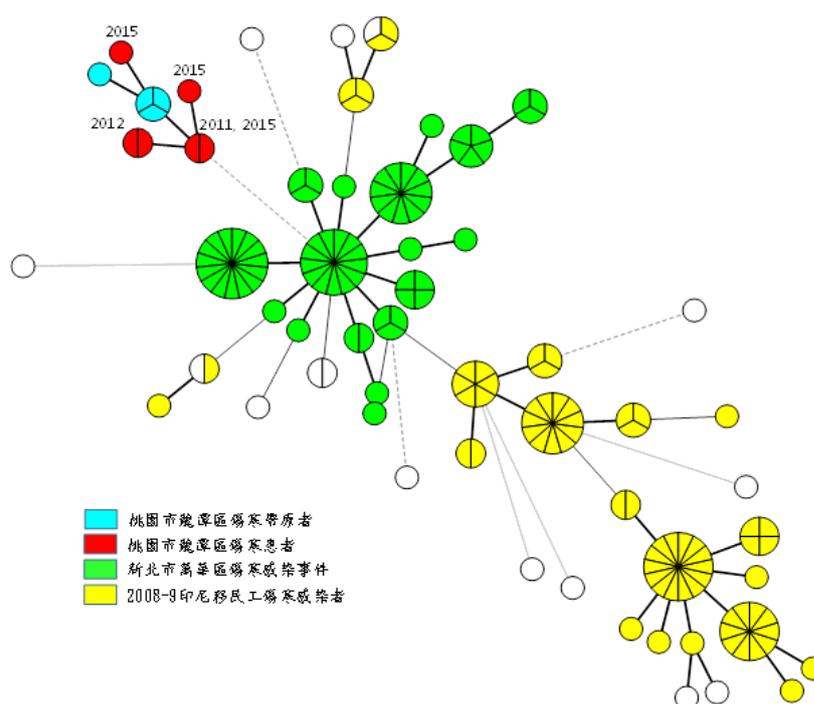
多起疾病爆發事件是由同一基因型病原菌株所引起；即單一個 strain 擴散出去引發流行，群聚感染會有許多不同感染源，但其菌株皆為同一基因型。

H3 可能是同一種基因型病原廣泛散播，引發多起不同感染來源的疾病爆發流行。對這類流行事件，菌株分型結果對追查感染來源的流病調查，難有實質貢獻。另外使用分型解析力不足的分型方法分析遺傳單形性的菌種也會出現 H3 的模式，其中 *S. Enteritidis* 即是一個明顯例子。*S. Enteritidis* 於 1980 年代初期在歐洲開始出現大規模流行；2015 年該血清型在臺灣引發近 40% 的 *Salmonella* 感染症病例。*S. Enteritidis* 菌株為遺傳單形性菌種，大多數菌株有相當接近的遺傳性狀，各地區的分離株呈現相當有限的 PFGE 與 MLVA 基因型別，因此許多具相同 PFGE 或 MLVA 基因型菌株其實來自不同感染來源，解決這個問題需使用更高分型解析力的 WGS 分型方法[20]。

### 細菌分型技術的發展趨勢

在公衛領域進行疾病的主動監測與疾病爆發事件的調查，需使用高分型解析力的分型方法，才能有效地偵測菌株經歷短時距的細微變化。例如 2015 年發生在桃園市龍潭區的傷寒群聚感染事件，病例、帶原者與之前 2011 與 2012 年的感染者分離株之 PFGE 圖譜皆是 SIX.001 型，在疾管署的資料庫中共有 161 株於 1998–2015 年分離之菌株是 SIX.001 基因型，其中 69 株來源為 2008–2009 年入境的印尼移民工，67 株來源為 2012 年新北市傷寒感染事件的患者與帶原者，這個事實降低 PFGE 分型結果對傷寒流病調查之參考價值。然而使用分型解析力較高的 MLVA 方法，則可明顯區別 2015 年龍潭傷寒群聚感染事件菌株與其它來源的 SIX.001 菌株（圖一）。

WGS 會是細菌最終的分型方法，美國疾病管制暨預防中心於 2015 年開始運用 WGS 分型方法進行 *Listeria monocytogenes* 的監測，發現 WGS 比 PFGE 可更明確辨認出群聚感染菌株，可偵測到更多與規模更小的群聚感染案件，因此該機構規劃在 2–3 年內全面使用 WGS 為分型工具，分析所有種類的食媒病原菌株。臺灣過去使用 PulseNet 的 PFGE 標準方法從事細菌株基因分型，除了進行國內疾病監測與流病調查，亦能和美國及其它使用 PulseNet 標準 PFGE 方法的國家進行菌株圖譜的比對，能取得國外食媒病原流行菌株情報。預估數年內 WGS 將成為疾病監測分析細菌株的標準分型工具，疾管署已於今(2016)年採購硬體設備與訓練操作人員，建立 WGS 的分析量能；然而在實務運作上，具有 WGS 序列資料分析能力的生物資訊人才是關鍵，公衛實驗室需重視生物資訊人才的招募與培養，以面對次世代定序時代的來臨。



圖一、傷寒菌株 MLVA8 基因型關係樹，關係樹係使用 **minimum spanning tree** 演算法所建立。菌株同屬 SIX.001 PFGE 基因型，主要來自 2008–2009 年入境之印尼移民工，新北市萬華區傷寒事件之病例與帶原者，2015 年桃園市傷寒群聚感染事件之病例與帶原者。MLVA8 基因型相差 1 個 VNTR 點者以短粗黑線連接，相差 2 個 VNTR 點者以長細黑線連接，相差 3 個或以上 VNTR 點者以長細虛線相連。

### 參考文獻

1. Maiden MC, Bygraves JA, Feil E, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* 1998; 95: 3140–5.
2. Feil EJ, Li BC, Aanensen DM, et al. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* 2004; 186: 1518–30.
3. Drake JW, Charlesworth B, Charlesworth D, et al. Rates of spontaneous mutation. *Genetics* 1998; 148: 1667–86.
4. Moodley Y, Linz B, Yamaoka Y, et al. The peopling of the Pacific from a bacterial perspective. *Science* 2009; 323: 527–30.
5. Swaminathan B, Barrett TJ, Hunter SB, et al. PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerg Infect Dis* 2001; 7: 382–9.

6. Swaminathan B, Gerner-Smidt P, Ng LK, et al. Building PulseNet International: an interconnected system of laboratory networks to facilitate timely public health recognition and response to foodborne disease outbreaks and emerging foodborne diseases. *Foodborne Pathog Dis* 2006; 3: 36–50.
7. Chiou CS, Watanabe H, Wang YW, et al. Utility of multilocus variable-number tandem-repeat analysis as a molecular tool for phylogenetic analysis of *Shigella sonnei*. *J Clin Microbiol* 2009; 47: 1149–54.
8. Chiou CS, Hung CS, Torpdahl M, et al. Development and evaluation of multilocus variable number tandem repeat analysis for fine typing and phylogenetic analysis of *Salmonella enterica* serovar Typhimurium. *Int J Food Microbiol* 2010; 142: 67–73.
9. Filliol-Toutain I, Chiou CS, Mammina C, et al. Global distribution of *Shigella sonnei* clones. *Emerg Infect Dis* 2011; 17: 1910–2.
10. Holt KE, Baker S, Weill FX, et al. *Shigella sonnei* genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. *Nat Genet* 2012; 44: 1056–9.
11. Sangal V, Holt KE, Yuan J, et al. Global phylogeny of *Shigella sonnei* strains from limited single nucleotide polymorphisms (SNPs) and development of a rapid and cost-effective SNP-typing scheme for strain identification by high-resolution melting analysis. *J Clin Microbiol* 2013; 51: 303–5.
12. Chiou CS. Multilocus variable-number tandem repeat analysis as a molecular tool for subtyping and phylogenetic analysis of bacterial pathogens. *Expert Rev Mol Diagn* 2010; 10: 5–7.
13. Tenover FC, Arbeit RD, Goering RV, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* 1995; 33: 2233–9.
14. Chiou CS, Wei HL, Mu JJ, et al. *Salmonella enterica* serovar Typhi variants in long-term carriers. *J Clin Microbiol* 2013; 51: 669–72.
15. 廖盈淑、王佑文、廖春杏等：苗栗縣 B 群 X 變異型桿菌性痢疾群突發事件。疫情報導 2009；25：717–25。
16. Liang SY, Watanabe H, Terajima J, et al. Multilocus variable-number tandem repeat analysis for molecular typing of *Shigella sonnei*. *J Clin Microbiol* 2007; 45: 3574–80.
17. 李佩璇、胡曉宇、許珮珊等：某精神專科醫院桿菌性痢疾群突發事件調查——外籍看護工之出入境省思。感染控制雜誌 2014；24：161–9。
18. McCollum JT, Cronquist AB, Silk BJ, et al. Multistate outbreak of listeriosis associated with cantaloupe. *N Engl J Med* 2013; 369: 944–53.

19. Wei SH, Huang AS, Liao YS, et al. A large outbreak of salmonellosis associated with sandwiches contaminated with multiple bacterial pathogens purchased via an online shopping service. *Foodborne Pathog Dis* 2014; 11: 230–3.
20. Taylor AJ, Lappi V, Wolfgang WJ, et al. Characterization of foodborne outbreaks of *Salmonella enterica* serovar Enteritidis with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection. *J Clin Microbiol* 2015; 53: 3334–40.