

thermore, while immunosubunit precursors-containing 15S proteasome assembly intermediates were detected in wild-type cells, under identical conditions 15S complexes were hardly detectable in PA28<sup>-/-</sup> cells (Fig. 4B). Notably, PA28 was found to be associated with the immunosubunit precursor-containing 15S complexes, suggesting that PA28 is required for the incorporation of immunosubunits into proteasomes.

It has been demonstrated that immunosubunits are required for efficient antigen processing (1) and that, in vitro, PA28 does not preferentially activate immunoproteasomes and does not alter proteasome substrate specificity (15). Thus, our findings strongly suggest that the inability to generate immune responses in PA28<sup>-/-</sup> mice was a consequence of a decreased cellular level of immunoproteasomes. Taken together with the fact that, in PA28-expressing cells, immunoproteasomes have greater stability than proteasomes not containing immunosubunits (3, 12) and that PA28 interacts with the α-subunit ring of the proteasome (1), we hypothesize that, by inducing conformational changes of the proteasome α-subunit ring (3, 16), PA28 promotes immunoproteasome assembly. PA28 might facilitate recruitment and assembly of the immunosubunit-containing proteasome β-subunit ring onto the α-subunit ring, which serves as a scaffold for the proteasome β-subunit ring formation and Ump1p-mediated maturation (17). As a consequence of increased cellular level of immunoproteasomes, MHC class I antigen processing and presentation are greatly enhanced by PA28.

References and Notes

1. K. L. Rock and A. L. Goldberg, *Annu. Rev. Immunol.* **17**, 739 (1999); K. Früh and Y. Yang, *Curr. Opin. Immunol.* **11**, 76 (1999).
2. C. Realini et al., *J. Biol. Chem.* **272**, 25483 (1997); X. Song, J. Von Kampen, C. A. Slaughter, G. N. Demartino, *J. Biol. Chem.* **272**, 27994 (1997).
3. K. Ahn et al., *J. Biol. Chem.* **271**, 18237 (1996); Y. Yang, K. Früh, K. Ahn, P. A. Peterson, *J. Biol. Chem.* **270**, 27687 (1995).
4. T. P. Dick et al., *Cell* **86**, 253 (1996); M. Groettrup et al., *Nature* **381**, 166 (1996).
5. Y. Li et al., *Immunogenetics* **49**, 149 (1998).
6. Supplemental Web figures are available at [www.sciencemag.org/feature/data/1043186.shl](http://www.sciencemag.org/feature/data/1043186.shl).
7. J. R. Knowlton et al., *Nature* **390**, 639 (1997).
8. M. R. Jackson, E. S. Song, Y. Yang, P. A. Peterson, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 12117 (1992).
9. Splenocytes were metabolically labeled for 60 min followed by a 60-min chase. NP-40 cell lysates were incubated at 31°, 37°, and 43°C for 45 min before being subjected to immunoprecipitation with anti-K<sup>b</sup> or D<sup>b</sup>. The amounts of class I molecules were quantified. It was found that while an equivalent amount of class I molecules was present in wild-type and PA28<sup>-/-</sup> splenocytes before temperature challenge, the amount of class I molecules in PA28<sup>-/-</sup> cells after the temperature challenge decreased to ~80% of the wild-type levels.
10. A. Porgador, J. W. Yewdell, Y. Deng, J. R. Bennink, R. N. Germain, *Immunity* **6**, 715 (1997). Ovalbumin-loaded LPS blasts from wild-type and PA28<sup>-/-</sup> mice were stained with 25D1.16 and analyzed by flow cytometry. It was found that ~13% of the wild-type cells

stained positive, whereas no staining of PA28<sup>-/-</sup> cells was observed.

11. S. C. Jameson, F. R. Carbone, M. J. Bevan, *J. Exp. Med.* **177**, 1541 (1993).
12. T. Preckel and Y. Yang, unpublished results.
13. M. Ho, *Cytomegalovirus: Biology and Infection* (Plenum, New York, ed. 2, 1991).
14. P. Borrow and M. B. A. Oldstone, in *Viral Pathogenesis*, N. Nathanson, Ed. (Lippincott-Raven, Philadelphia, 1997), pp. 593–627; M. J. Buchmeier and A. J. Zajac, in *Persistent Virus Infections*, R. Ahmed and I. Chen, Eds. (Wiley, New York, 1999), pp. 575–605.
15. G. Niedermann et al., *J. Exp. Med.* **186**, 209 (1997); M. Groettrup et al., *J. Biol. Chem.* **270**, 23808 (1995).
16. M. Conconi, L. Djavadi-Ohanian, W. Uerkvitz, K. B. Hendil, B. Friguet, *Arch. Biochem. Biophys.* **362**, 325 (1999).
17. P. C. Ramos, J. Hockendorff, E. S. Johnson, A. Varshavsky, R. J. Dohmen, *Cell* **92**, 489 (1998).
18. Antisera specific to the proteasome subunit C9, LMP2, PA28a, and PA28b have been described (3). Antibodies specific to CD4, CD8, CD3-e, TCR-α, CD23, CD25, CD28, CD45, CD69, class I, and class II molecules were

purchased from PharMingen (San Diego, CA). 25D1.16 (10) and a PA28a-specific antiserum (2) were kindly provided by A. Porgador and M. Rechsteiner, respectively. Metabolic labeling, immunoprecipitation, immunoblotting, and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) were performed as described (3).

19. T. Flad et al., *Cancer Res.* **58**, 5803 (1998).
20. M. W. Moore, F. R. Carbone, M. J. Bevan, *Cell* **54**, 777 (1988).
21. A. Vitiello et al., *Eur. J. Immunol.* **27**, 671 (1997).
22. H. Bluthmann et al., *Nature* **334**, 156 (1988).
23. W. Jacoby, P. V. Cammarata, S. Findlay, S. H. Pincus, *J. Invest. Dermatol.* **83**, 302 (1984).
24. All animal experiments were conducted in accordance with institutional guidelines. We thank S. Sutton, J. Culver and S. Courtney for technical assistance and J.-F. Huang, and G. Schoenhals for critical reading. The technical assistance of the DNA, vivarium, and peptide synthesis facilities of the R. W. Johnson Pharmaceutical Research Institute is gratefully acknowledged.

2 July 1999; accepted 2 November 1999

## Global Transposon Mutagenesis and a Minimal Mycoplasma Genome

Clyde A. Hutchison III,<sup>1,2\*</sup> Scott N. Peterson,<sup>1\*†</sup> Steven R. Gill,<sup>1</sup> Robin T. Cline,<sup>1</sup> Owen White,<sup>1</sup> Claire M. Fraser,<sup>1</sup> Hamilton O. Smith,<sup>1‡</sup> J. Craig Venter<sup>1‡§</sup>

*Mycoplasma genitalium* with 517 genes has the smallest gene complement of any independently replicating cell so far identified. Global transposon mutagenesis was used to identify nonessential genes in an effort to learn whether the naturally occurring gene complement is a true minimal genome under laboratory growth conditions. The positions of 2209 transposon insertions in the completely sequenced genomes of *M. genitalium* and its close relative *M. pneumoniae* were determined by sequencing across the junction of the transposon and the genomic DNA. These junctions defined 1354 distinct sites of insertion that were not lethal. The analysis suggests that 265 to 350 of the 480 protein-coding genes of *M. genitalium* are essential under laboratory growth conditions, including about 100 genes of unknown function.

One important question posed by the availability of complete genomic sequences (1–3) is how many genes are essential for cellular life. We are now in a position to approach this problem by rephrasing the question “What is life?” in genomic terms: “What is a minimal set of essential cellular genes?”

Interest in the minimal cellular genome predates genome sequencing [for a review, see (4)]. The smallest known cellular genome (5) is that of *Mycoplasma genitalium*, which is only 580 kb. This genome has been completely sequenced, and analysis of the se-

quence revealed 480 protein-coding genes plus 37 genes for RNA species (2).

The fraction of nonminimal genomes that is essential for cell growth and division has been experimentally measured in yeast (12%) and in the bacterium *Bacillus subtilis* (9%) (6). The indispensable portion of the *B. subtilis* genome was estimated to be 562 kb, close to the size of the *M. genitalium* genome. Theoretical approaches to defining a minimal gene set have also been attempted. With the availability of the first two complete genome sequences (*Haemophilus influenzae* and *M. genitalium*) and the assumption that genes conserved across large phylogenetic distances are likely to be essential, a minimal gene set of 256 genes was proposed (7).

*Mycoplasma pneumoniae* is the closest known relative of *M. genitalium*, with a genome size of 816 kb, 236 kb larger than that of *M. genitalium* (3). Comparison of the two genomes indicates that *M. pneumoniae* in-

<sup>1</sup>The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA. <sup>2</sup>Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

\*These authors contributed equally to this report.  
 †To whom reprint requests should be addressed.  
 ‡Present address: Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA.  
 §To whom correspondence should be addressed.

## REPORTS

cludes orthologs of virtually every one of the 480 *M. genitalium* protein-coding genes, plus an additional 197 genes (8). There is a substantial evolutionary distance between orthologous genes in the two species, which share an average of only 65% amino acid sequence identity. The existence of these two species with overlapping gene content provided an experimental paradigm to test whether the 480 protein-coding genes shared between the species were already close to a minimal gene set. We applied transposon mutagenesis to these completely sequenced genomes, which permitted precise localization of insertion sites with respect to each of the coding sequences.

Populations of 200 to 1000 viable mycoplasmas harboring independent transposon insertions were produced, and libraries of DNA fragments containing the junctions between the transposon and the chromosome were prepared and sequenced (9) (Table 1). Analysis of 2209 transposon junction fragments yielded 1354 different insertion sites. This data set is divided approximately equally between the two organisms. A total of 71% of the insertions were within genes in *M. genitalium* versus 61% in *M. pneumoniae*. This represents a substantial preference for intergenic insertion—because coding sequence constitutes 85% of the *M. genitalium* genome and 89% of the *M. pneumoniae* genome—and is consistent with the idea that intergenic sequences are less critical than protein-coding regions for viability. Transposon insertions have been identified in 140 different genes in *M. genitalium* and 179 different genes in *M. pneumoniae*.

The preference for insertion into the species-specific portion of the *M. pneumoniae* genome was striking (Fig. 1). The average density of distinct viable insertion events observed in *M. pneumoniae*-specific regions

(1.8 hits/kb) was about 5.5 times that found in the portion common to both species (0.33 hits/kb). This result supports our assumption that the *M. pneumoniae*-specific portion of the genome is fully dispensable. In addition to the species-specific insertions in *M. pneumoniae*, insertions were observed widely distributed throughout the shared regions of both genomes. The conspicuous absence of transposon insertions into certain regions expected to be essential—for example, the region containing a cluster of ribosomal genes (MP637 to MP668; Fig. 1)—provides additional support for the validity of transposon mutagenesis as an assay for dispensability. A paucity of hits in other genes involved in transcription, translation, and DNA metabolism was also apparent.

Not every transposon insertion within a gene is expected to disrupt gene function. An insertion near the 3' end of a gene may only remove a nonessential COOH-terminus of the protein. Similarly, an insertion near the 5' end of a gene may not always destroy gene function. Transposon Tn4001 (10) contains an outward-directed promoter that could drive transcription of flanking chromosomal DNA (11), leading to translation if an internal start site is located nearby downstream. For the purposes of cataloging potentially dispensable genes, we have considered an insertion disruptive if it is within the 5'-most 80% of the gene but downstream of nucleotide 9 of the protein-coding region. This criterion eliminates events in which the 5' end of the gene may actually be intact because of duplication of a short sequence at the target site (10), and eliminates potentially nondisruptive COOH-terminal insertions. The fraction of putatively disruptive insertions within *M. genitalium* genes is 66%, compared with 84% for *M. pneumoniae* (Table 1). This difference can be attributed to a higher proportion of nonessential genes in the *M. pneumoniae* genome. The majority of *M. genitalium* orthologs that have disruptive insertions are absent from the third fully sequenced myco-

plasma genome, *Ureaplasma urealyticum* (<http://genome.microbio.uab.edu>), consistent with the idea that they are not essential.

One approach to estimating the total number of nonessential *M. genitalium* genes is to determine the number of dispensable orthologs of these genes in *M. pneumoniae* (see Table 2). We obtain similar estimates of nonessential *M. genitalium* orthologs whether we use data from all genes disrupted in *M. pneumoniae* (~121 orthologs) or just from those independently disrupted more than once (~108 orthologs). Using the pooled data from both species (12), the total number of *M. genitalium* orthologs in which presumptively disruptive insertions are observed is 129, close to the estimates of the total number obtained from the *M. pneumoniae* data.

We have also estimated the number of nonessential genes to be ~180 to ~215 under the assumption that the number of sites hit per gene follows a Poisson distribution. These larger estimates fit reasonably with the observed proportion of orthologs hit in both species. Therefore, on the basis of our highest and lowest estimates for nonessential genes, we estimate that the number of essential mycoplasma protein-coding genes is between 265 and 350.

The 351 *M. genitalium* orthologs for which we have not yet identified a disruptive insertion constitute a first approximation to the true set of essential mycoplasma genes. From our estimate of the total number of essential genes ( $265/351 = \sim 3/4$ ), we predict that at least 3/4 of the undisrupted genes are essential. We also expect that most undisrupted genes within each functional class represent essential genes. Examination of the gene disruption data, organized by functional role, reveals that all functional classes of genes are not equally mutable under the selective growth conditions used in this study, which suggests that the genes are closer to a minimal set for some cellular functions than for others (12).

The portion of the mycoplasma genome dedicated to coding lipoproteins is relatively

**Table 1.** Summary of sequenced viable transposon insertion sites.

	<i>M. genitalium</i>	<i>M. pneumoniae</i>
Junctions sequenced*	1291	918
Distinct sites†	685	669
Intergenic sites‡	199	261
Sites in genes‡	484	408
Different genes§	140	179
"Disrupted" genes	93	150

\*Total numbers of interpretable sequences that include a junction between the end of the transposon and the mycoplasma genome sequence. †Numbers of unambiguously different positions in the genome sequence where a transposon junction was observed. The differences between these numbers represent multiple occurrences of identical insertion sites, which tended to cluster in certain genes, particularly in *M. genitalium*. ‡Annotated open reading frames and RNA genes in the two genomes, with respect to the total number of junctions sequenced. §Total number of annotated genes in which transposon insertions were observed. ||See the text for a discussion of the criteria for counting "disrupted" genes.

**Table 2.** Estimating the total number of dispensable *M. genitalium* orthologs in *M. pneumoniae*. Among the 150 putative gene disruptions in *M. pneumoniae*, 57 are in genes that have orthologs in the *M. genitalium* genome, and 93 are in *M. pneumoniae*-specific genes. We have identified disruptive insertions in 47% (93/197) of the *M. pneumoniae*-specific genes. It is then a reasonable assumption that 47% of the dispensable *M. pneumoniae* genes common to both genomes have also been disrupted. This leads to an estimate that ~121 *M. genitalium* orthologs are dispensable and ~318 genes in total (121 + 197) are dispensable in *M. pneumoniae*. An analogous calculation using only those genes that have been hit more than once leads to an estimate of ~108 dispensable *M. genitalium* orthologs.

Data from	Hits		Estimated total dispensable <i>M. genitalium</i> orthologs
	Species-specific	<i>M. genitalium</i> orthologs	
All genes disrupted in <i>M. pneumoniae</i>	93	57	121
Genes disrupted more than once in <i>M. pneumoniae</i>	42	23	108

## REPORTS

large and suggests that this class of membrane proteins is important to the cell. Among the 19 genes encoding putative lipoproteins, we have identified potential disruptions in 13. There are several plausible interpretations of these seemingly incongruent facts, but perhaps the most likely is that the importance of these proteins is limited to fulfilling essential functions in the human host. This idea is substantiated by the occur-

rence of disruptions of several genes [*M. genitalium*/*M. pneumoniae* orthologs MG191 (MP014), MG192 (MP013), MG218 (MP527), and MG317 (MP388)] that are involved either directly or indirectly in mediating adherence to host cells (13).

The large group of genes with no functional assignment includes many genes with no known homologs outside of the mycoplasmas. As expected, many genes in this group

have been disrupted (69 of 180). However, most of the 111 undisrupted genes of unknown function are apparently not dispensable and are expected to encode essential cellular functions.

Relatively few *M. genitalium* genes have functions related to biosynthesis and metabolism. This limited metabolic capacity has been compensated for by a proportionately greater dependence on transport of raw ma-



**Fig. 1.** Viable transposon insertions displayed on a composite *M. pneumoniae*-*M. genitalium* map. The *M. pneumoniae* genome is shown at a scale of 30 kb per line. Colored arrows above the line indicate annotated genes. Genes are colored according to their functional category, as indicated in the key. The genes are numbered sequentially from 1 to 677 as listed on Richard Herrmann's Web page ([http://mail.zmbh.uni-heidelberg.de/M\\_pneumoniae/genome/Sorted\\_Genes.html](http://mail.zmbh.uni-heidelberg.de/M_pneumoniae/genome/Sorted_Genes.html)) and are referred to in the text as MP001 to MP677. Red triangles below the line indicate positions of transposon insertions documented in *M. genitalium*, mapped onto the *M. pneumoniae* genome (20). Red triangles above the line indicate positions of transposon insertions documented in *M. pneumoniae*. Regions of the genome that are *M. pneumoniae*-specific (absent from *M. genitalium*) are highlighted in pink directly on the line (20).

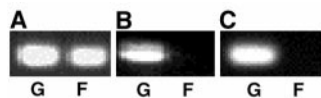
html) and are referred to in the text as MP001 to MP677. Red triangles below the line indicate positions of transposon insertions documented in *M. genitalium*, mapped onto the *M. pneumoniae* genome (20). Red triangles above the line indicate positions of transposon insertions documented in *M. pneumoniae*. Regions of the genome that are *M. pneumoniae*-specific (absent from *M. genitalium*) are highlighted in pink directly on the line (20).



terials from the extracellular environment. Although the ability to generate nucleotides by salvage pathways has been retained, as have a limited number of biosynthetic and metabolic enzymes, it is evident that these pathways are nonessential in the laboratory, where the organism is apparently able to import nucleosides, amino acids, and other metabolites. Likewise, genes involved in the biosynthesis of cofactors [MG270 (MP450)], fatty acid and phospholipid metabolism [MG310 (MP395)], and hexose conversion [MG118 (MP577)] appear to be dispensable.

Our data strongly support the idea that some metabolic pathways are essential. Glycolysis is thought to be the major source of adenosine triphosphate (ATP) and energy for *M. genitalium* and *M. pneumoniae*. We have not observed any disruptive insertions in any of the 10 genes involved in this pathway. Likewise, we have not identified any dispensable genes among the eight genes encoding ATP-proton-motive force interconversion activities.

ABC transporters are a heterotrimeric transport system made up of a specificity (ligand-binding) subunit, a permease, and an ATP-binding protein. ATP-binding subunits are distinct in that many appear to be "orphan" proteins, which are apparently overrepresented, compared to the other two subunits, in all genomes sequenced thus far. Likewise, there are specificity subunits with unknown partners within the *M. genitalium* genome, although their occurrence appears to be much more limited. Analysis of the *M. genitalium* genomic sequence data with less stringent searching parameters aimed at finding partners of the orphan specificity subunits led to the identification of potential transport partners (14). Because the sequence relatedness of these transporters was quite low, *M. genitalium* was thought to compensate for a reduced transporter spectrum by encoding transporters with broadened specificity (14). The current annotation of the *M. geni-*



**Fig. 2.** Analysis of conditionally dispensable genes. PCR was done using one transposon-specific primer directed toward the chromosomal junction, and a gene-specific primer priming toward the particular insertion site to be detected (21). (A) Control experiment in which an insertion in a gene of unknown function [MG296 (MP417)] was detected in the same DNAs used in (B). (B) PCR primers were designed to detect the presence of an insertion in MG062 in pool C of *M. genitalium* Tn4001 transformants. (C) PCR primers were designed to detect an insertion in the *M. pneumoniae* ortholog of MG062 (MP077) in pool E of *M. pneumoniae* Tn4001 transformants. Lane labels G (glucose) and F (fructose) indicate the sugar used to supplement the growth medium (21).

*talium* genome lists 12 "orphan" ATP-binding proteins. We have obtained central insertions in only three of these genes [MG014 (MP136), MG390 (MP271), and MG467 (MP159)]. The fact that only 25% of the ATP-binding subunits in our data set tolerate insertions suggests that at least some of these orphan subunits do serve an essential function within the cell.

Two of the three subunits of an ABC phosphate transporter [MG410 (MP233) and MG411 (MP232)] have been putatively disrupted. Phosphate transport is thought to be an essential function. The insertion data for these two genes appear to be quite solid, in that MG410 (MP233) insertions were uncovered in both *M. genitalium* and *M. pneumoniae* and MG411 (MP232) contained multiple independent insertion events in *M. genitalium*. This finding forces us to consider the possibility that some as yet undefined transport system exists in these mycoplasmas that can compensate for mutations in the putative phosphate transporter.

Both *Mycoplasma* species examined in this study include two genes homologous to DNA pol III subunits [MG261 (MP460) and MG031 (MP120)]. We have found an insertion into MG261 (MP460), which supports the idea that it may function as a repair enzyme (3, 15) rather than being the main replication enzyme in the cell. The MG261 (MP460) insertion, together with insertions in the *recA* and *uvrA* excision repair genes, represents a special class of dispensable cellular functions. It is not necessarily surprising that cells can tolerate transposon insertions in these genes. It is almost certain that cells bearing such gene disruptions in nature would be quickly selected against. Although it is difficult to address this idea quantitatively, it poses a relevant question for consideration when attempting to define a minimal gene set for cellular life.

As expected, the number of transposon insertions recovered in genes involved in transcription was small. We have identified apparent disruptions in two of the five genes annotated as putative ATP-dependent RNA helicases [MG017 (MP134) and MG308 (MP397)]. The level of functional redundancy within this cellular role is somewhat unexpected and may reflect a broader role for these genes with regard to replication, repair, or transcriptional regulation.

A few insertions were in genes generally believed to be essential (12). Such events represent less than 1% of the total number of mapped insertion sites (1354). They include single putatively disruptive insertion events in two aminoacyl tRNA synthetase genes, the gene for ribosomal protein L28, the DNA replication genes *dnaA* and *gidB*, and a sigma factor gene. These unexpected findings forced us to consider explanations other than the dispensability of a function presumed to be essential. Functional assignments for some of these

genes on the basis of sequence similarity may be incorrect. Also, some events that meet the criterion used in our analysis may not disrupt gene function. It is highly improbable that these events were recovered by cloning transposon junctions from nonviable cells. However, some cells might contain a functional duplicate copy of a gene in addition to the disrupted gene. It may also be that some functions can be supplied by unexpected uptake of enzymes or other compounds from the medium, or by cross-feeding. Conclusive proof of the dispensability of any specific gene requires cloning and detailed characterization of a pure population carrying the disrupted gene.

It is possible in some cases to verify the disruption of a specific gene function despite the presence of a transposon insertion mutant within a mixed population. For fructose-permease, insertions in both the *M. genitalium* gene (MG062) and the orthologous *M. pneumoniae* gene (MP077) were identified. We were able to detect the sequenced insertion events, using the polymerase chain reaction (PCR), in DNA from populations of cells grown in glucose-supplemented medium, but not when the medium was instead supplemented with fructose (Fig. 2).

The power of global transposon mutagenesis in our case benefited greatly from application to the fully sequenced and annotated genomes of the closely related *M. genitalium* and *M. pneumoniae*. Although *M. genitalium* may possess close to the minimal genome required for survival in its human host, it is clear from the results presented here that it contains a large number of genes that are dispensable under laboratory growth conditions. Our results imply that of the 111 genes of unknown function that have not been disrupted in our experiments, the majority are essential. The presence of so many genes of unknown function among the essential genes of the simplest known cell suggests that all the basic molecular mechanisms underlying cellular life may not yet have been described. The essential gene set is not the same as the minimal genome. It is clear that genes that are individually dispensable may not be simultaneously dispensable. The data presented here suggest some specific experiments that could be carried out as a first step in the engineering of a cell with a minimal genome in the laboratory environment. One way to identify a minimal gene set for self-replicating life would be to create and test a cassette-based artificial chromosome, an experiment pending ethical review (16).

**References and Notes**

1. R. D. Fleischmann *et al.*, *Science* **269**, 496 (1995); C. J. Bult *et al.*, *Science* **273**, 1058 (1996); T. Kaneko *et al.*, *DNA Res.* **3**, 109 (1996); A. Goffeau *et al.*, *Nature* **387** (suppl.) (1997); F. Blattner *et al.*, *Science* **277**, 1453 (1997); D. R. Smith *et al.*, *J. Bacteriol.* **179**, 7135 (1997); C. M. Fraser *et al.*, *Nature* **390**, 580 (1997); C. M. Fraser *et al.*, *Science* **281**, 375 (1998); H. P. Klenk *et al.*, *Nature* **390**, 364 (1997); F. Kunst *et al.*,

*Nature* **390**, 249 (1997); J. F. Tomb *et al.*, *Nature* **388**, 539 (1997).

2. C. M. Fraser *et al.*, *Science* **270**, 397 (1995). Updated annotation of the genome is available from the TIGR Microbial Database ([www.tigr.org/tdb/mdb/mdb.html](http://www.tigr.org/tdb/mdb/mdb.html)).
3. R. Himmelreich *et al.*, *Nucleic Acids Res.* **24**, 4420 (1996).
4. H. J. Morowitz, *Isr. J. Med. Sci.* **20**, 750 (1984); J. Maniloff, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10004 (1996).
5. S. D. Colman, p.-C. Hu, W. Litaker, K. F. Bott, *Mol. Microbiol.* **4**, 683 (1990).
6. M. G. Goebel and T. D. Petes, *Cell* **46**, 983 (1986); M. Itaya, *FEBS Lett.* **362**, 257 (1995).
7. A. R. Mushegian and E. V. Koonin, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1026 (1996).
8. R. Himmelreich, H. Plagens, H. Hilbert, B. Reiner, R. Herrmann, *Nucleic Acids Res.* **25**, 701 (1997).
9. Transposon Tn4001, originally from *Staphylococcus aureus*, was propagated in *Escherichia coli* plasmid pISM2062 (17) and then introduced into *M. genitalium* and *M. pneumoniae* cells by electroporation (18). Approximately one in 10<sup>3</sup> to 10<sup>4</sup> *M. pneumoniae* cells and one in 10<sup>5</sup> to 10<sup>6</sup> *M. genitalium* cells were transformed to resistance to gentamycin (Gm). Cultures were split immediately after electroporation to generate eight separate populations for each species. Each population harbored cells representing ~200 transposition events for *M. genitalium* and >1000 events for *M. pneumoniae*. These populations were allowed to recover in SP4 medium overnight, followed by growth in the presence of Gm for 2 to 4 weeks, resulting in the expansion of cell number by a factor of >10<sup>9</sup>. This procedure was designed to make the subsequent cloning of transposition events from nonviable cells highly improbable. Genomic DNA was isolated from mid-log cultures; 2 µg of DNA was digested with Dra I. The genomic DNA restriction digests were diluted to 5 ng/µl and fragments were circularized using DNA ligase. Transposon junctions were amplified using inverse PCR with two primers specific for the end of the transposon Tn4001. Reaction products containing oligonucleotide-encoded Eco RI and Hind III sites were digested with these enzymes and cloned into the corresponding sites in pUC18. DNA sequencing templates were prepared from selected colonies and sequences generated as described (2). Transposon junction sequences were aligned with the appropriate genomic sequence to establish the site in the genome of transposon insertion.
10. Tn4001 is a composite transposon with IS256 sequences at both ends. IS256 produces 8- and 9-base pair duplications of the target sequence at the insertion site [K. G. Dyke, S. Aubert, N. el Solh, *Plasmid* **28**, 235 (1992); L. B. Rice, L. L. Carias, S. H. Marshall, *Antimicrob. Agents Chemother.* **39**, 1147 (1995)].
11. C. E. Romero-Arroyo *et al.*, *J. Bacteriol.* **181**, 1079 (1999).
12. Supplementary data can be found at *Science Online* ([www.sciencemag.org/feature/data/1042937.shl](http://www.sciencemag.org/feature/data/1042937.shl)).
13. The proteins involved in mycoplasma adhesion have been most extensively studied in *M. pneumoniae*. For a review see D. C. Krause, *Mol. Microbiol.* **19**, 247 (1996).
14. W. Saurin and E. Dassa, *Mol. Microbiol.* **22**, 389 (1996).
15. E. V. Koonin and P. Bork, *Trends Biochem. Sci.* **21**, 128 (1996).
16. M. K. Cho *et al.*, *Science* **286**, 2087 (1999).
17. K. L. Knudson and F. C. Minion, *Gene* **137**, 217 (1993).
18. S. P. Reddy, W. G. Rasmussen, J. B. Baseman, *FEMS Immunol. Med. Microbiol.* **15**, 211 (1996).
19. A. L. Delcher *et al.*, *Nucleic Acids Res.* **27**, 2369 (1999).
20. The figure shows 600 *M. genitalium* sites that were unambiguously mapped onto the *M. pneumoniae* genome. The whole-genome alignment method of Delcher *et al.* (19) mapped 317 sites, but does not map sites close to insertion/deletion and rearrangement differences between the chromosomes. An additional 201 sites were mapped by searching the *M. pneumoniae* genome for matches to short sequences (200 to 400 base pairs) containing each *M. genitalium* insertion site. An additional 83 sites that were

not mappable by the above methods (because of matches to several related *M. pneumoniae* sequences) were mapped to the corresponding position within the orthologous *M. pneumoniae* gene. The *M. pneumoniae*-specific regions are from (3), with minor modifications. Each pink highlighted region contains a block of *M. pneumoniae*-specific genes and arbitrarily includes the intergenic regions flanking it.

21. In the *M. genitalium* experiments, the growth medium

is SP-4 supplemented with either glucose or fructose as indicated. In the *M. pneumoniae* experiment, the cells were grown in Hayflick's medium supplemented with glucose or fructose. Cells were grown five passages (split 1 to 10) under the indicated condition before preparation of DNA for PCR templates. This is approximately equivalent to nine doublings.

23 June 1999; accepted 20 October 1999

## Functional Human Corneal Equivalents Constructed from Cell Lines

May Griffith,<sup>1\*</sup> Rosemarie Osborne,<sup>2</sup> Rejean Munger,<sup>1</sup> Xiaojuan Xiong,<sup>1</sup> Charles J. Doillon,<sup>3</sup> Noelani L. C. Laycock,<sup>1</sup> Malik Hakim,<sup>1</sup> Ying Song,<sup>1</sup> Mitchell A. Watsky<sup>4</sup>

Human corneal equivalents comprising the three main layers of the cornea (epithelium, stroma, and endothelium) were constructed. Each cellular layer was fabricated from immortalized human corneal cells that were screened for use on the basis of morphological, biochemical, and electrophysiological similarity to their natural counterparts. The resulting corneal equivalents mimicked human corneas in key physical and physiological functions, including morphology, biochemical marker expression, transparency, ion and fluid transport, and gene expression. Morphological and functional equivalents to human corneas that can be produced *in vitro* have immediate applications in toxicity and drug efficacy testing, and form the basis for future development of implantable tissues.

The cornea comprises three major cellular layers: an outermost stratified squamous epithelium, a stroma with keratocytes, and an innermost monolayer of specialized endothelial cells. The structure of the cornea allows it to serve as a barrier to the outside environment and as a major element in the optical pathway of the eye (1). The cornea is transparent, avascular, and immunologically privileged (2), making it an excellent candidate for tissue engineering for transplantation. Various researchers have attempted to fabricate artificial corneas or parts of corneas *in vitro* (3), but there have been no reports of successfully reconstructed human corneas that mimic the anatomy and physiology of the human cornea.

Our objective was to develop a morphological and functional equivalent of the human cornea. Human cell lines were developed from cells isolated from the individual cellular layers of the cornea. Most were immortalized (4) by

infection with an amphotropic recombinant retrovirus containing HPV16 genes E6 and E7 (5, 6); others were immortalized by transfection (7) with mammalian expression vectors containing genes encoding SV40 large T antigen, pSV3neo (8), and adenovirus E1A 12S (9), separately or in combination. Immortalized cells had random chromosomal breaks, structural rearrangements, and, in several lines, aneuploidy (10); similar chromosomal anomalies associated with immortalization were reported in HPV E6/E7 immortalized vascular endothelial cells (6), although random structural rearrangements and aneuploidy are also present in normal human corneal cells (11). The immortalized cells also had significant telomerase activity (12) [associated with the immortalized phenotype (6)] compared to little or no activity in nonimmortalized cells.

Before use in corneal equivalents, cell lines were screened for morphological, biochemical, and electrophysiological similarities to freshly dissociated or low-passage corneal cells obtained from postmortem human corneas. Electrophysiological screening of epithelial cells by means of amphotericin-perforated patch clamping (13) showed that immortalized cells had whole-cell currents similar to those of cultured corneal epithelial cells (Fig. 1, A to F) (14). Cells with altered phenotypes and physiology (transformed cells), however, showed anomalous currents (Fig. 1, G to I). Patch clamping was also used in screening keratocyte and

<sup>1</sup>University of Ottawa Eye Institute and Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa Hospital-General Campus, Ottawa, Ontario K1H 8L6, Canada. <sup>2</sup>Human & Environmental Safety Division, Procter & Gamble Company, Miami Valley Laboratories, Cincinnati, OH 45253, USA. <sup>3</sup>Biomaterials Institute of Quebec, Pavillon St. Francois d'Assise, Centre Hospitalier et Universitaire de Quebec, Quebec G1L 3L5, Canada. <sup>4</sup>Department of Physiology, University of Tennessee College of Medicine, Memphis, TN 38163, USA.

\*To whom correspondence should be addressed. E-mail: [mgriffith@ogh.on.ca](mailto:mgriffith@ogh.on.ca)