2000; $n = 118$). We presumed that the effect of vaccination on the *B. pertussis* population was small in the early WCV period (15, 33). Obviously, the relationship between the periods and the vaccination history can only be approximate.

Two *fim2* alleles were observed in the worldwide collection of strains, *fim2-1* (the vaccine type) and *fim2-2*, the products of which differed in a single amino acid. The *fim2-1* allele predominated in all four periods (frequencies 77% to 98%), whereas the *fim2-2* allele was found at low frequencies (2% to 23%) in all four periods (Fig. 2A). Phylogenetic analysis (Fig. 1B) indicated that the mutation leading to the *fim2-2* allele arose twice within lineage IIb but also occurred on the branch leading to lineage IIa. Bayesian analysis suggested that, within lineage IIb, the mutation occurred between 1970 and 1984 (95% CI, 1956 to 1992) on the first occasion and between 1996 and 2002 (95% CI, 1995 to 2002) on the second. Thus, the first mutation arose in the WCV period and the most recent mutation occurred in the WCV/ACV period.

More variation was found in *fim3*, for which five alleles were identified. As one allele contains a silent mutation, the five alleles code for four distinct proteins: Fim3-1, Fim3-2, Fim3-3, and Fim3-6. The *fim3-1* (the vaccine type) and *fim3-2* alleles were predominant (Fig. 2B). The polymorphic amino acid residue in *fim3-2* relative to the sequence of *fim3-1* is located in a surface epitope that has been shown to interact with human serum (36). The *fim3-1* allele has always predominated, but the *fim3-2* allele, which was first detected in the WCV period (frequency 1%), increased in frequency to 37% in the ACV period. Our analyses agreed with this observation, with the mutation resulting in the *fim3-2* allele predicted to have occurred between 1986 and 1989 (95% CI, 1982 to 1992).

Eight *ptxA* alleles were found worldwide, two of which contained silent mutations. Thus, the eight alleles resulted in six protein variants (PtxA1, PtxA3, PtxA4, PtxA5, PtxA9, and PtxA10), mostly differing by one or two amino acids. Three alleles were predominant, *ptxA1*, *ptxA2*, and *ptxA4* (respective frequencies, 78%, 18%, and 2%). The *ptxA2* and *ptxA4* alleles predominated in the early WCV period (respective frequencies, 64% and 23%). Our analyses show that the *ptxA1* allele arose between 1921 and 1932 (95% CI, 1905 to 1942), before the introduction of vaccination. It increased in frequency from only 5% in the early WCV period to 68%, 92%, and 90% in subsequent periods (Fig. 2C). Although most (46%) of the vaccine strains harbor *ptxA2*, 17% do contain *ptxA1*.

Fourteen *ptxP* alleles were observed, of which *ptxP1* and *ptxP3* predominated (total frequencies of 60% and 32%, respectively). Strains with *ptxP1* were most common in the early WCV and WCV periods (respective frequencies, 68% and 83%) but were replaced by *ptxP3* strains in the last two periods (the *ptxP3* frequencies in the WCV/ACV and ACV periods were 48% and 57%, respectively) (Fig. 2D). Bayesian analysis suggested that the mutation resulting in the *ptxP3* allele arose between 1974 and 1977 (95% CI, 1970 to 1981), i.e., in the WCV period.

Twelve *prn* alleles were identified, of which 11 led to protein variants (Prn1 to -7, Prn10 to -12, and Prn16). Prn-deficient strains were not detected, presumably because these strains reached significant frequencies in a later period than analyzed in this study. Three alleles predominated in our worldwide collection, *prn1* (42%), *prn2* (38%), and *prn3* (12%). In the early WCV period, 67% of the strains harbored *prn1* (the vaccine type), with *prn2* and *prn3* alleles emerging in the WCV period. While the frequency of the *prn3* allele remained more or less constant (10% to 17%), *prn2* increased in frequency from 18% in the WCV period to 65% in the ACV period (Fig. 2E). Variation in *prn* mainly occurs by variation in numbers of repeats, a reversible process which is relatively frequent compared to point mutations. Therefore, many *prn* variants were homoplasic in our tree due to convergent evolution.

In conclusion, based on these five genes, it appears that the worldwide *B. pertussis* population has changed significantly in the last 60 years, consistent with other studies using temporally and geographically less diverse collections (15, 17–19, 21, 22, 32, 34, 37–40). Most changes resulted in genetic divergence from vaccine strains, consistent with vaccine-driven immune selection. Indeed, Bayesian analyses suggested that the non-vaccine-type alleles *ptxP3* and *fim3-2* arose in the period in which the WCV was used widely. Recently, strains have been identified which do not express Prn and/or FHA (17, 23, 24), and the emergence of these strains may be associated with the introduction of ACVs. In this and previous work, the largest number of alleles were observed for *ptxP* ($n = 14$), *prn* ($n = 12$), and *ptxA* ($n = 8$). The number of alleles may be related to the degree of diversifying selection caused, e.g., by the immune status of the host population or other (frequent) changes in the ecology of *B. pertussis*.

Previous studies have shown that changes in *fim3*, *ptxA*, *prn*, and *ptxP* are associated with selective sweeps (15, 19, 22, 32), implying a significant effect on strain fitness. Furthermore, variation in *ptxA*, *ptxP*, and *prn* has been shown to affect bacterial colonization of naive and vaccinated mice (40–45), underlining the biological significance of these changes. However, in one study, the effects were not observed (46).

Identification of additional loci potentially involved in adaptation. In addition to focusing on genes coding for vaccine components, we used a more comprehensive approach to identify putative adaptive loci. To detect genes important for adaptation, *dN*/*dS* ratios (ratio of nonsynonymous to synonymous substitution rates) are widely used. This method was originally developed for the analysis of divergent species and needs a large number of substitutions for a statistically reliable analysis (47–49). However, *B. pertussis* strains are highly related and differ by less than 0.1% in their genomic sequences. Recent studies have shown that the primary driver of *dN*/*dS* ratios in such closely related strains is time, not selection (48). Furthermore, the approach using *dN*/*dS* ratios assumes that silent mutations are neutral. However, silent mutations in genes can significantly affect gene expression (50). Finally, *dN*/*dS* ratios are not useful to detect diversifying selection in intergenic regions. Therefore, we chose to assess diversifying selection by focusing on SNP densities and homoplasy.

SNP densities. We explored whether particular gene categories had a significantly higher SNP density than the overall SNP density of the whole genome, 0.0013 SNPs/bp. The gene categories used were defined by Parkhill et al. (30), with modifications, i.e., pseudogenes and genes known or assumed to be associated with virulence were placed in separate categories. In all, 24 gene categories were defined (Fig. 3A; see Table S5 in the supplemental material). As expected, gene categories involved in housekeeping functions, which are generally conserved, showed the lowest SNP densities (0.0007 to 0.00012 SNPs/bp). The four categories with the highest SNP density were virulence associated (0.0016 SNPs/bp), transport/binding (0.0015 SNPs/bp), protection responses (0.0014 SNPs/bp), and pseudogenes (0.0014 SNPs/bp), which are

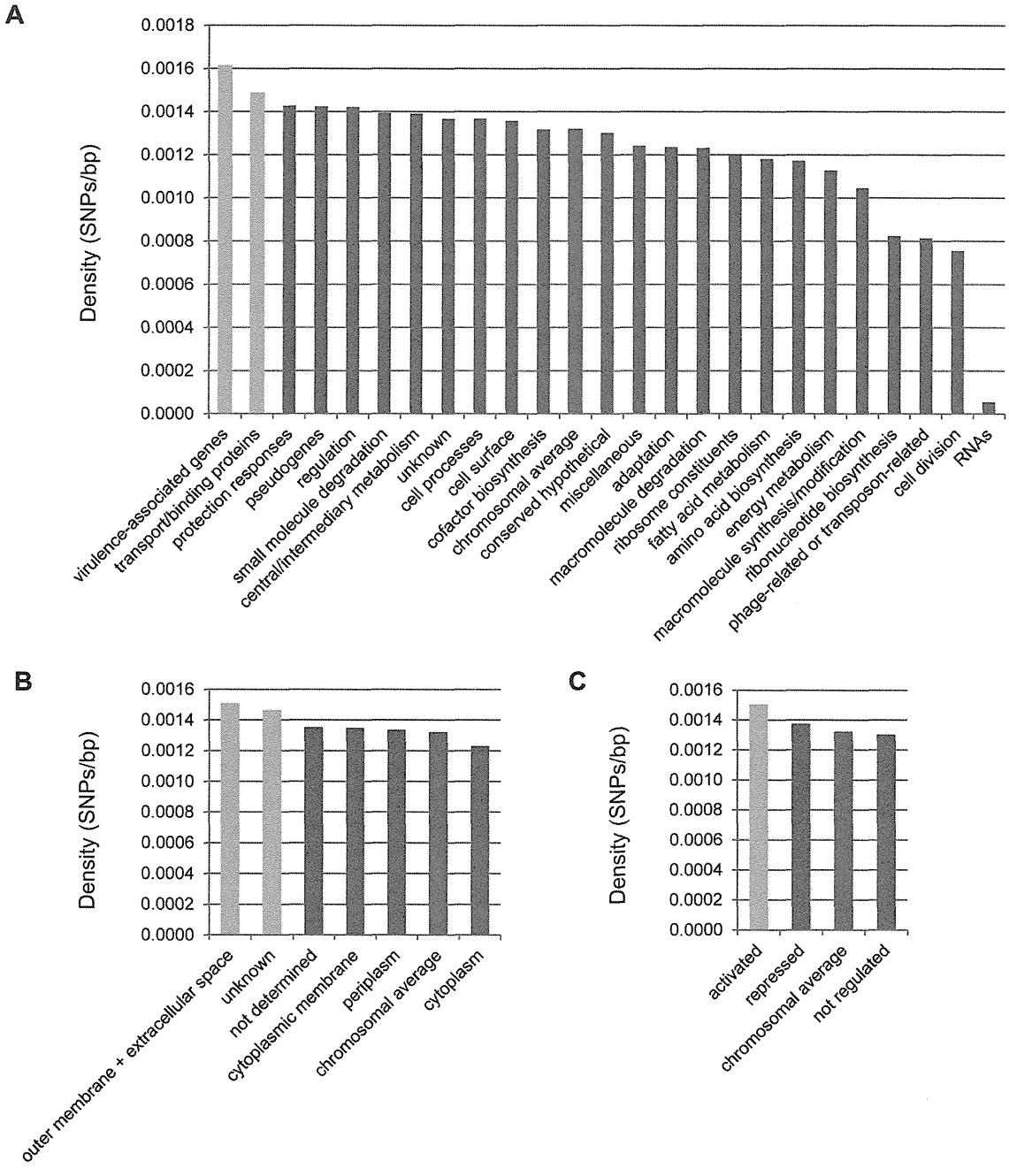

FIG 3 SNP densities per functional category (A), subcellular localization (B), and Bvg regulation (C). Red bars indicate the chromosomal average. Green bars refer to categories with an SNP density significantly higher than the chromosomal average ($P < 0.05$).

likely to be evolving neutrally since their inactivation. Only for the virulence-associated and transport/binding categories did the SNP density difference reach statistical significance, however ($P = 0.02$ and $P = 0.03$, respectively). The high SNP density in the transport/binding category was surprising, as this category mostly codes for housekeeping functions, including transport of molecules such as amino acids, small ions, and carbohydrates. The high SNP density may reflect changes in the physiology of *B. pertussis* or the surface exposure of membrane and periplasmic components of these systems.

To investigate this further, we tested whether the subcellular location of proteins would result in significantly different degrees of SNP density, as surface-exposed proteins are expected to be subject to a higher degree of immune selection than intracellular proteins. In line with this, we found that if categories were based on subcellular location prediction, genes coding for proteins ex-

TABLE 2 Genes and promoters with SNP densities significantly higher than the chromosomal average

Locus tag(s)	Gene(s)	Density (SNPs/bp)	*P* value	Product	Category[a]	Localization(s)[b]	Bvg[c]
3783BP	*ptxA*	0.01111	3.3E-03	Pertussis toxin subunit A precursor	Vir	E	+
2416BP	*cysB*	0.01053	2.9E-03	LysR family transcriptional regulator	Reg	C	
BP3783P	*ptxP*	0.07143	4.7E-18	Pertussis toxin promoter	Vir	E	+
BP2936P		0.03623	2.0E-02	Putative methylase promoter	Exp	CM	+
BP1878P, BP1879P	*bvgP*, *fhaBP*	0.02582	3.4E-05	Virulence factor transcription regulator promoter, filamentous hemagglutinin	Vir	C, OM	+, +
BP3723P, BP3724P		0.02047	1.8E-02	Hypothetical protein promoter	Hyp	U, C	

[a] Functional category: Vir, virulence-associated genes; Reg, regulation; Exp, exported proteins; Hyp, hypothetical proteins.
[b] Subcellular localization: E, extracellular; C, cytoplasmic; CM, cytoplasmic membrane; OM, outer membrane; U, unknown.
[c] Regulation by Bvg: +, activated; blank cells, not activated or repressed.

posed to the host environment (extracellular and outer membrane proteins) had the highest SNP density (0.0015 SNPs/bp; $P = 0.05$), whereas genes coding for cytoplasmic proteins showed the lowest SNP density (0.0012 SNPs/bp; $P = 1.0$) (Fig. 3B; see Table S5 in the supplemental material). In addition to the exposed category, only the category "unknown," which comprises proteins for which we could not predict a location, showed an SNP density which was significantly higher than the genomic average (0.0015 SNPs/bp; $P = 0.007$). For example, Ptx subunits 2 to 5 are included in the unknown category, although it is known that they are secreted (51). Possibly this category compromises more genes that encode surface-exposed proteins but for which the location could not be predicted.

We also assessed the SNP density in gene categories based on Bvg regulation (26, 27). For this, three categories were defined: genes activated, repressed, or unaffected by Bvg (Fig. 3C; see Table S5 in the supplemental material). The SNP density in these three categories decreased in the order Bvg activated, Bvg repressed, and not regulated by Bvg (SNP densities, 0.0015, 0.0014, and 0.0013 SNPs/bp, respectively; $P = 0.013$, $P = 0.40$, and $P = 1.0$, respectively). The relatively high SNP density in Bvg-activated genes was not unexpected, as genes encoding virulence-associated proteins and extracellular proteins are included in this category.

Focusing on gene categories increased the power of the statistical analyses but only gave a general picture and did not reveal individual loci that might be under selection. Therefore, we also identified particular loci which were highly polymorphic. For this, we calculated whether there was an overrepresentation of SNPs in a locus given its length (Table 2; see Table S5 in the supplemental material). Two genes showed a significantly higher SNP density than the chromosomal average of 0.0013 SNPs/bp in genes. One gene encodes Ptx subunit A (*ptxA*) (0.011 SNPs/bp; $P = 0.0033$). The other gene, *cysB* (0.011 SNPs/bp; $P = 0.0029$), encodes a LysR-like transcriptional regulator that acts as an activator of the *cys* genes and plays a role in sulfur metabolism (52, 53).

We also investigated SNP densities in intergenic regions, as these may be involved in transcription of downstream genes. We found four putative promoter regions with a significantly higher SNP density than the chromosomal average of 0.0026 SNPs/bp in intergenic regions (Table 2; see Table S5 in the supplemental material). Two promoter regions were located upstream from virulence-associated genes. One was upstream from the *ptx* operon (0.071 SNPs/bp; $P = 4.7 \times 10^{-18}$), and one was between the filamentous hemagglutinin gene (*fhaB*) and the *bvg* operon (0.026 SNPs/bp; $P = 3.4 \times 10^{-5}$). The extensive polymorphism in the Ptx promoter has been described previously (14, 16). Eleven SNPs were located in the intergenic region between the *bvg* operon and *fhaB*, which has been studied extensively (54–58). Seven and four SNPs were located in regions assumed to affect the transcription of *fhaB* and *bvgA*, respectively (Text S2). While the SNPs in the *fhaB* promoter may affect the expression of both *fha* and *fim* genes, which are part of a single operon (59), the SNPs in the *bvgA* promoter region may have a significant effect on the expression of many virulence factors. A high SNP density was also observed in the region upstream from a putative methylase possibly involved in ubiquinone/menaquinone biosynthesis (0.036 SNPs/bp; $P = 0.020$) and in the promoter region of two hypothetical proteins (0.020 SNPs/bp; $P = 0.018$).

In conclusion, we identified significantly higher SNP densities in virulence-associated genes, genes encoding surface-exposed proteins, and genes activated by Bvg. High SNP densities were also observed in the promoter regions for *ptx* and *bvg/fha*. The finding of a high SNP density in *cysB* was interesting, as a number of associations have been observed between sulfur metabolism and virulence (60). Indeed, in *B. pertussis*, the expression of virulence-associated genes is affected by the sulfate concentration (28). The identification of putative adaptive loci allows focused studies that may reveal novel strategies for pathogen adaptation.

Homoplasic SNPs. In a second approach to find loci possibly involved in adaptation, we identified homoplasic SNPs, that is, SNPs which arose independently on different branches of the tree. In our data set, 15 SNPs were homoplasic (Table 3). Thirty-three percent of the homoplasic SNPs were located in Bvg-activated genes, while this category only comprises 6% of the genome. The 5 SNPs found in Bvg-activated genes were located in genes for the serotype 2 and 3 fimbrial subunits (*fim2* and *fim3*), a type III secretion protein (*bscI*), a Ptx transport protein (*ptlB*), and a periplasmic solute-binding protein (*smoM*) involved in transport of mannitol. Of the remaining 10 homoplasic SNPs, 6 and 4 were located in genes and intergenic regions, respectively. Remarkably, one homoplasic SNP found in *cysM* was observed in five branches. The *cysM* gene codes for cysteine synthase, which is involved in cysteine biosynthesis and sulfate assimilation. All other homoplasic SNPs occurred in two branches.

Convergent evolution is extremely rare in monomorphic bacteria like *B. pertussis*. In other monomorphic bacteria, homoplasy

TABLE 3 Homoplasic SNPs

Position[a]	Locus tag(s)	Gene	Branches[b]	Bootstrap[c]	Change[d]	Product (distance to ATG in bp)	Functional category	Localization[e]	Bvg[f]
612075	BP0607	*gpm*	2 (1, 3)	99	Silent	Phosphoglycerate mutase 1	Energy metabolism	Cytoplasmic	
667028	BP0658		2 (1, 19)	55	Q30	Putative dehydrogenase	Miscellaneous	Cytoplasmic	
925864	BP0888		2 (7, 1)	100	Silent	GntR family transcriptional regulator	Regulation	Cytoplasmic	
997017	BP0958	*cysM*	5 (1, 1, 4, 2, 1)	100	G247E	Cysteine synthase B	Amino acid biosynthesis	Cytoplasmic	
1109310	1064BP	*maeB*	2 (6, 1)	100	Silent	NADP-dependent malic enzyme	Central/intermediary metabolism	Cytoplasmic	
1109312	1064BP	*maeB*	2 (6, 1)	100	Q28P	NADP-dependent malic enzyme	Central/intermediary metabolism	Cytoplasmic	
1175956	1119BP	*fim2*	2 (7, 9)	100	R177K	Serotype 2 fimbrial subunit precursor	Virulence-associated genes	Extracellular	+
1565529	1487BP	*smoM*	2 (1, 4)	100	R176K	Putative periplasmic solute-binding protein	Transport/binding proteins	Unknown	+
1647989	1568BP	*fim3*	2 (1, 1)	98	T130A	Serotype 3 fimbrial subunit precursor	Virulence-associated genes	Extracellular	+
2018882	BP1914P		2 (1, 2)	100	Intergenic	Transposase for IS*1663* (321)	Phage or transposon related	Unknown	
	BP1915P		2 (1, 2)	100	Intergenic	Conserved hypothetical protein (23)	Conserved hypothetical	Unknown	
2213448	BP2090P		2 (8, 1)	100	Intergenic	ABC transporter substrate-binding protein (306)	Transport/binding proteins	Periplasmic	−
	BP2091P		2 (8, 1)	100	Intergenic	Dioxygenase hydroxylase component (53)	Small molecule degradation	Cytoplasmic	−
2374322	2249BP	*bscI*	2 (1, 97)	60	Y114C	Type III secretion protein	Virulence-associated genes	Unknown	+
3041105	BP2862P		2 (6, 1)	100	Intergenic	Conserved hypothetical protein (174)	Unknown	Unknown	
	BP2863P		2 (6, 1)	100	Intergenic	Conserved hypothetical protein (148)	Unknown	Cytoplasmic	
3251279	BP3052P		2 (6, 2)	100	Intergenic	Putative gamma-glutamyl transpeptidase (242)	Miscellaneous	Periplasmic	
3992064	3789BP	*ptlB*	2 (1, 1)	69	Silent	Pertussis toxin transport protein	Virulence-associated genes	CM	+

[a] Position in reference genome *B. pertussis* Tohama I.
[b] Number of branches in which the homoplasic SNP occurred (number of strains/branch).
[c] Number of trees in which SNP is homoplasic (100 trees tested).
[d] Change in amino acid.
[e] Subcellular localization: CM, cytoplasmic membrane.
[f] Regulation by Bvg: + activated; − repressed; blank cells, not activated or repressed.

is usually only found in a few genes involved in antibiotic resistance (61). This suggests that the homoplasic SNPs we have identified may play an important role in the adaptation of *B. pertussis*.

Gene loss. Several studies have shown that some *B. pertussis* isolates contain DNA that is not in Tohama but is present in *Bordetella bronchiseptica* and *Bordetella parapertussis* (62–66). In this work, we performed a *de novo* assembly of all of the genomes and compared each assembly back against the reference Tohama I in order to identify any genomic DNA that may have been acquired since the origin of *B. pertussis*. This analysis showed no evidence of gene gain at any point in the phylogeny. All regions identified in the sample data set that were not in Tohama are present in other *Bordetella pertussis* genomes, such as 18323, consistent with gene loss in Tohama. Placing these regions onto the tree showed that progressive gene loss within multiple lineages can be observed (see Fig. S3 in the supplemental material).

Summary. With the determination of the global population structure of *B. pertussis* using whole-genome sequencing, we addressed key questions concerning the origin of pertussis, such as the forces that have driven the shifts in *B. pertussis* populations and the role of these shifts in the resurgence of pertussis. Despite a structure suggesting two relatively recent introductions of *B. pertussis* from an unknown reservoir, phylogenetic analysis did not reveal the ancient geographic origin of *B. pertussis*, possibly because rapid worldwide spread and selective sweeps have eliminated geographic signatures. Indeed, our results showed that the mutation that resulted in the *ptxP3* allele, which is associated with an increase in pertussis notifications in at least two countries (14, 20), occurred once and strains carrying this new allele spread worldwide in 25 to 30 years.

We confirmed and extended the observation that the worldwide *B. pertussis* population has changed significantly in the last

60 years, consistent with other studies using temporally and geographically less diverse collections (15, 17–19, 21, 32, 34, 37–40). We used several approaches to identify gene categories under selection, including SNP density and homoplasy. These approaches consistently suggested that Bvg-activated genes and genes coding for surface-exposed proteins were important for adaptation. At the individual gene level, four of the five genes for the components of current ACVs were found to be particularly variable, underlining their role in inducing protective immunity and consistent with vaccine-driven immune selection.

We identified other, less obvious genes which contained potentially adaptive mutations, such as two genes involved in cysteine and sulfate metabolism (*cysB* and *cysM*). Sulfate can be used to regulate virulence-associated genes *in vitro* (67), and our results suggest that sulfate may also be an important cue during natural infection. This result suggests that host-pathogen signaling and/or the physiology of *B. pertussis* has changed over time.

Temporal analyses showed that most mutations in genes encoding acellular vaccine components arose in the period in which the WCV was used. It should be noted, however, that the period in which the WCV was used (30 to 40 years) is much longer than the ACV period (7 to 15 years). These results are consistent with a significant effect of vaccination on the *B. pertussis* population, as suggested by previous studies (5, 20, 32, 39, 68). It seems plausible that the changes in the *B. pertussis* populations have reduced vaccine efficacy.

Pathogen adaptations may reveal weak spots in the bacterial defense, and hence, the loci under selective pressure may point to ways to improve pertussis vaccines. Furthermore, many of the putative adaptive loci we identified have a physiological role, and future studies of these loci may reveal less obvious ways in which the pathogen and host interact.

MATERIALS AND METHODS

Strains and sequencing. The clinical isolates used in this study are listed in Table S1 in the supplemental material. DNA was isolated by the participants and sequenced using Illumina technology (69). Nineteen isolates were sequenced using the Genome Analyzer II and resulting in single reads of 37 bp (sequencing method 1). Thirty-eight isolates were sequenced using the Genome Analyzer II and resulting in paired-end reads of 50 bp (sequencing method 2). The remaining isolates were sequenced using 12 multiplexed tags on the Genome Analyzer II, producing paired-end reads of 54 bp (sequencing method 3). The accession numbers of the raw sequence data are listed in Table S1.

SNP detection. Reads for all sequenced samples were mapped against the complete Tohama I reference genome sequence (accession number BX470248) using SMALT (http://www.sanger.ac.uk/resources/software/smalt/). Reads mapping with identical matches to two regions of the reference genome were left unmapped. The alignment of reads around insertions and deletions (indels) was improved using a combination of pindel (70) to identify short indels and dindel (71) to realign the reads. SNPs were identified using samtools mpileup (http://samtools.sourceforge.net) and filtered as described previously (72)

Information about promoters, genes, and proteins was retrieved from the sequenced genome of *B. pertussis* Tohama I. The annotation was updated using BLAST (73), and domain information was recovered from SMART (74) and Conserved Domain Database (75).

Homoplasic SNPs were identified by reconstructing base changes for each variable site onto the phylogenetic tree under the parsimony criterion. Any site for which the observed number of base changes for the maximum parsimony reconstruction on the tree was greater than the minimum possible number of changes for that site is homoplasic.

Phylogeny. The phylogenetic relationships of the entire data set were inferred under a maximum likelihood framework using PHYML (76) with an HKY85 model of evolution. The global phylogeny was rooted using *B. bronchiseptica* MO149 (sequence type 15 [ST15]), which was previously shown to be most closely related to *B. pertussis* (77, 78).

Mutation rates and ancestral node dates for lineage IIb were estimated using Bayesian analysis in the BEAST version 1.6.2 package (79). Analyses using the variable sites within lineage IIb isolates with isolation dates available were run under a general time reversible (GTR) model of evolution, with all combinations of constant, expansion, logistic and skyline population size models, and strict, relaxed exponential, and relaxed log-normal clock models. For each combination, three independent Markov chains were run for 100 million generations each, with parameter values sampled every 1,000 generations. Chains were manually checked for reasonable ESS values and for convergence between the three replicate chains using Tracer. Tracer was also used to identify a suitable burn-in period to remove from the beginning of each chain, as well as to assess the model with the best fit to the data using Bayes factors. A skyline population model with a relaxed exponential clock model was identified as the most appropriate, so this combination of models was used for all further analyses. It was found that, in each case, a burn-in of 10 million generations was clearly past the point where chains appeared to have converged, so this was chosen as the burn-in for all chains. The burn-in was removed and chains combined and down-sampled to every 10,000 generations using LogCombiner. A Bayesian skyline plot was calculated in Tracer using the default parameters, and a maximum clade credibility tree computed with TreeAnnotator.

SNP densities. The functional categories used were defined by Parkhill et al. (30), with modifications, i.e., pseudogenes and genes known or assumed to be associated with virulence were placed in separate categories. Subcellular localization was predicted by PSORTb version 3.0 (80). Bvg categories were defined based on the results of Streefland et al. (27) and Cummings et al. (26). For the length of a specific category or locus repeat, regions were excluded because SNPs in these regions are not reliable. To determine the number of bases in a specific category, the lengths of the included loci were added, excluding repeat regions. To determine whether the SNP density of a particular group or locus was significantly higher than the chromosomal average, Fisher's exact test was used. *P* values were corrected according to the method of Benjamini and Hochberg (81).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.01074-14/-/DCSupplemental.

Table S1, XLSX file, 0.1 MB.
Table S2, XLSX file, 0.1 MB.
Table S3, XLSX file, 0.1 MB.
Table S4, XLSX file, 0.2 MB.
Table S5, XLSX file, 0.5 MB.
Figure S1, PDF file, 0.4 MB.
Figure S2, PDF file, 0.2 MB.
Figure S3, PDF file, 0.5 MB.
Text S1, DOCX file, 0.1 MB.
Text S2, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

This effort was initiated during the *Bordetella* Workshop in Cambridge on 22 to 24 July 2008. Therefore, we are very grateful to the attendees and especially to Olivier Restif, who organized this meeting. We thank Gwendolyn L. Gilbert (Westmead Hospital, Australia) and Margaret Ip (the Chinese University of Hong Kong) for supplying strains.

This work was supported by the Wellcome Trust (grant number 098051), the RIVM (SOR project S/230446/01/BV), and the National Health and Medical Research Council of Australia.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. **Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, Jha P, Campbell H, Walker CF, Cibulskis R, Eisele T, Liu L, Mathers C, Child Health Epidemiology Reference Group of WHO and UNICEF.** 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet **375**:1969–1987. http://dx.doi.org/10.1016/S0140-6736(10)60549-1.
2. **Still GF.** 1965. The history of paediatrics. Oxford University Press, London, United Kingdom.
3. **Lapin JH.** 1943. Whooping cough. Charles C. Thomas Publisher Ltd., Springfield, IL.
4. **Magner LN.** 1993. Diseases of the premodern period in Korea, p 392–400. *In* Kiple KF (ed), The Cambridge world history of human disease. Cambridge University Press, Cambridge, United Kingdom.
5. **Mooi FR.** 2010. *Bordetella pertussis* and vaccination: the persistence of a genetically monomorphic pathogen. Infect. Genet. Evol. **10**:36–49. http://dx.doi.org/10.1016/j.meegid.2009.10.007.
6. **Spokes PJ, Quinn HE, McAnulty JM.** 2010. Review of the 2008-2009 pertussis epidemic in NSW: notifications and hospitalisations. N. S. W. Public Health Bull. **21**:167–173. http://dx.doi.org/10.1071/NB10031.
7. **Health Protection Agency.** 25 October 2012. Health protection report. Vol 6 No 43. Health Protection Agency, London England. http://www.hpa.org.uk/hpr/archives/2012/hpr4312.pdf.
8. **Conyn van Spaendock M, Van der Maas N, Mooi F.** 2013. Control of whooping cough in the Netherlands: optimisation of the vaccination policy. RIVM letter report 215121002. National Institute for Public Health and the Environment, Bilthoven, The Netherlands. http://www.rivm.nl/bibliotheek/rapporten/215121002.pdf.
9. **Winter K, Harriman K, Zipprich J, Schechter R, Talarico J, Watt J, Chavez G.** 2012. California pertussis epidemic, 2010. J. Pediatr. **161**:1091–1096. http://dx.doi.org/10.1016/j.jpeds.2012.05.041.
10. **DeBolt C, Tasslimi A, Bardi J, Leader B, Hiatt B, Quin X, Patel M, Martin S, Tondella ML, Cassiday P, Faulkner A, Messonnier NE, Clark TA, Meyer S.** 2012. Pertussis epidemic—Washington, 2012. MMWR Morb. Mortal. Wkly. Rep. **61**:517–522.
11. **Klein NP, Bartlett J, Rowhani-Rahbar A, Fireman B, Baxter R.** 2012. Waning protection after fifth dose of acellular pertussis vaccine in children. N. Engl. J. Med. **367**:1012–1019. http://dx.doi.org/10.1056/NEJMoa1200850.
12. **Sheridan SL, Ware RS, Grimwood K, Lambert SB.** 2012. Number and order of whole cell pertussis vaccines in infancy and disease protection. JAMA **308**:454–456. http://dx.doi.org/10.1001/jama.2012.6364.
13. **Misegades LK, Winter K, Harriman K, Talarico J, Messonnier NE, Clark TA, Martin SW.** 2012. Association of childhood pertussis with receipt of 5 doses of pertussis vaccine by time since last vaccine dose, California, 2010. JAMA **308**:2126–2132. http://dx.doi.org/10.1001/jama.2012.14939.
14. **Mooi FR, van Loo IH, van Gent M, He Q, Bart MJ, Heuvelman KJ, de Greeff SC, Diavatopoulos D, Teunis P, Nagelkerke N, Mertsola J.** 2009. *Bordetella pertussis* strains with increased toxin production associated with pertussis resurgence. Emerg. Infect. Dis. **15**:1206–1213. http://dx.doi.org/10.3201/eid1508.081511.
15. **van Gent M, Bart MJ, van der Heide HG, Heuvelman KJ, Mooi FR.** 2012. Small mutations in *Bordetella pertussis* are associated with selective sweeps. PLoS One **7**:e46407. http://dx.doi.org/10.1371/journal.pone.0046407.
16. **Advani A, Gustafsson L, Ahrén C, Mooi FR, Hallander HO.** 2011. Appearance of Fim3 and ptxP3-*Bordetella pertussis* strains, in two regions of Sweden with different vaccination programs. Vaccine **29**:3438–3442. http://dx.doi.org/10.1016/j.vaccine.2011.02.070.
17. **Hegerle N, Paris AS, Brun D, Dore G, Njamkepo E, Guillot S, Guiso N.** 2012. Evolution of French *Bordetella pertussis* and *Bordetella parapertussis* isolates: increase of bordetellae not expressing pertactin. Clin. Microbiol. Infect. **18**:E340–E346. http://dx.doi.org/10.1111/j.1469-0691.2012.03925.x.
18. **Kallonen T, Mertsola J, Mooi FR, He Q.** 2012. Rapid detection of the recently emerged *Bordetella pertussis* strains with the ptxP3 pertussis toxin promoter allele by real-time PCR. Clin. Microbiol. Infect. **18**:E377–E379. http://dx.doi.org/10.1111/j.1469-0691.2012.04000.x.
19. **Lam C, Octavia S, Bahrame Z, Sintchenko V, Gilbert GL, Lan R.** 2012. Selection and emergence of pertussis toxin promoter ptxP3 allele in the evolution of *Bordetella pertussis*. Infect. Genet. Evol. **12**:492–495. http://dx.doi.org/10.1016/j.meegid.2012.01.001.
20. **Octavia S, Sintchenko V, Gilbert GL, Lawrence A, Keil AD, Hogg G, Lan R.** 2012. Newly emerging clones of *Bordetella pertussis* carrying prn2 and ptxP3 alleles implicated in Australian pertussis epidemic in 2008-2010. J. Infect. Dis. **205**:1220–1224. http://dx.doi.org/10.1093/infdis/jis178.
21. **Petersen RF, Dalby T, Dragsted DM, Mooi F, Lambertsen L.** 2012. Temporal trends in *Bordetella pertussis* populations, Denmark, 1949-2010. Emerg. Infect. Dis. **18**:767–774. http://dx.doi.org/10.3201/eid1805.110812.
22. **Schmidtke AJ, Boney KO, Martin SW, Skoff TH, Tondella ML, Tatti KM.** 2012. Population diversity among *Bordetella pertussis* isolates, United States, 1935-2009. Emerg. Infect. Dis. **18**:1248–1255. http://dx.doi.org/10.3201/eid1808.120082.
23. **Barkoff AM, Mertsola J, Guillot S, Guiso N, Berbers G, He Q.** 2012. Appearance of *Bordetella pertussis* strains not expressing the vaccine antigen pertactin in Finland. Clin. Vaccine Immunol. **19**:1703–1704. http://dx.doi.org/10.1128/CVI.00367-12.
24. **Otsuka N, Han HJ, Toyoizumi-Ajisaka H, Nakamura Y, Arakawa Y, Shibayama K, Kamachi K.** 2012. Prevalence and genetic characterization of pertactin-deficient *Bordetella pertussis* in Japan. PLoS One **7**:e31985. http://dx.doi.org/10.1371/journal.pone.0031985.
25. **Queenan AM, Cassiday PK, Evangelista A.** 2013. Pertactin-negative variants of *Bordetella pertussis* in the United States. N. Engl. J. Med. **368**:583–584. http://dx.doi.org/10.1056/NEJMc1209369.
26. **Cummings CA, Bootsma HJ, Relman DA, Miller JF.** 2006. Species- and strain-specific control of a complex, flexible regulon by Bordetella BvgAS. J. Bacteriol. **188**:1775–1785. http://dx.doi.org/10.1128/JB.188.5.1775-1785.2006.
27. **Streefland M, van de Waterbeemd B, Happé H, van der Pol LA, Beuvery EC, Tramper J, Martens DE.** 2007. PAT for vaccines: the first stage of PAT implementation for development of a well-defined whole-cell vaccine against whooping cough disease. Vaccine **25**:2994–3000. http://dx.doi.org/10.1016/j.vaccine.2007.01.015.
28. **Stibitz S.** 2007. The bvg regulon, p 47–67. *In* Locht C (ed), *Bordetella* molecular microbiology. Horizon Bioscience, Norfolk, United Kingdom.
29. **van Gent M, Bart MJ, van der Heide HG, Heuvelman KJ, Kallonen T, He Q, Mertsola J, Advani A, Hallander HO, Janssens K, Hermans PW, Mooi FR.** 2011. SNP-based typing: a useful tool to study *Bordetella pertussis* populations. PLoS One **6**:e20340. http://dx.doi.org/10.1371/journal.pone.0020340.
30. **Parkhill J, Sebaihia M, Preston A, Murphy LD, Thomson N, Harris DE, Holden MT, Churcher CM, Bentley SD, Mungall KL, Cerdeño-Tárraga AM, Temple L, James K, Harris B, Quail MA, Achtman M, Atkin R, Baker S, Basham D, Bason N, Cherevach I, Chillingworth T, Collins M, Cronin A, Davis P, Doggett J, Feltwell T, Goble A, Hamlin N, Hauser H, Holroyd S, Jagels K, Leather S, Moule S, Norberczak H, O'Neil S, Ormond D, Price C, Rabbinowitsch E, Rutter S, Sanders M, Saunders D, Seeger K, Sharp S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Unwin L, Whitehead S, Barrell BG, Maskell DJ.** 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. Nat. Genet. **35**:32–40. http://dx.doi.org/10.1038/ng1227.
31. **Moran NA, Plague GR.** 2004. Genomic changes following host restriction in bacteria. Curr. Opin. Genet. Dev. **14**:627–633. http://dx.doi.org/10.1016/j.gde.2004.09.003.
32. **Litt DJ, Neal SE, Fry NK.** 2009. Changes in genetic diversity of the *Bordetella pertussis* population in the United Kingdom between 1920 and 2006 reflect vaccination coverage and emergence of a single dominant clonal type. J. Clin. Microbiol. **47**:680–688. http://dx.doi.org/10.1128/JCM.01838-08.
33. **Van Loo IH, Mooi FR.** 2002. Changes in the Dutch *Bordetella pertussis* population in the first 20 years after the introduction of whole-cell vaccines. Microbiology **148**:2011–2018.
34. **Weber C, Boursaux-Eude C, Coralie G, Caro V, Guiso N.** 2001. Polymorphism of *Bordetella pertussis* isolates circulating for the last 10 years in France, where a single effective whole-cell vaccine has been used for more than 30 years. J. Clin. Microbiol. **39**:4396–4403. http://dx.doi.org/10.1128/JCM.39.12.4396-4403.2001.
35. **Berbers GA, de Greeff SC, Mooi FR.** 2009. Improving pertussis vaccination. Hum. Vaccine **5**:497–503.
36. **Williamson P, Matthews R.** 1996. Epitope mapping the Fim2 and Fim3

proteins of *Bordetella pertussis* with sera from patients infected with or vaccinated against whooping cough. FEMS Immunol. Med. Microbiol. **13**:169–178. http://dx.doi.org/10.1111/j.1574-695X.1996.tb00231.x.
37. **Packard ER, Parton R, Coote JG, Fry NK.** 2004. Sequence variation and conservation in virulence-related genes of *Bordetella pertussis* isolates from the UK. J. Med. Microbiol. **53**:355–365. http://dx.doi.org/10.1099/jmm.0.05515-0.
38. **Octavia S, Maharjan RP, Sintchenko V, Stevenson G, Reeves PR, Gilbert GL, Lan R.** 2011. Insight into evolution of *Bordetella pertussis* from comparative genomic analysis: evidence of vaccine-driven selection. Mol. Biol. Evol. **28**:707–715. http://dx.doi.org/10.1093/molbev/msq245.
39. **Mooi FR, van Oirschot H, Heuvelman K, van der Heide HG, Gaastra W, Willems RJ.** 1998. Polymorphism in the Bordetella pertussis virulence factors, p 69/pertactin and pertussis toxin in The Netherlands: temporal trends and evidence for vaccine-driven evolution. Infect. Immun. **66**: 670–675.
40. **Bottero D, Gaillard ME, Fingermann M, Weltman G, Fernández J, Sisti F, Graieb A, Roberts R, Rico O, Ríos G, Regueira M, Binsztein N, Hozbor D.** 2007. Pulsed-field gel electrophoresis, pertactin, pertussis toxin S1 subunit polymorphisms, and surfaceome analysis of vaccine and clinical *Bordetella pertussis* strains. Clin. Vaccine Immunol. **14**: 1490–1498. http://dx.doi.org/10.1128/CVI.00177-07.
41. **King AJ, Berbers G, van Oirschot HF, Hoogerhout P, Knipping K, Mooi FR.** 2001. Role of the polymorphic region 1 of the *Bordetella pertussis* protein pertactin in immunity. Microbiology **147**:2885–2895.
42. **Watanabe M, Nagai M.** 2002. Effect of acellular pertussis vaccine against various strains of *Bordetella pertussis* in a murine model of respiratory infection. J. Health Sci. **48**:560. http://dx.doi.org/10.1248/jhs.48.560.
43. **Gzyl A, Augustynowicz E, Gniadek G, Rabczenko D, Dulny G, Slusarczyk J.** 2004. Sequence variation in pertussis S1 subunit toxin and pertussis genes in *Bordetella pertussis* strains used for the whole-cell pertussis vaccine produced in Poland since 1960: efficiency of the DTwP vaccine-induced immunity against currently circulating *B. pertussis* isolates. Vaccine **22**:2122–2128. http://dx.doi.org/10.1016/j.vaccine.2003.12.006.
44. **Komatsu E, Yamaguchi F, Abe A, Weiss AA, Watanabe M.** 2010. Synergic effect of genotype changes in pertussis toxin and pertactin on adaptation to an acellular pertussis vaccine in the murine intranasal challenge model. Clin. Vaccine Immunol. **17**:807–812. http://dx.doi.org/10.1128/CVI.00449-09.
45. **van Gent M, van Loo IH, Heuvelman KJ, de Neeling AJ, Teunis P, Mooi FR.** 2011. Studies on Prn variation in the mouse model and comparison with epidemiological data. PLoS One **6**:e18014. http://dx.doi.org/10.1371/journal.pone.0018014.
46. **Denoël P, Godfroid F, Guiso N, Hallander H, Poolman J.** 2005. Comparison of acellular pertussis vaccines-induced immunity against infection due to *Bordetella pertussis* variant isolates in a mouse model. Vaccine **23**:5333–5341. http://dx.doi.org/10.1016/j.vaccine.2005.06.021.
47. **Novichkov PS, Wolf YI, Dubchak I, Koonin EV.** 2009. Trends in prokaryotic evolution revealed by comparison of closely related bacterial and archaeal genomes. J. Bacteriol. **191**:65–73. http://dx.doi.org/10.1128/JB.01237-08.
48. **Rocha EP, Smith JM, Hurst LD, Holden MT, Cooper JE, Smith NH, Feil EJ.** 2006. Comparisons of dN/dS are time dependent for closely related bacterial genomes. J. Theor. Biol. **239**:226–235. http://dx.doi.org/10.1016/j.jtbi.2005.08.037.
49. **Yang Z, Bielawski JP.** 2000. Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. **15**:496–503. http://dx.doi.org/10.1016/S0169-5347(00)01994-7.
50. **Kudla G, Murray AW, Tollervey D, Plotkin JB.** 2009. Coding-sequence determinants of gene expression in *Escherichia coli*. Science **324**:255–258. http://dx.doi.org/10.1126/science.1170160.
51. **Locht C, Coutte L, Mielcarek N.** 2011. The ins and outs of pertussis toxin. FEBS J. **278**:4668–4682. http://dx.doi.org/10.1111/j.1742-4658.2011.08237.x.
52. **Imperi F, Tiburzi F, Fimia GM, Visca P.** 2010. Transcriptional control of the pvdS iron starvation sigma factor gene by the master regulator of sulfur metabolism CysB in *Pseudomonas aeruginosa*. Environ. Microbiol. **12**: 1630–1642. doi: 10.1111/j.1462-2920.2010.02210.x.
53. **Kredich NM.** 1992. The molecular basis for positive regulation of cys promoters in *Salmonella typhimurium* and *Escherichia coli*. Mol. Microbiol. **6**:2747–2753. http://dx.doi.org/10.1111/j.1365-2958.1992.tb01453.x.
54. **Boucher PE, Maris AE, Yang MS, Stibitz S.** 2003. The response regulator BvgA and RNA polymerase alpha subunit C-terminal domain bind simultaneously to different faces of the same segment of promoter DNA. Mol. Cell **11**:163–173. http://dx.doi.org/10.1016/S1097-2765(03)00007-8.
55. **Boucher PE, Murakami K, Ishihama A, Stibitz S.** 1997. Nature of DNA binding and RNA polymerase interaction of the *Bordetella pertussis* BvgA transcriptional activator at the *fha* promoter. J. Bacteriol. **179**:1755–1763.
56. **Boucher PE, Yang MS, Schmidt DM, Stibitz S.** 2001. Genetic and biochemical analyses of BvgA interaction with the secondary binding region of the *fha* promoter of *Bordetella pertussis*. J. Bacteriol. **183**:536–544. http://dx.doi.org/10.1128/JB.183.2.536-544.2001.
57. **Boucher PE, Yang MS, Stibitz S.** 2001. Mutational analysis of the high-affinity BvgA binding site in the fha promoter of *Bordetella pertussis*. Mol. Microbiol. **40**:991–999. http://dx.doi.org/10.1046/j.1365-2958.2001.02442.x.
58. **Decker KB, Chen Q, Hsieh ML, Boucher P, Stibitz S, Hinton DM.** 2011. Different requirements for σ region 4 in BvgA activation of the *Bordetella pertussis* promoters P(fim3) and P(fhaB). J. Mol. Biol. **409**:692–709. http://dx.doi.org/10.1016/j.jmb.2011.04.017.
59. **Mattoo S, Miller JF, Cotter PA.** 2000. Role of *Bordetella* bronchiseptica fimbriae in tracheal colonization and development of a humoral immune response. Infect. Immun. **68**:2024–2033. http://dx.doi.org/10.1128/IAI.68.4.2024-2033.2000.
60. **Łochowska A, Iwanicka-Nowicka R, Zielak A, Modelewska A, Thomas MS, Hryniewicz MM.** 2011. Regulation of sulfur assimilation pathways in *Burkholderia cenocepacia* through control of genes by the SsuR transcription factor. J. Bacteriol. **193**:1843–1853. http://dx.doi.org/10.1128/JB.00483-10.
61. **Achtman M.** 2012. Insights from genomic comparisons of genetically monomorphic bacterial pathogens. Philos. Trans. R. Soc. Lond. B Biol. Sci. **367**:860–867. http://dx.doi.org/10.1098/rstb.2011.0303.
62. **Brinig MM, Cummings CA, Sanden GN, Stefanelli P, Lawrence A, Relman DA.** 2006. Significant gene order and expression differences in Bordetella pertussis despite limited gene content variation. J. Bacteriol. **188**:2375–2382. http://dx.doi.org/10.1128/JB.188.7.2375-2382.2006.
63. **Caro V, Bouchez V, Guiso N.** 2008. Is the sequenced *Bordetella pertussis* strain Tohama I representative of the species? J. Clin. Microbiol. **46**: 2125–2128. http://dx.doi.org/10.1128/JCM.02484-07.
64. **Bouchez V, Caro V, Levillain E, Guigon G, Guiso N.** 2008. Genomic content of *Bordetella pertussis* clinical isolates circulating in areas of intensive children vaccination. PLoS One **3**:e2437. http://dx.doi.org/10.1371/journal.pone.0002437.
65. **King AJ, van Gorkom T, van der Heide HG, Advani A, van der Lee S.** 2010. Changes in the genomic content of circulating *Bordetella pertussis* strains isolated from the Netherlands, Sweden, Japan and Australia: adaptive evolution or drift? BMC Genomics **11**:64. http://dx.doi.org/10.1186/1471-2164-11-64.
66. **Bart MJ, van Gent M, van der Heide HG, Boekhorst J, Hermans P, Parkhill J, Mooi FR.** 2010. Comparative genomics of prevaccination and modern *Bordetella pertussis* strains. BMC Genomics **11**:627. http://dx.doi.org/10.1186/1471-2164-11-627.
67. **Bogdan JA, Nazario-Larrieu J, Sarwar J, Alexander P, Blake MS.** 2001. *Bordetella pertussis* autoregulates pertussis toxin production through the metabolism of cysteine. Infect. Immun. **69**:6823–6830. http://dx.doi.org/10.1128/IAI.69.11.6823-6830.2001.
68. **Njamkepo E, Cantinelli T, Guigon G, Guiso N.** 2008. Genomic analysis and comparison of *Bordetella pertussis* isolates circulating in low and high vaccine coverage areas. Microbes Infect. **10**:1582–1586. http://dx.doi.org/10.1016/j.micinf.2008.09.012.
69. **Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, et al.** 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature **456**: 53–59. http://dx.doi.org/10.1038/nature07517.
70. **Ye K, Schulz MH, Long Q, Apweiler R, Ning Z.** 2009. Pindel: a pattern growth approach to detect breakpoints of large deletions and medium sized insertions from paired-end short reads. Bioinformatics **25**: 2865–2871. http://dx.doi.org/10.1093/bioinformatics/btp394.

71. **Albers CA, Lunter G, MacArthur DG, McVean G, Ouwehand WH, Durbin R.** 2011. Dindel: accurate indel calls from short-read data. Genome Res. **21**:961–973. http://dx.doi.org/10.1101/gr.112326.110.
72. **Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD.** 2010. Evolution of MRSA during hospital transmission and intercontinental spread. Science **327**: 469–474. http://dx.doi.org/10.1126/science.1182395.
73. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
74. **Schultz J, Milpetz F, Bork P, Ponting CP.** 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. U. S. A. **95**:5857–5864.
75. **Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH.** 2013. CDD: conserved domains and protein three-dimensional structure. Nucleic Acids Res. **41**:D348–D352. http://dx.doi.org/10.1093/nar/gks1243.
76. **Guindon S, Gascuel O.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**: 696–704. http://dx.doi.org/10.1080/10635150390235520.
77. **Diavatopoulos DA, Cummings CA, Schouls LM, Brinig MM, Relman DA, Mooi FR.** 2005. Bordetella pertussis, the causative agent of whooping cough, evolved from a distinct, human-associated lineage of B. bronchiseptica. PLoS Pathog. 1:e45. http://dx.doi.org/10.1371/journal.ppat.0010045.
78. **Park J, Zhang Y, Buboltz AM, Zhang X, Schuster SC, Ahuja U, Liu M, Miller JF, Sebaihia M, Bentley SD, Parkhill J, Harvill ET.** 2012. Comparative genomics of the classical Bordetella subspecies: the evolution and exchange of virulence-associated diversity amongst closely related pathogens. BMC Genomics **13**:545. http://dx.doi.org/10.1186/1471-2164-13-545.
79. **Drummond AJ, Suchard MA, Xie D, Rambaut A.** 2012. Bayesian phylogenetics with BEAUti and the BEAST. Mol. Biol. Evol. **29**:1969–1973. http://dx.doi.org/10.1093/molbev/mss075.
80. **Yu NY, Laird MR, Spencer C, Brinkman FS.** 2011. PSORTdb—an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. Nucleic Acids Res. **39**:D241–D244. http://dx.doi.org/10.1093/nar/gkq1093.
81. **Benjamini Y, Hochberg Y.** 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B Stat. Methodol. **57**:289–300.

Heat shock protein 70 regulates degradation of the mumps virus phosphoprotein via the ubiquitin-proteasome pathway

Hiroshi Katoh[1]*, Toru Kubota[1], Shunsuke Kita[2], Yuichiro Nakatsu[1], Natsuko Aoki[1], Yoshio Mori[1], Katsumi Maenaka[2], Makoto Takeda[1], Minoru Kidokoro[1]

[1]Department of Virology III, National Institute of Infectious Diseases, Tokyo, and [2]Laboratory of Biomolecular Science, Faculty of Pharmaceutical Sciences, Hokkaido University, Hokkaido, Japan

+To whom correspondence should be addressed:
Hiroshi Katoh, D.V.M., Ph.D
Department of Virology III
National Institute of Infectious Diseases
4-7-1 Gakuen, Musashimurayama-shi
Tokyo 208-0011, Japan
Tel: 81-42-848-7064
Fax: 81-42-567-5631
E-mail: kato0704@nih.go.jp

Running title: Hsp72 regulates degradation of MuV P protein
Key Words: Heat shock protein 70 / Mumps virus / Phosphoprotein / Ubiquitin proteasome

The authors declare no conflict of interest.

Abstract: 161 words
Importance: 114 words
Text: 6,822 words

Abstract

Mumps virus (MuV) infection induces formation of cytoplasmic inclusion bodies (IBs). Growing evidence indicates that IBs are the site where RNA viruses synthesize their viral RNA. However, in the case of MuV infection, little is known about the viral and cellular compositions and biological functions of the IBs. In this study, pulldown purification and N-terminal amino acid sequencing revealed that stress inducible heat shock protein 70 (Hsp72) was a binding partner of MuV phosphoprotein (P protein), which was an essential component of the IBs formation. Immunofluorescence and immunoblotting analyses revealed that Hsp72 was colocalized with the P protein in the IBs, and its expression was increased during MuV infection. Knockdown of Hsp72 using siRNAs had little, if any, effect on the viral propagation in cultured cells. Knockdown of Hsp72 caused accumulation of ubiquitinated P protein and delayed the P protein degradation. These results show that Hsp72 is recruited to IBs and regulates degradation of MuV P protein through the ubiquitin-proteasome pathway.

Importance

Formation of cytoplasmic inclusion bodies (IBs) is a common characteristic feature in mononegavirus infections. IBs are considered to be the site of viral RNA replication and transcription. However, there have been few studies focused on host factors recruited to the IBs and their biological functions. Here, we identified stress inducible heat shock protein 70 (Hsp72) as the first cellular partner of mumps virus (MuV) phosphoprotein (P), which is an essential component of the IBs and involved in viral RNA replication/transcription. We found that the Hsp72 mobilized to the IBs promoted degradation of the MuV P protein thorough the ubiquitin-proteasome pathway. Our data provide new insight into the role played by IBs in mononegavirus infection.

59

60 **Introduction**

61 One of the characteristic features of mononegavirus infection is formation of
62 cytoplasmic inclusion bodies (IBs), which can be observed by light microscopy (1),
63 fluorescence microscopy (2-6) and electron microscopy (7-9). It is well known that IBs
64 contain nucleocapsid-like structures, but the detailed compositions and biological
65 functions of IBs remain to be elucidated. In the case of Ebolavirus (family *Filoviridae,*
66 order *Mononegavirales*), the IBs have been reported to be the site of viral RNA
67 replication (6). Similar findings were also reported for other RNA viruses, including
68 rabies virus (RV) (4) and vesicular stomatitis virus (VSV) (5) (both of the family
69 *Rhabdoviridae,* order *Mononegavirales*). In regard to the significance of IB formation, it is
70 currently considered that the IBs concentrate the machinery for viral RNA synthesis. In
71 the present study we studied mumps virus (MuV), which is also known to form IBs.
72 MuV is the causative agent of mumps, a common childhood illness characterized by
73 fever and swelling of the salivary glands (10). It often causes neurological
74 complications, including aseptic meningitis, encephalitis, and deafness. MuV belongs
75 to the genus *Rubulavirus* within the family *Paramyxoviridae* (order *Mononegavirales*) (11).
76 The viral non-segmented negative strand RNA genome encodes eight viral proteins:
77 the nucleocapsid (N), V, phospho- (P), matrix (M), haemagglutinin-neuraminidase
78 (HN), fusion (F), large (L), and small hydrophobic (SH) proteins. The genome is
79 encapsidated by the N protein and forms an active template for RNA replication and
80 transcription, a viral ribonucleoprotein (vRNP), with viral polymerases composed of
81 the P and L proteins (12). The F and HN proteins are envelope glycoproteins, and the
82 M protein is an intra-virion protein that associates with the cytoplasmic tails of the
83 envelope glycoproteins and vRNP. The SH protein is also a structural integral
84 membrane protein with unknown function. The V protein is a nonstructural protein
85 that counteracts the host antiviral responses. The V and P proteins are encoded in the
86 same gene using overlapping reading frames. The V protein is translated from the V
87 mRNA, a faithful transcript of the V/P gene, whereas the P protein is translated from
88 the P mRNA possessing two additional nontemplated guanine residues inserted by an
89 RNA editing mechanism. Therefore, the resulting P and V proteins have an identical
90 N-terminal region and unique C-terminal regions.
91 Heat shock protein 70 (Hsp70) family proteins are molecular chaperones that

92 comprise a set of abundant cellular machines (13). Under normal unstressed conditions,
93 Hsp70 proteins play central roles in protein homeostasis, such as assisting in the
94 folding or assembly of newly translated proteins, guiding the intracellular trafficking
95 of client proteins, disassembling oligomeric protein structures, and facilitating the
96 proteolytic degradation of unstable proteins. Under conditions of stress, they prevent
97 abnormal protein aggregation and assist in the renaturation or degradation of
98 misfolded proteins. Human Hsp70 proteins are comprised of at least eight gene
99 products with different amino acid sequences, expression levels and subcellular
100 localizations (14). Among the major Hsp70 proteins expressed at high levels in a wide
101 range of tissues, stress inducible heat shock protein 70 (Hsp72) and constitutively
102 expressed heat shock cognate protein 70 (Hsc70) are present in the cytoplasm and
103 nucleus, while glucose-regulated protein (GRP78) and GRP75 are localized in the
104 lumen of the ER and the mitochondrial matrix, respectively. During virus infections,
105 Hsp70 family proteins are frequently mobilized to the viral replication sites and play
106 roles in all steps of the life cycle of many DNA and RNA viruses (4, 15-17). In the case
107 of negative-stranded RNA viruses, Hsp70 has been found to have both positive and
108 negative regulatory effects on viral propagation. For example, Hsp72 interacts with the
109 N protein of measles virus (MV), which is another member of the paramyxovirus
110 family, and enhances viral RNA replication (18). Hsp70 is also recruited to the IBs of
111 RV and positively regulates RV infection (19). On the other hand, Hsp70 interferes with
112 the polymerase activity of influenza virus and negatively regulates viral RNA
113 replication (20, 21), highlighting the complexity of the virus-chaperone interaction.
114 Our data revealed that MuV-infected cells also recruited Hsp72 to the IBs. In the
115 present study, we analyzed the molecular basis and significance of this event in
116 MuV-infected cells.
117
118 **Materials & Methods**
119 **Cells and virus.** Vero (African green monkey kidney), 293T (human kidney) and Huh7
120 (human hepatocellular carcinoma) cells were maintained in Dulbecco's modified
121 Eagle's minimal essential medium (DMEM) (Nacalai Tesque, Kyoto, Japan)
122 supplemented with 100 U/ml penicillin, 100 mg/ml streptomycin and 10% fetal
123 bovine serum (FBS).
124 The highly neuropathogenic strain MuV Odate strain was isolated from a patient who

125 developed aseptic meningitis (22) and used in this study.
126 **Plasmids.** The cDNA of the P protein was amplified from 293T cells infected with MuV
127 Odate by reverse transcription-PCR (RT-PCR) and cloned into pCAGGS,
128 pCAGPM-N-HA and pCAG-MCS2-FOS for expression in mammalian cells as
129 non-tagged, HA-tagged and FLAG-One-STrEP (FOS)-tagged proteins, respectively (23).
130 The resulting plasmids were designated pCAGGS-P, pCAGPM-HA-P and
131 pCAG-P-FOS, respectively. The cDNAs of the N, L and V proteins were also amplified
132 and cloned into pCAGGS and/or pCAG-MCS2-FOS, resulting in pCAGGS-N, L and V
133 and pCAG-N-FOS, respectively. The cDNAs of human Hsp72, Hsc70, GRP78 and
134 ubiquitin were amplified from 293T cells by RT-PCR and cloned into pcDNA3.1-FLAG
135 or pCAGPM-N-HA for expression in mammalian cells as a FLAG- or HA-tagged
136 protein. The resulting plasmids were designated pcDNA-FLAG-Hsp72,
137 pcDNA-FLAG-Hsc70, pcDNA-FLAG-GRP78 and pCAGPM-HA-Ub, respectively. A
138 series of deletion mutants of the P protein and Hsp72 was generated by PCR-based
139 mutagenesis. All plasmids were confirmed by sequencing with an ABI Prism 3130*xl*
140 genetic analyzer (Life Technologies Inc., Rockville, MD).
141 **Reagents and antibodies.** MG-132 and cycloheximide (CHX) were purchased from
142 Cell Signaling Technology (Danvers, MA) and Sigma (St. Louis, MO), respectively.
143 Lactacystin and epoxomicin were purchased from Peptide Institute Inc. (Osaka, Japan).
144 Anti-N (23D), P (57A), M (79D), F (170C) and HN (78) mouse monoclonal antibodies
145 (MAbs) and anti-MuV V (T60), V/P (T61) and L (L17) rabbit polyclonal antibodies
146 (PAbs) were prepared as described previously (24-26). Anti-MuV N rabbit PAb was
147 generated with a synthetic peptide derived from the MuV N protein at Sigma.
148 Anti-FLAG (M2) and anti-α-tubulin mouse MAbs were purchased from Sigma.
149 Anti-Hsp70 (C92F3A-5) and anti-Hsc70 (1F2-H5) mouse MAbs were purchased from
150 StressMarq Bioscience Inc. (Victoria, Canada). Anti-HA mouse MAb (HA11),
151 anti-GRP78 rabbit PAb (ab21685), and anti-ubiquitin rabbit PAb (#3933) were
152 purchased from Covance (Richmond, CA), Abcam (Cambridge, United Kingdom), and
153 Cell Signaling Technology, respectively.
154 **Virus titration.** Virus titers were determined by plaque assay in triplicate using Vero
155 cells in 12-well plates. After 1 to 2 hr of virus adsorption, the cells were cultured in
156 DMEM with 5% FBS and 1% agarose. At 6 days post-inoculation, the cells were stained
157 with Neutral Red Solution (Sigma), and the plaque counts were determined.

158 **Cell extracts, immunoblotting and immunoprecipitation.** For the preparation of cell
159 extracts, cells were washed twice with cold phosphate-buffered saline (PBS) and then
160 lysed in cell lysis buffer (20 mM Tris-HCl, pH 7.5, 135 mM NaCl, 1% Triton-X 100, and
161 protease inhibitor cocktail [Complete Mini; Roche, Mannheim, Germany]). For
162 immunoblotting, the cell lysate was boiled in sodium dodecyl sulfate (SDS) sample
163 buffer and subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The
164 proteins were transferred to polyvinylidene difluoride membranes (Millipore, Bedford,
165 MA) and incubated with the appropriate antibodies. Each protein was visualized with
166 SuperSignal West Femto Maximum Sensitivity Substrate (Life Technologies Inc.) and
167 detected by use of an LAS-3000 image analyzer system (Fuji Film, Tokyo, Japan). For
168 immunoprecipitation, the cell lysate was pre-cleaned with protein G-sepharose (GE
169 Healthcare, Buckinghamshire, United Kingdom). Antibody-protein complexes were
170 purified with protein G beads and washed with cell lysis buffer three times. After
171 boiling in SDS sample buffer, the proteins were separated by SDS-PAGE and processed
172 for immunoblotting.
173 **Immunofluorescence microscopy.** Vero cells were fixed in 4% paraformaldehyde in
174 PBS for 15 min at room temperature. Then, the cells were permeabilized with 0.2%
175 Triton X-100 in PBS for 10 min, blocked with PBS containing 2% bovine serum albumin
176 (BSA) for 30 min at room temperature, and incubated with the appropriate antibodies.
177 Nuclei were stained with 4', 6-diamidino-2-phenylindole (DAPI). The samples were
178 examined under an FV1000D confocal laser-scanning microscope (Olympus, Tokyo,
179 Japan).
180 **Fluorescence *in situ* hybridization.** MuV genomic RNA was detected using the
181 QuantiGene ViewRNA ISH cell assay kit (Affymetrix, Santa Clara, CA) and mumps
182 virus probe set (Affymetrix) that can hybridize to nucleotides 13,501-14,518 of the MuV
183 genome. Vero cells infected with MuV were fixed in 4% paraformaldehyde in PBS for
184 30 min at room temperature. Then, cells were permeabilized and hybridized according
185 to the manufacturer's protocol. The samples were examined under an FV1000D
186 confocal laser-scanning microscope.
187 **FOS-tagged purification and N-terminal amino acid sequencing.** pCAG-N-FOS,
188 pCAG-P-FOS or empty vector was transfected into 293T cells by use of TransIT LT1
189 (Mirus, Madison, WI), harvested at 24 hr post-transfection, washed twice with ice-cold
190 PBS, suspended in cell lysis buffer, and centrifuged at 14,000 x *g* for 20 min at 4 °C. The

supernatant was pulled down using 50 µl of STrEP-Tactin Sepharose (IBA, Gottingen, Germany) equilibrated with cell lysis buffer for 2 hr at 4 °C. The affinity beads were washed three times with cell lysis buffer and suspended in 2 x SDS-PAGE sample buffer. The proteins were subjected to SDS-PAGE and transferred to membrane, followed by Coomassie brilliant blue (CBB) staining using CBB Stain One (Nakalai Tesque). Each band was spliced out and subjected to N-terminal amino acid sequencing (Procise 491cLC; Applied Biosystems).

Gene silencing. Commercially available small interfering RNA (siRNA) pool targeting Hsp72 (siGENOME SMARTpool, human Hsp72) and control nontargeting siRNA were purchased from Dharmacon (Buckinghamshire, United Kingdom) and transfected using Lipofectamine RNAiMAX (Life Technologies Inc.) according to the manufacturer's protocol.

Quantitative RT-PCR (qRT-PCR). Total RNA was prepared by use of an RNeasy Mini Kit (Qiagen), and first-strand cDNA was synthesized using PrimeScript II RTase and an oligo(dT) primer (Takara Bio, Shiga, Japan). The amount of each cDNA was measured using the Universal ProbeLibrary and the LightCycler 480 system (Roche) according to the manufacturer's instructions. Primers for qRT-PCR were designed by using the Probe Finder software (Roche). The value of each RNA was normalized to that of hypoxanthine phosphoribosyltransferase 1 (HPRT1) mRNA.

TUNEL staining. Vero cells were fixed in 4% paraformaldehyde in PBS for 15 min at room temperature. Then, the cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min at room temperature and incubated with terminal deoxynucleotidyltransferase-mediated dUTP nick end labeling (TUNEL) reaction mixture (*In situ* Apoptosis Detection Kit, Takara Bio) for 90 min at 37°C. Nuclei were stained with DAPI. The samples were examined under a BZ-8000 fluorescence microscope (Keyence Co., Osaka, Japan).

Caspase activity and cell viability assays. Caspase 3/7 activity in Vero cells in 96-well plates was measured by using a Caspase-Glo 3/7 Assay Kit (Promega, Madison, WI) according to the manufacturer's protocol. Cell viability was measured by using a CellTiter-Glo Luminescent Cell Viability Assay Kit (Promega) and used for normalization.

Results

Coexpression of the MuV N and P proteins induces formation of an IB-like structure, where these proteins were concentrated. A previously reported analysis using electron microscopy suggested that the IBs observed in MuV-infected cells are aggregates of nucleocapsids (7). In order to identify viral components of the MuV-induced IBs, the intracellular localizations of MuV proteins and genomic RNA were analyzed by immunofluorescence microscopy and fluorescence *in situ* hybridization (Fig. 1A and B). The N, V, P, and L proteins were localized mainly to the IBs as well as viral genomic RNA, while the M protein was detected not only in the IBs but also in the nucleoli and the pericellular region. The F and HN proteins were localized mainly in the pericellular region. To further clarify the intracellular localizations of the vRNP components, the N, P and L proteins were expressed in cells alone or in combination using expression plasmids. Each protein showed a diffuse distribution pattern throughout the cytoplasm, when expressed alone (Fig. 1C). Co-expression of the N and P proteins but not other combinations led to the formation of IB-like structures, where these proteins were concentrated (Fig. 1D).

MuV P protein associates with Hsp70 family proteins. Data by immunofluorescence assay showed that the N, P, V, M, and L proteins were concentrated in the IBs, and co-expression of the N and P proteins induced the formation of IB-like structures. As an initial step in the search for host factors involved in the formation of IBs, host proteins associated with the N and P proteins were analyzed by an FOS affinity tag purification method (Fig. 2A). Co-expression of N- and P-FOS proteins led to the formation of IB-like structures similar to the untagged N and P proteins, indicating that placement of the FOS tag at C-terminus of N and P protein had minimal effects on the function of IB-like structure formation (Fig. 2B). The N- and P-FOS proteins were expressed in 293T cells and purified together with associated proteins. Several polypeptides, including three polypeptides with molecular weights (MWs) of ~72, 73 and 78 kDa, were co-purified with the P protein, whereas no N protein-associated host proteins were detected (Fig. 2C). Analysis by immunoblotting confirmed that the purified P-FOS protein was observed at an MW of ~45 kDa, and suggested that several other bands were different forms of the P protein, seemingly correspondent to phosphorylated and cleaved P-FOS products (Fig. 2D). In order to identify the three polypeptides with MWs of ~72, 73 and 78 kDa, individual bands were isolated and

257 analyzed by N-terminal amino acid sequencing. The 78 kDa polypeptide was identified
258 as GRP78 with a sequence of EEEDKKEDVG (residues 19 to 28), whereas the amino
259 acid sequences of the other two polypeptides of 72 and 73 kDa were not identified by
260 this assay. It was postulated that they could be Hsp72 and Hsc70, because Hsp70
261 family proteins have conserved domains and the MWs of Hsp72 and Hsc70 are 72 kDa
262 and 73 kDa, respectively. We investigated this possibility by immunoblotting using
263 specific antibodies against Hsp72, Hsc70 and GRP78. As shown in Fig. 2E, Hsp72 and
264 Hsc70 as well as GRP78 were clearly detected. These data thus showed that all three
265 Hsp70 family proteins were associated with the P protein.
266
267 **Hsp72 is up-regulated and recruited to IBs during MuV infection.** The interaction
268 between the P protein and the three Hsp70 family proteins (Fig. 3A) was analyzed by
269 co-immunoprecipitation assay. 293T cells expressing the HA-tagged P protein (HA-P)
270 and FLAG-tagged Hsp70 family proteins were used. As shown in Fig. 3B, massive
271 amount of FLAG-Hsp72 was co-immunoprecipitated with HA-P, whereas only a small
272 amount of FLAG-GRP78 was co-precipitated. Although the binding affinities between
273 the P protein and Hsp70 family proteins might be different, all three Hsp70 family
274 proteins were indeed capable of associating with the P protein in cell lysates. However,
275 it was still unclear whether all three Hsp70 family proteins interact with the P protein
276 in living cells. To clarify this point, we investigated the intracellular localizations of
277 Hsp70 family proteins in MuV-infected cells by immunofluorescence microscopy. The
278 results showed that Hsp72 expression was upregulated, and the protein was
279 redistributed to the IBs in the MuV-infected cells (Fig. 3C). On the other hand, the
280 expression levels and localization of Hsc70 and GRP78 were unchanged (Fig. 3C).
281 These results suggested that Hsp72, but not Hsc70 or GRP78, interacts with the P
282 protein in MuV-infected cells. Data by immunoblotting assay also demonstrated that
283 the expression of Hsp72 was increased by MuV infection (Fig. 3D). To better
284 understand the requirements for the up-regulation and localization change of Hsp72,
285 the P protein was expressed in cells alone or in combination with the N protein. While
286 ectopic expression of the P protein alone did not induce the up-regulation of Hsp72
287 and was not co-localized with Hsp72 (Fig. 3E and F), the formation of IB-like structures
288 caused by the co-expression of N and P proteins dramatically induced the expression
289 of Hsp72 and recruited Hsp72 to the IB-like structures (Fig. 3G). Taken together, the

up-regulation and recruitment to IBs of Hsp72 occurred with the IB formation during MuV infection.

The N-terminal region of the P protein and the C-terminal region of Hsp72 are responsible for their interaction. To determine the interacting regions of Hsp72 and the P protein, N-terminally FLAG-tagged Hsp72 (FLAG-Hsp72F), HA-tagged P protein (HA-P-Full) and their truncation polypeptides were expressed in cells, and their interaction was analyzed by co-immunoprecipitation assays. FLAG-Hsp72N was comprised of the N-terminal ATPase domain, while FLAG-Hsp72C was comprised of the C-terminal peptide binding domain and variable region (Fig. 4A). HA-P-Full was co-precipitated with FLAG-Hsp72F and FLAG-Hsp72C (Fig. 4A), indicating that the C-terminal region of Hsp72 interacted with the P protein. HA-PΔN was comprised of the oligomerization domain and C-terminal region, while HA-PΔC was comprised of the N-terminal region and oligomerization domain (Fig. 4B). The co-immunoprecipitation assay showed that FLAG-Hsp72F was co-precipitated with HA-P-Full and HA-PΔC (Fig. 4B). The data, taken together, indicated that the N-terminal region of the P protein and the C-terminal region of Hsp72 were responsible for their interaction. A co-immunoprecipitation assay was also performed for the V protein, as the P and V proteins possess the common N-terminal region. The data showed that Hsp72 was also associated with the V protein (Fig. 4C).

Hsp72 is nonessential for MuV replication, but suppresses apoptotic cell death of MuV-infected cells. To determine the roles for Hsp72 in MuV infection, the expression of Hsp72 was suppressed by Hsp72-specific siRNAs (siHsp72). Transfection of siHsp72 efficiently knocked down the Hsp72 expression (Fig. 5A). However, the levels of viral RNAs, virus production and IB formation were not affected (Fig. 5B-D), demonstrating that Hsp72 was nonessential for MuV replication. Alternatively, it was noted that ~10% of Hsp72-knockdown cells infected with MuV were positive for the TUNEL stain indicating apoptosis induction while only ~2.5% of control cells were TUNEL positive (Fig. 6A and B). The induction of apoptosis in Hsp72-knockdown cells was confirmed by the elevated caspase 3/7 activity (Fig. 6C). These data indicated that Hsp72 was needed to suppress apoptotic cell death of MuV-infected cells.

The P and V proteins were ubiquitinated in MuV-infected cells. Since IBs are comprised of abundant viral proteins that could lead to deleterious consequences for the cells, Hsp72 may play a role in ubiquitin-mediated degradation of the accumulated viral proteins. To examine this possibility, HA-Ub-expressing 293T cells were infected with MuV, and subjected to immunoprecipitation assays. Polypeptides were immunoprecipitated with the anti-MuV V/P antibody (T61), and detected by immunoblotting using an anti-HA antibody. They showed broad size distributions (Fig. 7A). The signals of ubiquitinated proteins were increased in the presence of MG132, a proteasomal inhibitor. Since the T61 antibody detects both the P and V proteins, these data suggested that the V protein, the P protein, or both were ubiquitinated in MuV-infected cells. Furthermore, endogenous ubiquitinated P and V proteins were detected in Vero cells infected with MuV (Fig. 7B). Taken together, these results demonstrated that both the P and V proteins were ubiquitinated in MuV-infected cells.

Hsp72 targets the P protein for degradation through the ubiquitin-proteasome pathway. To investigate the roles of Hsp72 in ubiquitin-mediated degradation of the P and V proteins, the effects of Hsp72 knockdown were analyzed. 293T cells expressing HA-Ub were transfected with either siHsp72 or control siNC. The cells were then infected with MuV, and subjected to immunoprecipitation assays, in which the polypeptides were immunoprecipitated with the T61 antibody and detected by immunoblotting using an anti-HA antibody. The signals were clearly higher in the Hsp72-knockdown cells than in control cells (Fig. 8A). In fact, the signal levels were as high as those in MG-132-treated cells (Fig. 8A). These data suggested that the ubiquitinated V protein, P protein, or both were accumulated in Hsp72-knockdown cells. Similar experiments were performed using cells expressing the V or P protein individually. Signals for the ubiquitinated P protein were low in control (siNC-transfected) cells, but increased in Hsp72-knockdown (siHsp72-transfected) cells (Fig. 8B). On the other hand, signals for the ubiquitinated V protein were similar between Hsp72-knockdown and control cells (Fig. 8B). These data suggested that Hsp72 was involved in the P protein degradation, but not in the V protein degradation. Next, the kinetics of the P and V protein degradation was analyzed. Hsp72-knockdown and control Vero cells were infected with MuV, and cultured for 24 hr. Then, these cells were incubated for 0 to 12 hr in the presence of CHX. Degradation of the P protein was

suppressed in Hsp72-knockdown cells (Fig. 8C, DMSO-treated lanes). On the other hand, degradation of the V protein was minimally affected. To further confirm the proteasomal degradation of P protein, we treated the MuV-infected cells with specific proteasome inhibitors, lactacystin and epoxomicin, which have higher degree of specificity than MG-132 (27). Degradation of the P and V proteins was blocked by treatment of proteasome inhibitors (Fig. 8C). Taken together, these data show that Hsp72 binds to both the V and P proteins, but specifically promotes proteasomal degradation of the P protein.

Discussion

Many RNA viruses form IBs. However, the precise functions and complete compositions of IBs remain to be elucidated. Studies of negative-stranded RNA viruses have reported that the viral genomic RNA and mRNA were present along with the machineries for viral RNA synthesis in IBs (3-6). Therefore, IBs are likely the site of viral RNA replication and transcription (3-6). Also, in the case of MuV infection the vRNP components were concentrated in the IBs (7). This compartmentalization may facilitate virus replication. In addition to viral proteins, host factors involved in innate immune responses are localized to the IBs. This is thought to be a virus strategy to sequester cellular detectors of viral infections (3, 28).

Our data demonstrated that Hsp72 was recruited to the IBs and interacted with the P protein during MuV infection. Many viruses use cellular chaperones for their genome replication, protein synthesis and virion assembly (29). In the paramyxovirus family, Hsp72 has been shown to enhance MV RNA replication and transcription through the interaction with the C terminal region of N protein (18). Hsp72 also associates with polymerase complexes of respiratory syncytial virus (RSV) to positively affect viral RNA synthesis (30). Therefore, it was possible that MuV actively used Hsp72 for its replication. However, we considered this unlikely, since knockdown of Hsp72 showed little, if any, effect on MuV propagation in cultured cells. Recent study revealed that MuV P protein forms a unique tetramer structure different from other paramyxovirus (31, 32). Thus, MuV might differ in the requirements for viral RNA replication and propagation from MV and RSV. Further investigation will be required to define the roles of Hsp72 in MuV infection and to explain these differences.

Abnormally accumulated proteins disrupt the cellular functions and implicate in