Direct molecular typing of *Bordetella pertussis* from clinical specimens submitted for diagnostic quantitative (real-time) PCR

David J. Litt,1 Elita Jauneikaite,1 Draga Tchipeva,1,2 Timothy G. Harrison1 and Norman K. Fry1

1Respiratory and Systemic Infection Laboratory, Health Protection Agency, 61 Colindale Avenue, London NW9 5EQ, UK
2Department of Biomedical Sciences, School of Life Sciences, University of Westminster, London W1W 6UW, UK

Molecular typing of *Bordetella pertussis* is routinely performed on bacterial isolates, but not on DNA extracted from nasopharyngeal aspirates or pernasal swabs submitted for diagnostic real-time PCR (qPCR). We investigated whether these DNA extracts were suitable for multilocus variable-number tandem repeat analysis (MLVA) and DNA sequence-based typing. We analysed all the available qPCR-positive samples received by our laboratory from patients <1 year of age between January 2008 and August 2010. Eighty-one per cent (106/131) of these generated a complete MLVA profile. This rose to 92 % (105/114) if only samples positive for both of the two targets used for the *B. pertussis* PCR (insertion element IS481 and pertussis toxin promoter ptxP) were analysed. Sequence-based typing of the pertactin, pertussis toxin S1 subunit and pertussis promoter regions (prn, ptxA and ptxP) was attempted on 89 of the DNA extracts that had generated a full MLVA profile. Eighty-three (93 %) of these produced complete sequences for all three targets. Comparison of molecular typing data from the 89 extracts with those from 111 contemporary bacterial isolates showed that the two sources yielded the same picture of the *B. pertussis* population [dominated by the MLVA-27 prn(2) ptxA(1) ptxP(3) clonal type]. There was no significant difference in MLVA type distribution or diversity between the two sample sets. This suggests that clinical extracts can be used in place of, or to complement, bacterial cultures for typing purposes (at least, in this age group). With small modifications to methodology, generating MLVA and sequence-based typing data from qPCR-positive clinical DNA extracts is likely to generate a complete dataset in the majority of samples from the <1 year age group. Its success with samples from older subjects remains to be seen. However, our data suggest that it is suitable for inclusion in molecular epidemiological studies of the *B. pertussis* population or as a tool in outbreak investigations.

INTRODUCTION

The bacterium *Bordetella pertussis* causes potentially fatal pertussis disease in humans. Despite vaccination programmes that are highly effective in reducing serious disease and mortality, there has been some concern in developed countries that pertussis disease is on the increase. Various explanations have been proposed, including increased ascertainment due to improved diagnostic methods, poor efficacy of particular batches of vaccine, and genetic variation occurring in the bacterial population (Campbell et al., 2012; Celentano et al., 2005; Mooi et al., 1999; Mooi, 2010; Ntezayabo et al., 2003; Poynten et al., 2004; Tanaka et al., 2003). Protein sequence variation has been observed in many *B. pertussis* virulence factors, including components of acellular pertussis vaccines such as pertussis toxin (Ptx), pertactin (Prn) and fimbriae (Fim) (Mooi et al., 2007; Mooi, 2010). Experiments using animal models have shown that divergent Ptx, Prn and Fim types can result in immune evasion (Robinson et al., 1989; King et al. 2001; Gzyl et al., 2004; Bottero et al., 2007; Komatsu et al., 2010) [although they also suggest that commercial acellular vaccines adequately protect against *B. pertussis* strains expressing divergent antigens (Boursaux-Eude et al., 1999; Denoël et al., 2005)].

**Abbreviations:** DI, diversity index; MLVA, multilocus variable-number tandem repeat analysis; MT, MLVA type; NPA, nasopharyngeal aspirate; PNS, pernasal swab; qPCR, quantitative (real-time) PCR; RSIL, Respiratory and Systemic Infection Laboratory; SNP, single nucleotide polymorphism.
A number of methods are used to monitor variation in circulating \textit{B. pertussis} isolates. Serotyping measures the expression of the fimbriae Fim2 and Fim3 (Mooi \textit{et al.}, 2000). Characterizing changes in genes encoding known vaccine antigens or virulence factors such as the S1 and S3 subunits of pertussis toxin (\textit{ptxA} and \textit{ptxC}), pertactin (\textit{prn}), tracheal colonization factor (\textit{tcfA}) and fimbrial antigens 2 and 3 (\textit{fim2} and \textit{fim3}) is also valuable (Mooi \textit{et al.}, 2007; Packard \textit{et al.}, 2004; Tsang \textit{et al.}, 2004; van Loo \textit{et al.}, 2002). Polymorphisms in the pertussis toxin promoter (\textit{ptxP}) have also been proposed to affect the virulence of \textit{B. pertussis} through increased Ptx production (Mooi \textit{et al.}, 2009; Mooi, 2010), although an increase in virulence was not seen in a recent study using a murine model of infection (Hegerle \textit{et al.}, 2012). Other methods designed to distinguish isolates, which are not linked to virulence factors, include PFGE, multilocus variable-number tandem repeat analysis (MLVA) and genome-wide single nucleotide polymorphism (SNP) analysis (Mooi \textit{et al.}, 2000; Octavia \textit{et al.}, 2011; Schouls \textit{et al.}, 2004; van Gent \textit{et al.}, 2011).

Bacterial typing methods are usually performed using clinical isolates that have been cultured from patients during pertussis disease. Acute \textit{B. pertussis} infections are traditionally diagnosed by isolating live bacteria from pernasal swabs (PNSs) or nasopharyngeal aspirates (NPAs) (Mattoo & Cherry, 2005). However, bacteria can also be detected in these samples using conventional PCR or real-time PCR (qPCR) (Riffelmehn \textit{et al.}, 2005; Mattoo & Cherry, 2005). The UK Health Protection Agency’s Respiratory and Systemic Infection Laboratory (RSIL) offers a diagnostic qPCR service, aimed primarily at hospitalized children under 1 year of age. RSIL performs a two-target qPCR on DNA extracted from PNSs or NPAs (Fry \textit{et al.}, 2009). The first target, within the pertussis toxin S1 promoter (\textit{ptxP}, previously denoted \textit{ptxA}-pr), is specific for \textit{B. pertussis} and is present at 1 copy per bacterial genome. The second target, the insertion sequence IS481, occurs in multiple copies in \textit{B. pertussis} strains, but may also be present in strains of \textit{Bordetella parapertussis}, \textit{Bordetella bronchiseptica} and \textit{Bordetella holmesii} (Muyldermans \textit{et al.}, 2005; Register & Sanden, 2006; Reischl \textit{et al.}, 2001; Fry \textit{et al.}, 2009). Despite the age restriction on the Health Protection Agency’s diagnostic qPCR service, the number of pertussis cases confirmed by qPCR is increasing each year. Meanwhile, the number confirmed by bacterial isolation has been falling (data from Health Protection Agency Health Protection Report; www.hpa.org.uk/HPR).

Using current methodology, a reduction in culturing would compromise our attempts to analyse the population of circulating clinical isolates. Hence, we decided to assess the feasibility of characterizing \textit{B. pertussis} strains infecting patients using DNA extracted from uncultured clinical samples that had been submitted to the laboratory for diagnostic qPCR. Furthermore, we tested the appropriateness of using these samples by comparing the data generated by clinical samples from patients <1 year of age with the results produced by isolates taken from a separate (but equivalent) group of patients. MLVA and DNA sequence-based typing were selected due to their applicability to samples containing low concentrations of target DNA.

**METHODS**

**DNA from clinical samples and bacterial strains.** A review of \textit{B. pertussis}-specific diagnostic qPCR results (method of Fry \textit{et al.}, 2009) generated between January 2008 and August 2010 by the RSIL revealed that samples from 189 patients were positive, either for both the \textit{ptxP} and IS481 qPCR targets or for IS481 alone. [For reporting purposes, the lower limit of detection for the \textit{ptxP} and IS481 components of the PCR are 20 fg \textit{u}l\textsuperscript{-1} and 2 fg \textit{u}l\textsuperscript{-1} \textit{B. pertussis} genomic DNA, respectively (Fry \textit{et al.}, 2009).] This equates to \textasciitilde4.8 and \textasciitilde0.48 genome copies \textit{u}l\textsuperscript{-1}, respectively (calculations not shown; based on published genome size and mol\%GC content of Tohama I strain; Parkhill \textit{et al.}, 2003). From these 189 patients, 131 DNA extracts were available, and were retrieved from \textasciitilde20 °C storage. The DNA had originally been extracted from 45 NPAs, 82 PNSs, 3 bronchoalveolar lavages (BALs) and 1 endotracheal secretion (ETS) using the QIAamp DNA mini kit (Qiagen) or MagNA PURE Compact (Roche), essentially as previously described (Fry \textit{et al.}, 2009). In the case of NPAs, BALs and the ETS, DNA was extracted from 200 \textit{u}l aspirate and eluted in 100 \textit{u}l elution buffer. For PNSs, material was released from the swab into 400 \textit{u}l PBS. DNA was then extracted from 200 \textit{u}l of this and eluted in 100 \textit{u}l elution buffer.

For comparison, 111 clinical \textit{B. pertussis} isolates were randomly selected (out of 226 possible isolates) from patients <1 year of age. These had been submitted to the RSIL between January 2008 and August 2010 by hospital laboratories from England and Wales (as part of a routine surveillance programme). DNA was prepared from each bacterial isolate as previously described (Litt \textit{et al.}, 2009).

**MLVA typing.** MLVA was performed on DNA extracted from \textit{B. pertussis} isolates as previously described (Litt \textit{et al.}, 2009). The method was modified when used with DNA extracted from clinical samples as follows: the number of PCR cycles was increased from 28 to 40; the dilution factor for the resulting PCR products in water was varied between 1/100 and 1/10, when required, to increase the chances of detecting a weakly amplified DNA fragment. MLVA types were assigned using the international reference database at www.mlva.net. BioNumerics 6.1 software (Applied Maths) was used to generate minimum spanning trees from the data, according to the rules previously described by Litt \textit{et al.} (2009).

**Sequence-based typing.** \textit{prn} and \textit{ptxA} sequence-based typing was performed on DNA extracted from bacterial isolates or clinical samples using the method previously described (Litt \textit{et al.}, 2009) with the following changes: the number of PCR cycles was increased to 40; the variable region 2 of \textit{prn} was not sequenced, hence, \textit{prn}(1) and \textit{prn}(7) types were indistinguishable and were designated \textit{prn}(1/7).

For \textit{ptxP} sequence-based typing, the promoter region was amplified using primers and conditions based on those previously described (Mooi \textit{et al.}, 2009; and M. van Gent, personal communication). Each 20 \textit{u}l PCR contained 10 \textit{u}l HotStarTaq Master Mix (Qiagen), 1 \textit{m}M each primer, 1 \textit{m}M betaine (Sigma) and 2 \textit{m}l DNA extract (for both bacterial isolates and clinical samples). Cycling conditions were as follows: 15 min at 95 °C; followed by 5 cycles of 15 s at 95 °C, 30 s at 70/68/66/64/62 °C (decreasing each cycle) and 1 min at 72 °C; followed by 35 cycles of 15 s at 95 °C, 30 s at 60 °C and 1 min at 72 °C. After the last cycle, a final step of 10 min at 72 °C was added. The product was sequenced using the same primers.
Statistical analysis. The age-ranges or distributions of MLVA types in different populations were compared using a two-tailed Fisher’s Exact test via Monte Carlo methods. Genetic diversity indexes (DIs) and their associated 95% confidence intervals were calculated from MLVA types using the methods of Hunter and Gaston, and Simpson, as previously described (Litt et al., 2009). DIs from different populations were also compared statistically as previously described (Litt et al., 2009).

RESULTS

MLVA on qPCR-positive DNA extracts from clinical samples

MLVA was performed on 131 DNA extracts from clinical samples that had previously scored positive in the *B. pertussis* diagnostic qPCR (for the IS481 target alone or for both the IS481 and ptxP targets). As shown in Table 1, a complete six allele MLVA profile was obtained from 106 (81%) of the samples, and a partial profile (between one and five alleles) was generated from another 14 (11%).

Eleven samples yielded no MLVA result. These had only scored positive for the IS481 target in the original qPCR, suggesting that they contained either a low concentration of *B. pertussis* template DNA or DNA from a different *Bordetella* species. All but one had produced a result in the IS481 qPCR equivalent to <6 fg µl⁻¹ of the *B. pertussis* control DNA (~1.4 genome copies µl⁻¹). The 11th sample had generated a result in the qPCR equivalent to ~2 ng µl⁻¹ (~4.8 × 10⁵ genome copies µl⁻¹) of the *B. pertussis* control DNA. Unfortunately, further characterization of this DNA sample by *B. pertussis* qPCR and 16S rRNA gene sequencing were inconclusive (results not shown).

When solely the 114 samples positive for both qPCR targets were analysed, 105 (92%) samples generated a complete MLVA profile (Table 1). A matched bacterial culture was available for six of the qPCR DNA samples. In each case, the MLVA profile generated from the DNA sample and culture was identical (not shown).

DNA sequence-based typing on qPCR samples with a full MLVA profile

Eighty-nine of the 106 qPCR samples that yielded a full MLVA profile were randomly selected for further typing by sequencing the *prn* (region 1), *ptxA* and *ptxP* genetic loci. (All 89 samples had previously scored positive for both the IS481 and *ptxP* target in the qPCR.) Eighty-three (93%) of these samples generated complete, unambiguous, sequences for all three sequencing loci; five (6%) were missing only the *ptxA* result and one sample failed to generate the *ptxA* and *ptxP* results.

Comparison of typing results from qPCR DNA samples and contemporary bacterial cultures

Typing results from the 89 qPCR samples that had undergone MLVA and sequence-based typing were compared with those from 111 randomly selected contemporary *B. pertussis* clinical isolates from patients aged <1 year. Analysis of the age ranges of the patients from the two sample groups showed that they were not statistically different (P=0.95, using the groupings in Table 2). There was no overlap of subjects between the two patient groups.

As shown in Fig. 1, both groups of samples were dominated by the MLVA type 27 (MT-27) *prn*(2) *ptxA*(1) *ptxP*(3) genotype. Overall, 73% (65/89) of qPCR samples and 81% (90/111) of cultures were MT-27 (Fig. 1). Despite differences, the overall distributions of MLVA types in the two populations were not statistically different (P=0.22). The genetic diversity, deduced from the MLVA type distributions, was higher amongst the qPCR samples than amongst the bacterial isolates (Fig. 1). However, this difference was not significant (P=0.16).

When the two groups were combined, 92% (183/200) of the samples possessed all three sequence-based typing alleles *prn*(2) *ptxA*(1) *ptxP*(3); 78% (155/200) of samples were MT-27; and 75% (149/200) were confirmed to have the combined genotype of MT-27 *prn*(2) *ptxA*(1) *ptxP*(3) (not shown). The DI of the MLVA types for the combined population was 0.40 (95% CI: 0.31–0.48).

DISCUSSION

We investigated the feasibility of using molecular typing data generated from clinical samples sent in for *B. pertussis* diagnostic qPCR to supplement or substitute those from *B. pertussis* cultures. Supplementing data from patients’ cultures with those from qPCR samples is desirable, as

<table>
<thead>
<tr>
<th>Original diagnostic qPCR result</th>
<th>MLVA result*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Complete profile</td>
</tr>
<tr>
<td>IS481 positive only</td>
<td>1</td>
</tr>
<tr>
<td>IS481 + <em>ptxP</em> positive</td>
<td>105</td>
</tr>
<tr>
<td>Total</td>
<td>106</td>
</tr>
</tbody>
</table>

*Complete profile, six alleles; partial profile, between one and five alleles.
Table 2. Frequencies in various age ranges of the patient groups from which clinical samples were submitted for diagnostic qPCR or bacterial isolates were cultured

<table>
<thead>
<tr>
<th>Patient age range</th>
<th>qPCR samples group (%)</th>
<th>Bacterial isolates group (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to &lt;1 month</td>
<td>6 (7)</td>
<td>10 (9)</td>
</tr>
<tr>
<td>1 to &lt;2 months</td>
<td>35 (39)</td>
<td>45 (41)</td>
</tr>
<tr>
<td>2 to &lt;3 months</td>
<td>26 (29)</td>
<td>33 (30)</td>
</tr>
<tr>
<td>3 to &lt;4 months</td>
<td>14 (16)</td>
<td>13 (12)</td>
</tr>
<tr>
<td>4 to &lt;5 months</td>
<td>3 (3)</td>
<td>5 (4)</td>
</tr>
<tr>
<td>5 to &lt;12 months</td>
<td>5 (6)</td>
<td>5 (4)</td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>111</td>
</tr>
</tbody>
</table>

the overlap between patients producing a bacterial culture and a positive PCR result is very small. For example, of the samples sent to our laboratory during the time period of this study, 257 patients yielded a bacterial isolate and 189 a positive qPCR result, and, of these, only 25 generated both (data not shown). Furthermore, diagnostic qPCR is generally accepted to be more sensitive than culture, and is gaining in popularity (Mattoo & Cherry, 2005). The ability to type bacterial DNA in qPCR samples will also be valuable in outbreak investigations, when PCR-positive samples, but no cultures, may be available.

Proof of concept for generating MLVA typing data from DNA extracts of clinical samples was demonstrated by Schouls et al. (2004) when they generated 10 complete profiles from 17 samples that were positive for the IS481 PCR target. Using a modified MLVA method, we generated full profiles on >80% of a collection of respiratory samples that scored positive for IS481 (using our in-house qPCR). This rose to >90% if the analysis was limited to samples that were also positive for the single copy ptxP target. For reporting purposes, the limits of detection for the IS481 and the ptxP component of our PCR are 2 fg μl⁻¹ and 20 fg μl⁻¹ (~0.48 and ~4.8 genome copies μl⁻¹) B. pertussis

Fig. 1. Minimum spanning trees showing the MLVA types and prn, ptxA and ptxP genotypes generated from (a) 89 DNA extracts from clinical specimens submitted for B. pertussis diagnostic qPCR and (b) 111 B. pertussis isolates collected from pertussis patients. Samples were received between January 2008 and August 2010. Trees were derived from the six MLVA alleles. Each circle represents a unique MT (shown by the number in the circle). The size of each circle illustrates the number of strains with that MT (the smallest circle in each tree represents one isolate). Thick lines separate MTs that differ at a single locus and thin lines separate those that differ at two loci. Colours illustrate the combination of prn, ptxA and ptxP genotypes (see key). When more than one combination is present for a given MT, the circle is divided proportionally in the form of a pie chart. The genetic diversity index (DI), derived from the MT distribution, is displayed beneath each tree.
genomic DNA, respectively (post-extraction concentrations) (Fry et al., 2009). Eleven samples that were positive for IS481 did not generate any MLVA fragments. In 10 of these cases, this was most likely due to very low concentrations of template DNA. In the 11th case, it appeared that a species other than B. pertussis may have been present. Unfortunately, it was not possible to confirm the identity of the species present in this particular sample. However, it is worth noting that B. bronchiseptica and B. parapertussis should generate complete profiles under this MLVA scheme and B. holmesii produces products in the VNTR2 and VNTR3 PCRs (Schouls et al., 2004; and our unpublished results).

On the assumption that MLVA is more sensitive than sequence-based typing, we attempted prn, ptxA and ptxP typing only on DNA extracts that had generated a complete MLVA profile. With only minor methodological changes, we successfully produced a genotype for all three targets from 93 % (83/89) of extracts tested. We generated data for at least one target for the remaining six samples. Hence, clinical extracts containing >20 fg $\mu l^{-1}$ (~4.8 genome copies $\mu l^{-1}$) B. pertussis DNA (i.e. ptxP-positive) should reliably provide sufficient DNA template for sequence-based typing. We did not attempt this method with samples containing lower concentrations of template DNA. It may be necessary to use a nested PCR method (Nakamura et al., 2011) or a qPCR strategy (Chan et al., 2009; Mäkinen et al., 2002) to obtain alleles of every gene target from samples with very low concentrations of DNA. Our samples were taken from children <1 year of age. Low concentrations of B. pertussis DNA appear to be more common in samples from older children and adults (Nakamura et al., 2011). Another possible explanation for a failure to generate results for ptxP or ptxA sequence-based typing targets from six of our samples is that these B. pertussis isolates carried deletions in their ptx operon that prevented their amplification by PCR. Deletions have been recently described in the ptx operon and in prn of clinical isolates that knock out the expression of pertussis toxin and pertactin, respectively (Bouchez et al., 2009; Hegerle et al., 2012). This was not investigated in our study.

We compared the molecular typing results from qPCR-positive samples to those from contemporary bacterial isolates to determine whether (a) they generated a different distribution of types when performed on the two different patient groups, and (b) combining the data from the two groups enriched the epidemiological data for that time period. The results showed that there was no significant difference in the distribution or diversity of MLVA types between the qPCR samples and the bacterial isolates. However, some rare MLVA types were detected solely in the clinical extracts or the isolates. We propose that this is merely the effect of increasing the overall sample size rather than any bias in MLVA types linked to sample type. Overall, our results suggest that clinical DNA extracts are equivalent to bacterial isolates for the purpose of generating epidemiological data (at least, in the <1 year age group). Adding molecular typing data from clinical extracts to those from bacterial isolates will provide a more complete picture of the circulating B. pertussis population.

Because the samples used in this study were collected as part of our routine surveillance programme, the results also reveal that between January 2008 and August 2010, the UK clinical B. pertussis population in the <1 year age group was dominated by the genotype of MT-27 prn(2) ptxA(1) ptxP(3). This is similar to recently published data from 2005–2006 (Litt et al., 2009) (which was heavily dominated by this age group), and adds ptxP typing data. Hence, it appears that no major changes have occurred to the bacterial population since the UK switched to an acellular pertussis vaccine for primary vaccination in 2004, and its genetic diversity (from MLVA type distribution) continues to fall.

In conclusion, generating MLVA and sequence-based typing data from PCR-positive clinical DNA extracts is straightforward and is likely to generate a complete dataset in the majority of samples in the <1 year age group. The results are representative of those generated from bacterial cultures (at least, in this age group). Hence, positive clinical extracts are suitable for inclusion in molecular epidemiological studies of the B. pertussis population or as a tool in outbreak investigations. This approach may also prove successful with other typing methods, such as qPCR-based SNP analysis (Octavia et al., 2011).

ACKNOWLEDGEMENTS

We are grateful to John Duncan and Lalita Vaghi for assistance in the laboratory, Marjolein van Gent for advice regarding the ptxP sequence-based typing, Steve Platt for his expertise with BioNumerics, Eve Arnold and Ross Harris for performing statistical analyses, and to the clinicians and laboratory staff from hospitals for submitting B. pertussis isolates to our laboratory. Some of this work was conducted by Draga Tchipeva as part of a BSc degree, under the supervision of Norman Fry and Dr Patrick Kimmitt (University of Westminster, UK). Part of this work was presented orally at the 9th International Bordetella Symposium, 30th September–3 October 2010, Baltimore, USA.

REFERENCES


variation and protection induced by a tricomponent acellular vaccine. *Vaccine* 17, 2651–2660.


