OPEN ACCESS Freely available online

PLoS PATHOGENS

Opinion

Systematic Genetic Nomenclature for Type VII Secretion Systems

Wilbert Bitter[1]*, Edith N. G. Houben[1], Daria Bottai[2], Priscille Brodin[3], Eric J. Brown[4], Jeffery S. Cox[5], Keith Derbyshire[6], Sarah M. Fortune[7], Lian-Yong Gao[8], Jun Liu[9], Nicolaas C. Gey van Pittius[10], Alexander S. Pym[11], Eric J. Rubin[7], David R. Sherman[12], Stewart T. Cole[13], Roland Brosch[14]*

1 VU University Medical Centre, Amsterdam, The Netherlands, **2** Dipartimento di Patologia Sperimentale, Biotecnologie Mediche, Infettivologia ed Epidemiologia, University of Pisa, Pisa, Italy, **3** Inserm Avenir Group, Institut Pasteur Korea, Seoul, Korea, **4** Department of Microbial Pathogenesis, Genentech Inc., San Francisco, California, United States of America, **5** Department of Microbiology and Immunology, University of California, San Francisco, California, United States of America, **6** Wadsworth Center, New York State Department of Health, Albany, New York, United States of America, **7** Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts, United States of America, **8** Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland, United States of America, **9** Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada, **10** Department of Biomedical Sciences, Stellenbosch University, Stellenbosch, South Africa, **11** Unit for Clinical and Biomedical TB Research, South African MRC, Durban, South Africa, **12** Seattle Biomedical Research Institute, Seattle, Washington, United States of America, **13** Global Health Institute, EPFL, Lausanne, Switzerland, **14** Institut Pasteur, Integrated Mycobacterial Pathogenomics, Paris, France

Mycobacteria, such as the etiological agent of human tuberculosis, *Mycobacterium tuberculosis*, are protected by an impermeable cell envelope composed of an inner cytoplasmic membrane, a peptidoglycan layer, an arabinogalactan layer, and an outer membrane. This second membrane consists of covalently linked, tightly packed long-chain mycolic acids [1,2] and non-covalently bound shorter lipids involved in pathogenicity [3–5]. To ensure protein transport across this complex cell envelope, mycobacteria use various secretion pathways, such as the SecA1-mediated general secretory pathway [6,7], an alternative SecA2-operated pathway [8], a twin-arginine translocation system [9,10], and a specialized secretion pathway variously named ESAT-6-, SNM-, ESX-, or type VII secretion [11–16]. The latter pathway, hereafter referred to as type VII secretion (T7S), has recently become a large and competitive research topic that is closely linked to studies of host–pathogen interactions of *M. tuberculosis* [17] and other pathogenic mycobacteria [16]. Molecular details are just beginning to be revealed [18–22] showing that T7S systems are complex machineries with multiple components and multiple substrates. Despite their biological importance, there has been a lack of a clear naming policy for the components and substrates of these systems. As there are multiple paralogous T7S systems within the Mycobacteria and orthologous systems in related bacteria, we are concerned that, without a unified nomenclature system, a multitude of redundant and obscure gene names will be used that will inevitably lead to confusion and hinder future progress. In this opinion piece we will therefore propose and introduce a systematic nomenclature with guidelines for name selection of new components that will greatly facilitate communication and understanding in this rapidly developing field of research.

The first T7S-associated protein to be identified was the 6-kD early secreted antigenic target ESAT-6 [23]. This small, highly immunogenic protein lacks a classical N-terminal signal sequence and is present in large amounts in the culture filtrate of *M. tuberculosis* [23], but is missing from the closely related attenuated live vaccine *Mycobacterium bovis* bacille Calmette-Guérin (BCG) [24] due to the deletion of region of difference 1 (RD1) [25]. ESAT-6 and its protein partner, the 10-kD culture filtrate protein CFP-10 [26], form a 1:1 protein complex [27] that involves hydrophobic interaction [18,28]. Secretion of ESAT-6 and CFP-10 is required for the pathogenicity of *M. tuberculosis* [29–31]. The absence of ESAT-6 secretion is responsible in part for the attenuation of the BCG and *Mycobacterium microti* vaccines [13,32,33], as well as for the decrease in virulence of the attenuated *M. tuberculosis* H37Ra strain [34].

In *M. tuberculosis*, ESAT-6 and CFP-10 belong to the WXG100 family of 23 small secreted proteins that share a size of approximately 100 amino acids, a helical structure, and a characteristic hairpin bend formed by the conserved Trp-Xaa-Gly (W-X-G) motif [35]. The genes encoding these proteins, many of which represent immunodominant T cell antigens [36], are called *esx* genes in *M. tuberculosis* (*esxA-W*, Table 1) and are arranged in tandem pairs at 11 genomic loci [37]. In five of these genomic loci (ESX-1–ESX-5), the *esx* genes are flanked by genes coding for components of secretion machineries involved in the export of the corresponding ESX proteins (Figure 1). These proteins constitute the major building blocks of the T7S systems [11,12,15,16,19]. Four of these regions are also characterized by the presence of genes encoding PE and/or PPE proteins (Figure 1, Table 2), named after their characteristic N-terminal motifs proline-glutamic acid (PE) and proline-proline-glutamic acid (PPE) [38]. Apart from genes localized in these core ESX regions, additional genes situated elsewhere on the chromosome may be required for the function of T7S systems. For example, the *rv3616c-rv3614c* genes are required for

Citation: Bitter W, Houben ENG, Bottai D, Brodin P, Brown EJ, et al. (2009) Systematic Genetic Nomenclature for Type VII Secretion Systems. PLoS Pathog 5(10): e1000507. doi:10.1371/journal.ppat.1000507

Editor: Glenn F. Rall, The Fox Chase Cancer Center, United States of America

Published October 30, 2009

Funding: This work has received funding from the European Community's Seventh Framework Programme ([FP7/2007–2013]) under grant agreement n°201762. In addition, the authors acknowledge their individual external funding for T7S/ESX related research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: w.bitter@vumc.nl (WB); roland.brosch@pasteur.fr (RB)

Table 1. Overview of *esx* Genes (WXG100 Family) of *M. tuberculosis* H37Rv, Also Showing Previously Used Gene Names in Brackets.

Gene Family	ESX-1	ESX-2	ESX-3	ESX-4	ESX-5	No Similarity To Cluster
ESAT-6	***esxA*** *(esat-6, rv3875)*	***esxC*** *(rv3890c)*	***esxH*** *(cfp7, tb10.4, rv0288)*	***esxT*** *(rv3444c)*	***esxN*** *(mtb9.9A, Rv1793)*	
CFP-10	***esxB*** *(lhp, cfp-10, rv3874)*	***esxD*** *(rv3891c)*	***esxG*** *(tb9.8, rv0287)*	***esxU*** *(rv3445c)*	***esxM*** *(tb11.0, rv1792)*	
ESAT-6 homologues elsewhere in the genome			***esxR*** *(tb10.3, rv3019c)*, ***esxQ*** *(tb.9, rv3017c)*		***esxI*** *(mtb9.9D, rv1037c)*, ***esxL*** *(mtb9.9C, rv1198)*, ***esxO*** *(mtb9.9E, rv2346c)*, ***esxV*** *(mtb9.9D, rv3619c)*	***esxE*** *(rv3904c)*
CFP-10 homologues elsewhere in the genome			***esxS*** *(rv3020c)*		***esxJ*** *(tb11.0, Rv1038c)*, ***esxK*** *(tb11.0, Rv1197)*, ***esxP*** *(rv2347c)*, ***esxW*** *(rv3620c)*	***esxF*** *(rv3905c)*

doi:10.1371/journal.ppat.1000507.t001

secretion of ESAT-6 and CFP-10 by ESX-1 [39–41].

Apart from members of the *M. tuberculosis* complex, the ESX-1 cluster is also present in a range of mycobacteria, including *Mycobacterium kansasii* [23] and *Mycobacterium leprae* [42]. However, experimental work has mainly focused on the ESX-1 system of *Mycobacterium marinum* [21,22,43–47], a fish pathogen that shows high homology in its ESX loci with *M. tuberculosis* [48], and the fast grower *Mycobacterium smegmatis* [49–51]. *M. marinum* has also been used to define a role for the paralogous system ESX-5, which is required for the secretion of PE and PPE proteins [16,52,53]. For the remaining ESX-2, ESX-3, and ESX-4 systems, only very limited predictions of their putative functions can be made. ESX-3 transcriptome data suggest that this system is involved in iron/zinc homeostasis [54,55], which would be consistent with the essential role of ESX-3 in *M. tuberculosis* [56]. The putative functions of ESX-2 and ESX-4 remain unknown. ESX-4, which harbors a smaller number of genes than other ESX loci (Table 2), appears to represent the most ancestral T7S system in mycobacteria [12]. This hypothesis is based on the observation that ESX-4-like loci are the only ESX clusters that are found in other high GC Gram-positive bacteria, suggesting that the last common ancestor of mycobacteria already harbored an ESX-4 T7S system. Other ESX clusters may have evolved later by gene duplication and gene diversification events. However, the finding that *Nocardia farcinica* (http://nocardia.nih.go.jp/) contains two T7S systems, one orthologous to ESX-4 and one locus that shows some similarity to all the conserved components of larger T7S systems, suggests that evolution of T7S systems is more complex than previously anticipated. This second T7S locus in *N. farcinica* even contains two PPE-like genes that were originally thought to be specific for the mycobacteria [38].

T7S-like systems are also found outside the high GC Gram-positive bacteria, since a number of Firmicutes have WXG100 members [35]. However, the loci containing these WXG100 genes are only weakly similar to the mycobacterial T7S systems: in fact, only the gene encoding the FtsK/SpoIIIE-like protein is present. Therefore, these systems should be called WXG100 systems to differentiate them from true T7S systems. Both *Staphylococcus aureus* and *Bacillus anthracis* have an active WXG100 system, and the WXG100 system encoded by *S. aureus* is important for virulence [57,58].

Research in the T7S/ESX field is relatively new, but is now rapidly expanding and we therefore would like to propose a systematic nomenclature for all components involved. Until now a small number of genes within the different ESX loci of mycobacteria have been named, but for most genes the original genome annotation numbers are used. These gene numbers vary between different species and even between different strains of the same species, and therefore make comparative studies confusing. Our nomenclature is appropriate for all T7S systems in high GC-Gram-positive species. Extending this nomenclature to the T7S-like systems of Firmicutes is not recommended, since there are only a very few conserved components.

As a starting point for the new nomenclature, we focus on the most studied system, the ESX-1 system of *M. tuberculosis*, which is the paradigm T7S system. The new nomenclature is given for ESX-1 in *M. tuberculosis* (Figure 1 and Table 2) and for all ESX systems in various Mycobacteria (Table S1). The proposed rules for the nomenclature are as follows:

- Only genes that have homologues in at least four of the mycobacterial ESX systems will get a general name, whereas the locus-specific genes have a more restricted name reflecting their specificity. The reason for this distinction is that the conserved genes are most likely to represent the core components of the secretion system. Moreover, all of the conserved ESX-1 components have been shown to be essential for ESAT-6/CFP-10 secretion in at least one of the mycobacterial species studied (See below). In contrast, many of the locus-specific genes encode secreted proteins, as has been shown for the ESX-1 system (see below). Furthermore, in *M. leprae*, an organism with an extreme reductive evolution of its genome, almost all of the non-conserved ESX-1 components are pseudogenes, whereas all of the conserved components seem to be intact [42].
- The three letter acronym for the conserved components will be *ecc*, for esx conserved component (Figure 1, Table 2). This abbreviation has not been used for other genes in bacteria.
- The ESAT-6 and CFP-10 encoding genes, *esxA* and *esxB*, respectively, and the other *esx* genes (Table 1) will not be renamed. These gene names are informative, well-accepted, and frequently used in the literature. Furthermore, the *esx* gene products seem to be secreted proteins and do not seem to be components of the secretion system itself, although their presence is required for the secretion of other substrates. The same reasoning is used for the *pe* and *ppe* genes. Four of the five systems harbor *pe* and *ppe* genes, but for the moment their functions within the T7S systems remain uncertain. Furthermore, various mycobacte-

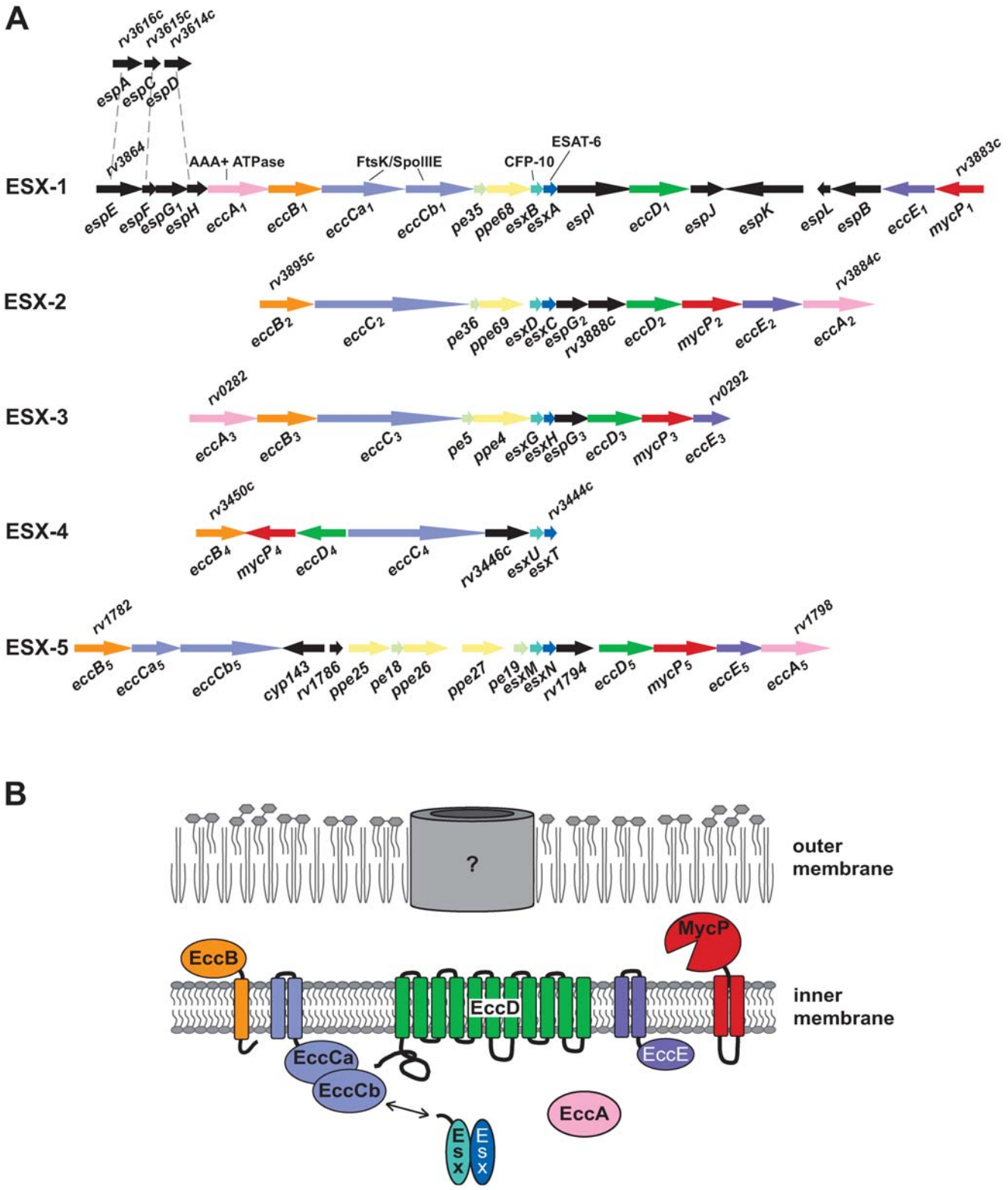

Figure 1. Genetic organization of the 5 ESX loci and the *espA* operon in *M. tuberculosis* H37Rv with the proposed nomenclature and predicted cellular localization of the conserved ESX gene products and their interactions. (A) Genetic organization. (B) Model. The abbreviation *ecc* stands for esx conserved component, whereas *esp* stands for ESX-1 secretion-associated proteins. The topology of the different proteins in the cytoplasmic membrane shown in (B) refers to the ESX-1 cluster and is based on predictions made using the MEMSAT3 algorithm [60]. Note that the channel drawn in the outer membrane of our model refers to a hypothetical pore, whose existence has not been experimentally demonstrated.

doi:10.1371/journal.ppat.1000507.g001

Table 2. New and Old Nomenclature of the Different Esx Conserved Components (*ecc* Genes) and Genes Encoding ESX-1 Secretion-Associated Proteins (*esp* Genes) of the T7S Systems of *M. tuberculosis* H37Rv.

New Gene Name[a]	Putative Function of Gene Products	Previously Proposed Gene Names				
		ESX-1	ESX-2	ESX-3	ESX-4	ESX-5
eccA	AAA+ ATPase	*rv3868*	*rv3884c*	*rv0282*	-	*rv1798*
eccB	Transmembrane protein (1 TM)	*rv3869*	*rv3895c*	*rv0283*	*rv3450c*	*rv1782*
eccC	FtsK/SpoIIIE-like transmembrane protein (1–3 TMs)	-	*rv3894c*	*rv0284*	*rv3447c*	-
eccCa	FtsK/SpoIIIE-like transmembrane protein (1–3 TMs)	*rv3870 snm1*	-	-	-	*rv1783*
eccCb	FtsK/SpoIIIE-like transmembrane protein (1–3 TMs)	*rv3871 snm2*	-	-	-	*rv1784*
eccD	Transmembrane protein (10–11 TMs)	*rv3877 snm4*	*rv3887c*	*Rv0290*	*rv3448c*	*rv1795*
eccE	Transmembrane protein (2 TMs)	*rv3882c*	*rv3885c*	*rv0292*	-	*rv1797*
mycP	Subtilisin-like serine protease (mycosin) (1 TM)	*rv3883c*	*rv3886c*	*rv0291*	*rv3449c*	*rv1796*
espA	Secreted protein	*rv3616c*	-	-	-	-
espB	Secreted protein	*rv3881c*	-	-	-	-
espC	Secreted protein	*rv3615c*	-	-	-	-
espD	Unknown	*rv3614c*	-	-	-	-
espE	Secreted protein	*rv3864*	-	-	-	-
espF	Secreted protein	*rv3865*	-	-	-	-
espG	Soluble protein	*rv3866*	*rv3889c*	*rv0289*	-	-
espH	Unknown	*rv3867*	-	-	-	-
espI	Pro and Ala rich protein	*rv3876 snm3*	-	-	-	-
espJ	Unknown	*rv3878*	-	-	-	-
espK	Pro and Ala rich protein	*rv3879c*	-	-	-	-
espL	Unknown	*rv3880c*	-	-	-	-
espR	Regulation	*rv3849*	-	-	-	-

The number of transmembrane domains varies depending on the prediction programme used (for details see Table S2).
[a]The numeral suffix indicating the ESX cluster to which this gene belongs is not shown in this table.
doi:10.1371/journal.ppat.1000507.t002

rial species contain a large number of genes belonging to the *pe* and *ppe* families, and it would be confusing to rename some of them. Finally, the subtilisin-like proteases already have an established and descriptive name in literature, i.e., the mycosins [59]. Therefore, we will not change this name.

- The alphabetic suffix of conserved genes will be based on the gene order in the paradigm ESX-1 system (see Figure 1). This decision is mainly based on the fact that the ESX-1 system is the most studied. The gene order of the different T7S systems is highly variable and it is therefore difficult to propose a logical ordering that would be satisfactory for all systems. The genes of ESX-2/-3/-4 and -5 will therefore be named according to their paralogue in ESX-1 (Table 2 and Table S1), allowing for a direct and relevant comparison. The gene names of each mycobacterial T7S will include a numeral suffix indicating the ESX cluster to which this gene belongs. In order to avoid confusion with numbering of alleles, the ESX cluster number is indicated in subscript. As shown in Figure 1, the first conserved gene of the ESX-1 cluster will be $eccA_1$.
- In some of the T7S clusters, the gene encoding the FtsK/SpoIIIE-like protein is split in two genes. Since these gene products clearly form a functional unit, as has also been shown for the two FtsK/SpoIIIE-like proteins of the ESX-1 system [14], the split genes will get a lower case alphabetic suffix, i.e., $eccCa_1$ and $eccCb_1$ for the ESX-1 system of *M. tuberculosis* (Figure 1 and Table 2).
- When working with several different organisms, it can also be useful to indicate the origin of the respective genes. For this we recommend using a two-letter subscript at the end of the gene name. For example, the orthologues of the *M. tuberculosis* genes $eccCa_{1mt}$ and $eccCb_{1mt}$ would be $eccCa_{1ms}$ and $eccCb_{1ms}$ in *M. smegmatis*.
- The gene names can be converted into their proteins by capitalization, e.g., $EccCa_1$. Alternatively, once the true function of a protein is known, the name could be changed to indicate this function, as has been done for the secretins of type II and type III secretion systems. If in the future new genes are identified that are essential for the functioning of several T7S systems, these genes could be named similarly using the next alphabetical suffix (*eccG*, *eccH*, etc.).
- As discussed above, in addition to the conserved genes, there are also region-specific genes. The role of these genes in ESAT-6/CFP-10 secretion is not entirely clear: some of the encoded proteins seem to be involved in the secretion of T7S substrates in *M. marinum*, whereas their orthologues show less or no effect on secretion in *M. tuberculosis*. Recently, it has been shown that a subset of these proteins are in fact also substrates of the ESX-1 system. Thus far, four ESX-1 substrates have been identified in addition to ESAT-6 and CFP-10. These substrates are called EspA [39], EspB

[46], EspR [41], and the *M. marinum* homologue of Rv3864 [22]. The acronym Esp stands for ESX-1 secretion-associated protein. Both *rv3864* and *espB* are located within the ESX-1 cluster, whereas EspA and the secreted regulatory protein EspR are encoded by genes outside the ESX-1 locus. However, the *espA* gene is part of an operon (*rv3616-3614*) that has paralogues in the 5′ region of the ESX-1 locus. Therefore, we propose naming all the region-specific genes of the ESX-1 system and the rest of the *espA* operon *esp* genes with alphabetical suffixes (see Table 2 and Figure 1). We will follow the *espA* operon and ESX-1 gene order, with the exception of *espB* and *espR*, which are already named. This means that the first gene in the *esx-1* operon, whose gene product was recently shown to be secreted protein in *M. marinum*, will be named *espE*. One of the new *esp* genes, *espG*, is present with low but significant homology in two other ESX systems (ESX-2 and ESX-3) and should therefore also have a numeral suffix (Figure 1, Table 2).

- The nomenclature of *esp* genes in *M. marinum* is more complicated, in particular for *espA*. The genome of *M. marinum* contains a large gene cluster upstream of the ESX-1 locus, among which are 15 *espA*-like genes [48]. In addition, there are three more paralogues at other locations in the genome. These genes should all be named *espA* with a superscript numeral suffix to indicate the exact gene and a subscript "*mm*" to indicate the species.
- Region-specific genes or genes encoding secreted proteins of the other ESX loci and T7S systems should not be called *esp*, as this name should be reserved for ESX-1 related genes. If there are important region-specific genes for ESX-2/-3/-4 or -5, a new name has to be introduced.

In order to ensure wide visibility for this new nomenclature it will be included in the most extensively used mycobacterial genome databases. As a first step, selected genome browsers available at the Institut Pasteur (http://genolist.pasteur.fr/) and/or the Ecole Polytechnique Federale de Lausanne (http://tuberculist.epfl.ch/) will adopt these new rules; other databases could follow this example.

In conclusion, we would like to emphasize that the introduction of a uniform gene nomenclature for other secretion systems in Gram-negative bacteria (type II, type III) has facilitated comparative analysis of these systems. We anticipate that the acceptance/implementation of this proposal will provide similar advantages for the T7S systems.

Supporting Information

Table S1 New and old nomenclature of the different conserved components of the T7S systems in selected mycobacteria (*M. tuberculosis* H37Rv, *M. marinum* M, *M. smegmatis* mc^{2}155, *M. leprae* TN, *M. avium paratuberculosis* K10). The numeral suffices to indicate the ESX clusters to which the genes belong are omitted. Note that the ESX-2 genes of *M. avium paratuberculosis* are located in two separate genomic loci. TM, transmembrane domain.
Found at: doi:10.1371/journal.ppat.1000507.s001 (0.09 MB DOC)

Table S2 Comparison of the transmembrane topologies and signal sequence predictions of the *M. tuberculosis* H37Rv Ecc membrane proteins. Amongst the different topology prediction programs that were used (TMHMM Server v. 2.0, MEMSAT3, Philius, SCAMPI, HMMTOP and Phobius) MEMSAT3 gave the correct prediction for the highest number of Ecc membrane proteins. Therefore, only the topology prediction results of TMHMM (used on the TubercuList server) and MEMSAT3 are shown. The clearly incorrect predictions are depicted in gray. TM, transmembrane domain; in, cytoplasmic location; out, periplasmic location; C, C-terminus; N, N-terminus; ss, signal sequence.
Found at: doi:10.1371/journal.ppat.1000507.s002 (0.07 MB DOC)

Acknowledgments

We are grateful to our many collaborators and colleagues, who have contributed to the rapid progress in this field.

References

1. Hoffmann C, Leis A, Niederweis M, Plitzko JM, Engelhardt H (2008) Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. Proc Natl Acad Sci U S A 105: 3963–3967.
2. Zuber B, Chami M, Houssin C, Dubochet J, Griffiths G, et al. (2008) Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. J Bacteriol 190: 5672–5680.
3. Camacho LR, Ensergueix D, Perez E, Gicquel B, Guilhot C (1999) Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. Mol Microbiol 34: 257–267.
4. Cox JS, Chen B, McNeil M, Jacobs WR Jr (1999) Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. Nature 402: 79–83.
5. Trivedi OA, Arora P, Vats A, Ansari MZ, Tickoo R, et al. (2005) Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid. Mol Cell 17: 631–643.
6. Pugsley AP (1993) The complete general secretory pathway in gram-negative bacteria. Microbiol Rev 57: 50–108.
7. Wiker HG, Harboe M (1992) The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis*. Microbiol Rev 56: 648–661.
8. Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR Jr (2003) SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol 48: 453–464.
9. McDonough JA, McCann JR, Tekippe EM, Silverman JS, Rigel NW, Braunstein M (2008) Identification of functional Tat signal sequences in *Mycobacterium tuberculosis* proteins. J Bacteriol 190: 6428–6438.
10. Saint-Joanis B, Demangel C, Jackson M, Brodin P, Marsollier L, et al. (2006) Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases beta-lactam susceptibility and virulence. J Bacteriol 188: 6669–6679.
11. Tekaia F, Gordon SV, Garnier T, Brosch R, Barrell BG, et al. (1999) Analysis of the proteome of *Mycobacterium tuberculosis* in silico. Tuber Lung Dis 79: 329–342.
12. Gey Van Pittius NC, Gamieldien J, Hide W, Brown GD, Siezen RJ, et al. (2001) The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. Genome Biol 2: RESEARCH0044.
13. Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, et al. (2003) Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. Nat Med 9: 533–539.
14. Stanley SA, Raghavan S, Hwang WW, Cox JS (2003) Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. Proc Natl Acad Sci U S A 100: 13001–13006.
15. Brodin P, Rosenkrands I, Andersen P, Cole ST, Brosch R (2004) ESAT-6 proteins: protective antigens and virulence factors? Trends Microbiol 12: 500–508.
16. Abdallah A, Gey van Pittius N, Champion P, Cox J, Luirink J, et al. (2007) Type VII secretion–mycobacteria show the way. Nat Rev Microbiol 5: 883–891.
17. Simeone R, Bottai D, Brosch R (2009) ESX/type VII secretion systems and their role in host-pathogen interaction. Curr Opin Microbiol 12: 4–10.
18. Brodin P, de Jonge MI, Majlessi L, Leclerc C, Nilges M, et al. (2005) Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of *Mycobacterium tuberculosis*, reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. J Biol Chem 280: 33953–33959.
19. Champion PA, Stanley SA, Champion MM, Brown EJ, Cox JS (2006) C-terminal signal sequence promotes virulence factor secretion in *Mycobacterium tuberculosis*. Science 313: 1632–1636.
20. de Jonge MI, Pehau-Arnaudet G, Fretz MM, Romain F, Bottai D, et al. (2007) ESAT-6 from *Mycobacterium tuberculosis* dissociates from its putative chaperone CFP-10 under acidic conditions and exhibits membrane-lysing activity. J Bacteriol 189: 6028–6034.
21. Smith J, Manoranjan J, Pan M, Bohsali A, Xu J, et al. (2008) Evidence for pore formation in host cell membranes by ESX-1-secreted ESAT-6 and

its role in *Mycobacterium marinum* escape from the vacuole. Infect Immun 76: 5478–5487.
22. Carlsson F, Joshi SA, Rangell L, Brown EJ (2009) Polar localization of virulence-related Esx-1 secretion in mycobacteria. PLoS Pathog 5: e1000285. doi:10.1371/journal.ppat.1000285.
23. Sorensen AL, Nagai S, Houen G, Andersen P, Andersen AB (1995) Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*. Infect Immun 63: 1710–1717.
24. Harboe M, Oettinger T, Wiker HG, Rosenkrands I, Andersen P (1996) Evidence for occurrence of the ESAT-6 protein in *Mycobacterium tuberculosis* and virulent *Mycobacterium bovis* and for its absence in *Mycobacterium bovis* BCG. Infect Immun 64: 16–22.
25. Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK (1996) Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. J Bacteriol 178: 1274–1282.
26. Berthet FX, Rasmussen PB, Rosenkrands I, Andersen P, Gicquel B (1998) A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). Microbiology 144: 3195–3203.
27. Renshaw PS, Panagiotidou P, Whelan A, Gordon SV, Hewinson RG, et al. (2002) Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6*CFP-10 complex. Implications for pathogenesis and virulence. J Biol Chem 277: 21598–21603.
28. Renshaw PS, Lightbody KL, Veverka V, Muskett FW, Kelly G, et al. (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. Embo J 24: 2491–2498.
29. Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, et al. (2003) Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation. J Infect Dis 187: 117–123.
30. Hsu T, Hingley-Wilson SM, Chen B, Chen M, Dai AZ, et al. (2003) The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. Proc Natl Acad Sci U S A 100: 12420–12425.
31. Guinn KI, Hickey MJ, Mathur SK, Zakel KL, Grotzke JE, et al. (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. Mol Microbiol 51: 359–370.
32. Pym AS, Brodin P, Brosch R, Huerre M, Cole ST (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. Mol Microbiol 46: 709–717.
33. Brodin P, Majlessi L, Marsollier L, de Jonge MI, Bottai D, et al. (2006) Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence. Infect Immun 74: 88–98.
34. Frigui W, Bottai D, Majlessi L, Monot M, Josselin E, et al. (2008) Control of *M. tuberculosis* ESAT-6 secretion and specific T cell recognition by PhoP. PLoS Pathog 4: e33. doi:10.1371/journal.ppat.0040033.
35. Pallen MJ (2002) The ESAT-6/WXG100 superfamily – and a new Gram-positive secretion system? Trends Microbiol 10: 209–212.
36. Skjot RL, Oettinger T, Rosenkrands I, Ravn P, Brock I, et al. (2000) Comparative evaluation of low-molecular-mass proteins from *Mycobacterium tuberculosis* identifies members of the ESAT-6 family as immunodominant T-cell antigens. Infect Immun 68: 214–220.
37. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393: 537–544.
38. Gey van Pittius NC, Sampson SL, Lee H, Kim Y, van Helden PD, et al. (2006) Evolution and expansion of the *Mycobacterium tuberculosis* PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions. BMC Evol Biol 6: 95.
39. Fortune SM, Jaeger A, Sarracino DA, Chase MR, Sassetti CM, et al. (2005) Mutually dependent secretion of proteins required for mycobacterial virulence. Proc Natl Acad Sci U S A 102: 10676–10681.
40. MacGurn JA, Raghavan S, Stanley SA, Cox JS (2005) A non-RD1 gene cluster is required for Snm secretion in *Mycobacterium tuberculosis*. Mol Microbiol 57: 1653–1663.
41. Raghavan S, Manzanillo P, Chan K, Dovey C, Cox J (2008) Secreted transcription factor controls *Mycobacterium tuberculosis* virulence. Nature 454: 717–721.
42. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, et al. (2001) Massive gene decay in the leprosy bacillus. Nature 409: 1007–1011.
43. Gao LY, Guo S, McLaughlin B, Morisaki H, Engel JN, et al. (2004) A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. Mol Microbiol 53: 1677–1693.
44. Tan T, Lee WL, Alexander DC, Grinstein S, Liu J (2006) The ESAT-6/CFP-10 secretion system of *Mycobacterium marinum* modulates phagosome maturation. Cell Microbiol 8: 1417–1429.
45. Xu J, Laine O, Masciocchi M, Manoranjan J, Smith J, et al. (2007) A unique Mycobacterium ESX-1 protein co-secretes with CFP-10/ESAT-6 and is necessary for inhibiting phagosome maturation. Mol Microbiol 66: 787–800.
46. McLaughlin B, Chon JS, MacGurn JA, Carlsson F, Cheng TL, et al. (2007) A mycobacterium ESX-1-secreted virulence factor with unique requirements for export. PLoS Pathog 3: e105. doi:10.1371/journal.ppat.0030105.
47. Davis JM, Ramakrishnan L (2009) The role of the granuloma in expansion and dissemination of early tuberculous infection. Cell 136: 37–49.
48. Stinear TP, Seemann T, Harrison PF, Jenkin GA, Davies JK, et al. (2008) Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. Genome Res 18: 729–741.
49. Flint JL, Kowalski JC, Karnati PK, Derbyshire KM (2004) The RD1 virulence locus of *Mycobacterium tuberculosis* regulates DNA transfer in *Mycobacterium smegmatis*. Proc Natl Acad Sci U S A 101: 12598–12603.
50. Converse SE, Cox JS (2005) A protein secretion pathway critical for *Mycobacterium tuberculosis* virulence is conserved and functional in *Mycobacterium smegmatis*. J Bacteriol 187: 1238–1245.
51. Coros A, Callahan B, Battaglioli E, Derbyshire KM (2008) The specialized secretory apparatus ESX-1 is essential for DNA transfer in *Mycobacterium smegmatis*. Mol Microbiol 69: 794–808.
52. Abdallah AM, Verboom T, Hannes F, Safi M, Strong M, et al. (2006) A specific secretion system mediates PPE41 transport in pathogenic mycobacteria. Mol Microbiol 62: 667–679.
53. Abdallah AM, Savage ND, van Zon M, Wilson L, Vandenbroucke-Grauls CM, et al. (2008) The ESX-5 secretion system of *Mycobacterium marinum* modulates the macrophage response. J Immunol 181: 7166–7175.
54. Rodriguez GM, Voskuil MI, Gold B, Schoolnik GK, Smith I (2002) ideR, An essential gene in *Mycobacterium tuberculosis*: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. Infect Immun 70: 3371–3381.
55. Maciag A, Dainese E, Rodriguez GM, Milano A, Provvedi R, et al. (2007) Global analysis of the *Mycobacterium tuberculosis* Zur (FurB) regulon. J Bacteriol 189: 730–740.
56. Sassetti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol 48: 77–84.
57. Burts ML, Williams WA, Debord K, Missiakas DM (2005) EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus aureus* infections. Proc Natl Acad Sci U S A 102: 1169–1174.
58. Garufi G, Butler E, Missiakas D (2008) ESAT-6-like protein secretion in *Bacillus anthracis*. J Bacteriol 190: 7004–7011.
59. Brown GD, Dave JA, Gey van Pittius NC, Stevens L, Ehlers MR, et al. (2000) The mycosins of *Mycobacterium tuberculosis* H37Rv: a family of subtilisin-like serine proteases. Gene 254: 147–155.
60. Jones DT (2007) Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics 23: 538–544.