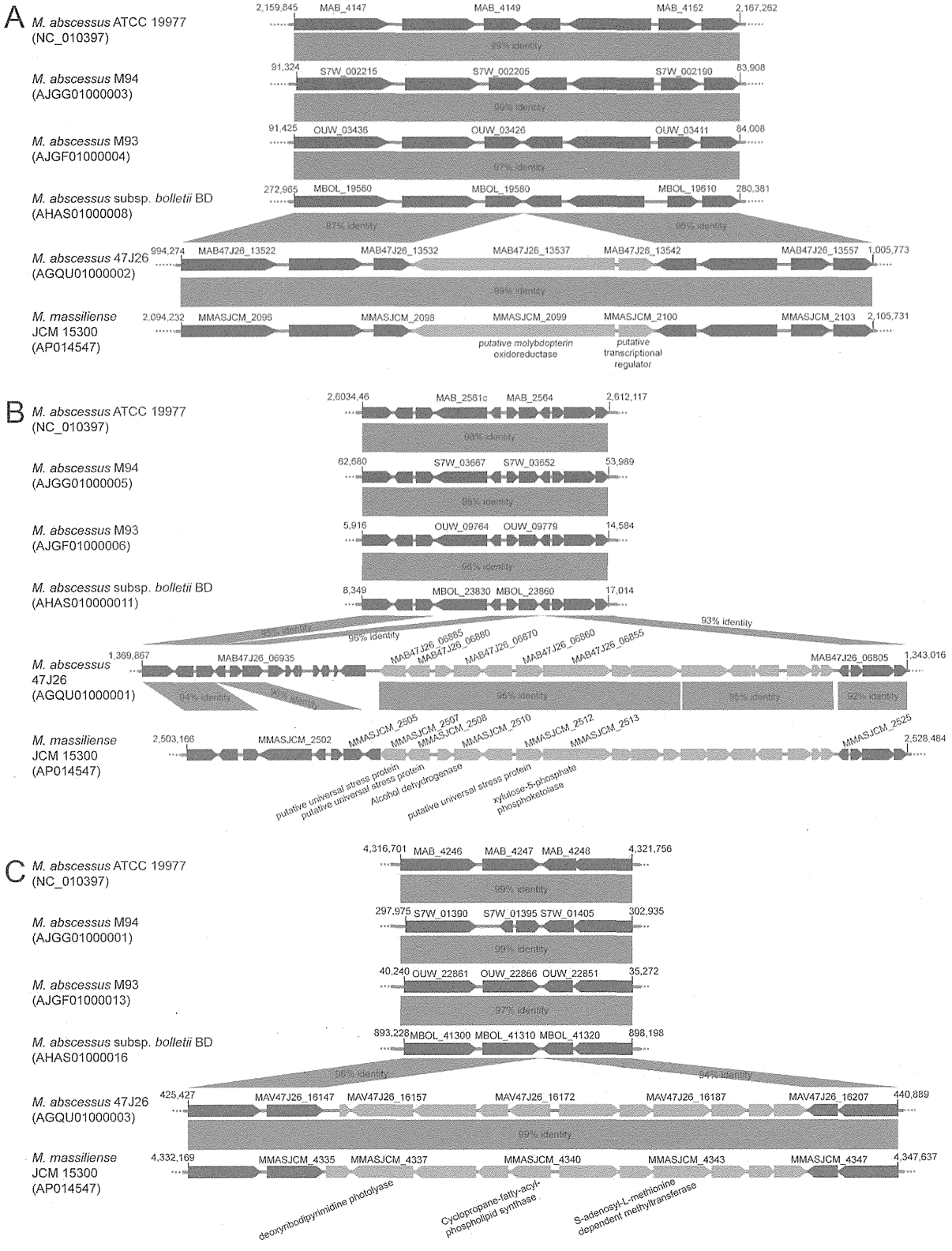

Fig. 4. Comparison of unique genes and flanking regions in the massiliense cluster. GenBank accession numbers are given in parentheses. The orange arrows indicate the unique genes in the massiliense cluster. BLASTN match scores less than 200 are not shown.

doi:10.1371/journal.pone.0114848.g004

Diego, CA), according to the manufacturer's instructions. A paired-end sequencing run for 83 mers was performed using an Illumina Genome Analyzer IIx (GA IIx) with the TruSeq SBS Kit v5. Fluorescent images were analyzed using the Illumina RTA1.8/SCS2.8 base-calling pipeline to obtain FASTQ-formatted sequence data.

De novo assembly of short DNA reads and gap-closing

Prior to *de novo* assembly, the obtained 80-mer reads were assembled using ABySS-pe v1.2.5 [36] with the following parameters: k60, n60, c68.4, t10, e10 and q20. Predicted gaps were amplified with specific PCR primer pairs followed by Sanger DNA sequencing with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA).

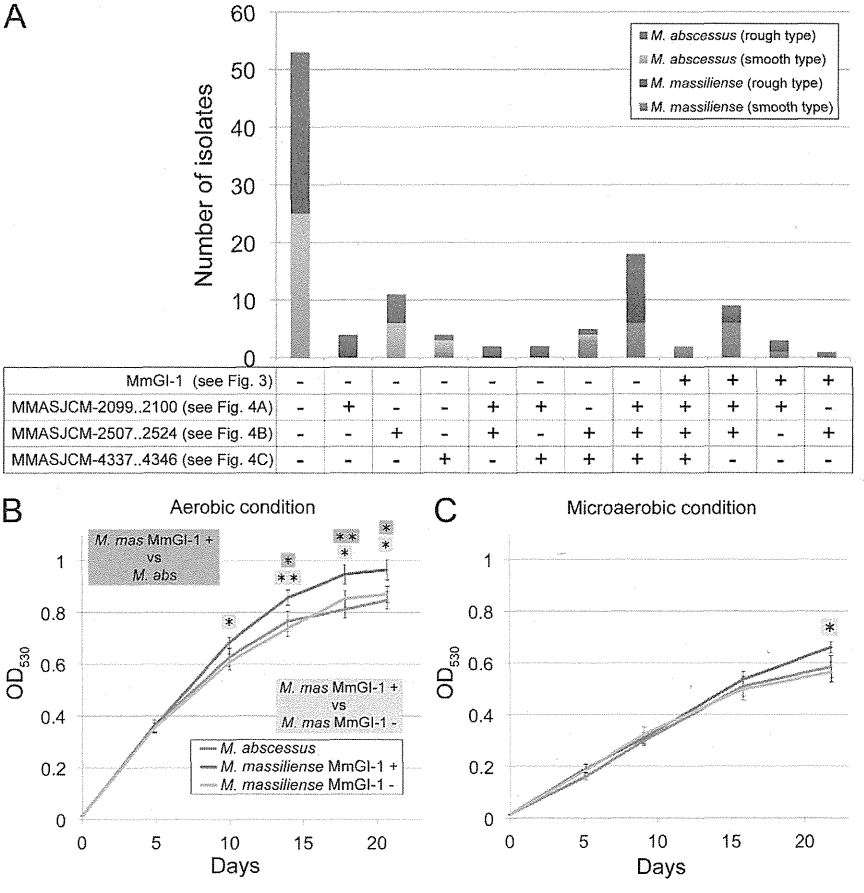

Fig. 5. Prevalence of massiliense cluster unique regions and growth curve analysis in Japanese *M. massiliense* and *M. abscessus* isolates. A bar chart showing the prevalence of MmGI-1 and three massiliense cluster unique regions in Japanese *M. massiliense* and *M. abscessus* isolates (A). The curves represent *in vitro* growth (OD at 530 nm) over a period of 21 days at 37°C in aerobic (B) and microaerobic (C) conditions. Data represent the means ± SE from 6 MmGI-1 positive *M. massiliense*, 8 MmGI-1 negative *M. massiliense* and 12 *M. abscessus* isolates. *M. mas* and *M. abs* shows *M. massiliense* and *M. abscessus*, respectively. Key: +, positive; -, negative. * $P<0.05$; ** $P<0.01$ (Student's t-test).

doi:10.1371/journal.pone.0114848.g005

Validation of gap closing and sequencing errors by short-read mapping

To determine whether mis-assembled sequences and incorrect gap-closing remained after reference-assisted gap-closing, 40-mer short reads were aligned to the tentative complete chromosomal DNA sequence using Maq software (ver. 0.7.1) with the easyrun Perl command [37]. We then performed a read alignment to validate possible errors using the MapView graphical alignment viewer [38].

Annotation

Gene prediction was performed for the complete genomic sequence with the RAST annotation server [39], followed by InterProScan [40] search and BLASTP search using nr database for validation. Genomic information, such as nucleotide variations and circular representations, was analyzed with gview software [41].

Pairwise alignment of chromosomal sequences

Pairwise alignment was performed by BLASTN and TBLASTN homology searches [42] followed by visualization of the aligned images with the ACT [43] or EMBOSS dottup program [44].

BLAST atlas

A BLAST atlas was generated by a BLASTP homology search [42] using the gview program [41]. The atlas displays BLASTP comparison results. The visualized area shows that the length of similar genes covers at least 80% between *M. massiliense* JCM 15300 and other *Mycobacterium* spp.

SNP analysis

To construct simulated paired-end reads from the available genomic sequences of *M. abscessus* group strains, SimSeq software [45] was used with "SimSeq.jar" and "SamToFastq.jar" commands with the following default parameter modifications: number of pairs of reads, "—read_number 2000000"; mean library insert size, "—insert_size 150"; and paired-end reads length of 120 mer, "−1 120 −2 120". These parameters indicated that 4 million hypothetical 120-mer reads were generated without mutations or indels from the genomic sequences used for SNP identification. To generate short-read mapping data of all *M. abscessus* group strains compared with the reference chromosomal sequence of *M. massiliense* JCM 15300, bwasw [46] and samtools [47] software was used with the default parameters. All SNPs were extracted by VarScan v2.3.4 [48] with the default parameters. All SNPs were concatenated to generate a pseudo sequence for phylogenetic analysis. The DNA maximum-likelihood program (RAxML v7.25) [49] was used for phylogenetic analysis with 1,000-fold bootstrapping. FigTree v. 1.2.3 software was used to display the generated tree.

Phylogenetic analysis

Nucleotide and amino acid sequences were aligned with mafft v6.86 [50] followed by phylogenetic analysis using the neighbor-joining method or maximum-likelihood method with 1,000-fold bootstrapping in clustalW2 [51] or RAxML v7.25 software [49]. FigTree v. 1.2.3 software was used to display the generated tree.

PCR amplification

The PCR mixture contained approximately 1 ng of template DNA, 1 × PrimeSTAR GXL Buffer (Takara Biochem. Shiga, Japan), 200 μM of each dNTP, 200 nM of each primer, and a total of 2.5 unit of PrimeSTAR GXL DNA polymerase (Takara Biochem.). The primer sequences for PCR amplification are shown in S4 Table. PCR was performed in 25 μl volumes under the following conditions: at 98℃ for 20 sec followed by 30 cycles at 98℃ for 15 sec, 65℃ for 15 sec and 68℃ for 1 min (for below 1.5 kb amplicons) or 5 min (for over 1.5 kb amplicons). Amplified PCR products were electrophoresed in 1.0% (w/v) agarose gel at 100 V and detected by staining with GelRed (Biotium Inc. Hayward, CA).

Bacterial culture

The *M. abscessus* and *M. massiliense* type strains were cultured at 37℃ in Middlebrook 7H9 broth (Difco) supplemented with 10% OADC (BD) and 0.05% Tween 80 under aerobic or microaerobic (6% aerobic O_2 tension) conditions with AnaeroPack (Mitsubishi Gas Chemical Company, Inc., Tokyo, Japan). Growth was monitored by removing aliquots at the indicated time points and measuring the OD at 530 nm.

Statistical analysis

The statistical test between MmGI-1 positive *M. massiliense* and *M. abscessus* was calculated by Fisher's Exact Test. Data of bacterial culture are expressed as mean ± standard error (SE) from 7 MmGI-1 positive *M. massiliense*, 8 MmGI-1 negative *M. massiliense* and 12 *M. abscessus* isolates. Statistical analysis was performed using the student's t-test. The t-test was used to investigate whether the means of two groups are statistically different from each other. Differences were considered significant with a p-value of <0.05 and 0.01.

Nucleotide sequence accession numbers

The complete genomic sequence of *M. massiliense* JCM 15300 has been deposited into the DNA Data Bank of Japan (DDBJ; accession number: AP014547).

Supporting Information

S1 Figure. Comparative analysis between the complete genomic sequence of the *M. massiliense* JCM 15300 strain and draft genomic sequences of *M. massiliense* CCUG 48898. The upper dot plot represents synteny between JCM 15300 and CCUG 48898, and the yellow vertical bars indicate gap regions in the draft genome of CCUG 48898. The bottom table shows gaps between contigs in CCUG 48898.
doi:10.1371/journal.pone.0114848.s001 (TIF)

S2 Figure. Genomic comparison and BLAST atlas of 3 clusters in the *M. abscessus* group. Comparative analysis of *M. massiliense* JCM 15300 and *M. abscessus* ATCC 19977 using a BLASTN homology search visualized by the ACT program (middle) and a BLAST atlas of *M. massiliense* JCM 15300 and *M. abscessus* ATCC 19977. In the comparative analysis, the red and blue bars between chromosomal DNA sequences represent nucleotide matches in the forward and reverse directions, respectively. BLASTN match scores less than 999 are not shown. In the BLAST atlas, the coding regions of JCM 15300 or ATCC 19977 were aligned against those of other *M. abscessus* group strains using BLASTP, and the results are displayed as colored bars (as in Fig. 1A). The three yellow boxes represent prophages on each chromosome. Specific features are represented by characters: †, unique region in the massiliense cluster; •, unique region in JCM 15300; §, unique region in the abscessus cluster; ¶, MmGI-1 (also see blue bars in Fig. 1A).
doi:10.1371/journal.pone.0114848.s002 (TIF)

S3 Figure. Visualization for *M. abscessus* group pan-genomes and core genomes. A. Curve for pan-genomes and core genomes of *M. abscessus* group. The box plots indicate the pan- or core genome size for each genome comparison. The median values were connected to represent the relationship between genome number and gene cluster number. B. Curve for the new gene cluster number observed with every increase in the number of *M. abscessus* group genomes.
doi:10.1371/journal.pone.0114848.s003 (TIF)

S4 Figure. Phylogenetic tree of mycolic acid cyclopropane synthetase domain (CMAS, pfam02353) proteins in *Mycobacterium* using the maximum-likelihood method with 1,000-fold bootstrapping. The scale indicates that a branch length of 0.3 is 30 times as long as one that would show a 1% difference between the amino acid sequences at the beginning and end of the branch. The number at each branch node represents the bootstrapping value. The proteins in red indicate proteins that are conserved only in the massiliense cluster.
doi:10.1371/journal.pone.0114848.s004 (TIF)

S1 Table. Mutation sites in the complete genomic sequence of *M. massiliense* JCM 15300 compared with those in draft genomic sequences of *M. massiliense* CCUG 48898.
doi:10.1371/journal.pone.0114848.s005 (PDF)

S2 Table. The unique gene loci in *M. massiliense* JCM15300.
doi:10.1371/journal.pone.0114848.s006 (PDF)

S3 Table. The deleted genes of massiliense and bolletii clusters among *M. abscessus* group.
doi:10.1371/journal.pone.0114848.s007 (PDF)

S4 Table. Oligonucleotide primer sequences used in PCR assays and the judging method for presence of MmGI-1 and other *M. massiliense* unique regions.
doi:10.1371/journal.pone.0114848.s008 (PDF)

S5 Table. Isolates analyzed in the present study and results of conventional PCR based detection against MmGI-1 and other *M. massiliense* unique regions.
doi:10.1371/journal.pone.0114848.s009 (PDF)

Author Contributions

Conceived and designed the experiments: TS M. Kai YH M. Kuroda. Performed the experiments: TS M. Kai KN NN YK SM YH M. Kuroda. Analyzed the data: TS M. Kai M. Kuroda. Contributed reagents/materials/analysis tools: TS M. Kai MM YH M. Kuroda. Wrote the paper: TS M. Kuroda. Performed genomic sequencing: TS M. Kai M. Kuroda.

References

1. **Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, et al.** (2007) An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. American journal of respiratory and critical care medicine 175: 367–416.
2. **Brown-Elliott BA, Nash KA, Wallace RJ Jr** (2012) Antimicrobial susceptibility testing, drug resistance mechanisms, and therapy of infections with nontuberculous mycobacteria. Clinical microbiology reviews 25: 545–582.
3. **Olivier KN, Weber DJ, Wallace RJ Jr, Faiz AR, Lee JH, et al.** (2003) Nontuberculous mycobacteria. I: multicenter prevalence study in cystic fibrosis. American journal of respiratory and critical care medicine 167: 828–834.
4. **Bryant JM, Grogono DM, Greaves D, Foweraker J, Roddick I, et al.** (2013) Whole-genome sequencing to identify transmission of Mycobacterium abscessus between patients with cystic fibrosis: a retrospective cohort study. Lancet 381: 1551–1560.
5. **Iseman MD, Marras TK** (2008) The importance of nontuberculous mycobacterial lung disease. American journal of respiratory and critical care medicine 178: 999–1000.
6. **Chan ED, Bai X, Kartalija M, Orme IM, Ordway DJ** (2010) Host immune response to rapidly growing mycobacteria, an emerging cause of chronic lung disease. American journal of respiratory cell and molecular biology 43: 387–393.
7. **Falkinham JO 3rd** (1996) Epidemiology of infection by nontuberculous mycobacteria. Clinical microbiology reviews 9: 177–215.
8. **Primm TP, Lucero CA, Falkinham JO 3rd** (2004) Health impacts of environmental mycobacteria. Clinical microbiology reviews 17: 98–106.
9. **Bastian S, Veziris N, Roux AL, Brossier F, Gaillard JL, et al.** (2011) Assessment of clarithromycin susceptibility in strains belonging to the Mycobacterium abscessus group by erm(41) and rrl sequencing. Antimicrobial agents and chemotherapy 55: 775–781.

10. **Macheras E, Roux AL, Bastian S, Leao SC, Palaci M, et al.** (2011) Multilocus sequence analysis and rpoB sequencing of Mycobacterium abscessus (sensu lato) strains. Journal of clinical microbiology 49: 491–499.

11. **Nakanaga K, Sekizuka T, Fukano H, Sakakibara Y, Takeuchi F, et al.** (2014) Discrimination of Mycobacterium abscessus subsp. massiliense from Mycobacterium abscessus subsp. abscessus in Clinical Isolates by Multiplex PCR. Journal of clinical microbiology 52: 251–259.

12. **Furuya EY, Paez A, Srinivasan A, Cooksey R, Augenbraun M, et al.** (2008) Outbreak of Mycobacterium abscessus wound infections among "lipotourists" from the United States who underwent abdominoplasty in the Dominican Republic. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America 46: 1181–1188.

13. **Medjahed H, Gaillard JL, Reyrat JM** (2010) Mycobacterium abscessus: a new player in the mycobacterial field. Trends in microbiology 18: 117–123.

14. **Villanueva A, Calderon RV, Vargas BA, Ruiz F, Aguero S, et al.** (1997) Report on an outbreak of postinjection abscesses due to Mycobacterium abscessus, including management with surgery and clarithromycin therapy and comparison of strains by random amplified polymorphic DNA polymerase chain reaction. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America 24: 1147–1153.

15. **Otsuki T, Izaki S, Nakanaga K, Hoshino Y, Ishii N, et al.** (2012) Cutaneous Mycobacterium massiliense infection: a sporadic case in Japan. The Journal of dermatology 39: 569–572.

16. **Nakanaga K, Hoshino Y, Era Y, Matsumoto K, Kanazawa Y, et al.** (2011) Multiple cases of cutaneous Mycobacterium massiliense infection in a "hot spa" in Japan. Journal of clinical microbiology 49: 613–617.

17. **Cho YJ, Yi H, Chun J, Cho SN, Daley CL, et al.** (2013) The Genome Sequence of 'Mycobacterium massiliense' Strain CIP 108297 Suggests the Independent Taxonomic Status of the Mycobacterium abscessus Complex at the Subspecies Level. PloS one 8: e81560.

18. **Ripoll F, Pasek S, Schenowitz C, Dossat C, Barbe V, et al.** (2009) Non mycobacterial virulence genes in the genome of the emerging pathogen Mycobacterium abscessus. PloS one 4: e5660.

19. **Choo SW, Wee WY, Ngeow YF, Mitchell W, Tan JL, et al.** (2014) Genomic reconnaissance of clinical isolates of emerging human pathogen Mycobacterium abscessus reveals high evolutionary potential. Sci Rep 4: 4061.

20. **Pawlik A, Garnier G, Orgeur M, Tong P, Lohan A, et al.** (2013) Identification and characterization of the genetic changes responsible for the characteristic smooth-to-rough morphotype alterations of clinically persistent Mycobacterium abscessus. Mol Microbiol 90: 612–629.

21. **Joseph J, Rajendran V, Hassan S, Kumar V** (2011) Mycobacteriophage genome database. Bioinformation 6: 393–394.

22. **Casali N, Riley LW** (2007) A phylogenomic analysis of the Actinomycetales mce operons. BMC genomics 8: 60.

23. **Arruda S, Bomfim G, Knights R, Huima-Byron T, Riley LW** (1993) Cloning of an M. tuberculosis DNA fragment associated with entry and survival inside cells. Science 261: 1454–1457.

24. **Pandey AK, Sassetti CM** (2008) Mycobacterial persistence requires the utilization of host cholesterol. Proceedings of the National Academy of Sciences of the United States of America 105: 4376–4380.

25. **Klepp LI, Forrellad MA, Osella AV, Blanco FC, Stella EJ, et al.** (2012) Impact of the deletion of the six mce operons in Mycobacterium smegmatis. Microbes and infection/Institut Pasteur 14: 590–599.

26. **Kim SH, Harzman C, Davis JK, Hutcheson R, Broderick JB, et al.** (2012) Genome sequence of Desulfitobacterium hafniense DCB-2, a Gram-positive anaerobe capable of dehalogenation and metal reduction. BMC microbiology 12: 21.

27. **Suzuki R, Katayama T, Kim BJ, Wakagi T, Shoun H, et al.** (2010) Crystal structures of phosphoketolase: thiamine diphosphate-dependent dehydration mechanism. The Journal of biological chemistry 285: 34279–34287.

28. **Boes N, Schreiber K, Hartig E, Jaensch L, Schobert M** (2006) The Pseudomonas aeruginosa universal stress protein PA4352 is essential for surviving anaerobic energy stress. Journal of bacteriology 188: 6529–6538.

29. **Klotz T, Vorreuther R, Heidenreich A, Zumbe J, Engelmann U** (1996) Testicular tissue oxygen pressure. The Journal of urology 155: 1488–1491.

30. **Shahidi M, Wanek J, Blair NP, Little DM, Wu T** (2010) Retinal tissue oxygen tension imaging in the rat. Investigative ophthalmology & visual science 51: 4766–4770.

31. **Wang W, Vadgama P** (2004) O2 microsensors for minimally invasive tissue monitoring. Journal of the Royal Society, Interface/the Royal Society 1: 109–117.

32. **Ponce LL, Pillai S, Cruz J, Li X, Julia H, et al.** (2012) Position of probe determines prognostic information of brain tissue PO2 in severe traumatic brain injury. Neurosurgery 70: 1492–1502; discussion 1502–1493.

33. **Barkan D, Rao V, Sukenick GD, Glickman MS** (2010) Redundant function of cmaA2 and mmaA2 in Mycobacterium tuberculosis cis cyclopropanation of oxygenated mycolates. Journal of bacteriology 192: 3661–3668.

34. **Banerjee R, Vats P, Dahale S, Kasibhatla SM, Joshi R** (2011) Comparative genomics of cell envelope components in mycobacteria. PloS one 6: e19280.

35. **Adekambi T, Reynaud-Gaubert M, Greub G, Gevaudan MJ, La Scola B, et al.** (2004) Amoebal coculture of "Mycobacterium massiliense" sp. nov. from the sputum of a patient with hemoptoic pneumonia. Journal of clinical microbiology 42: 5493–5501.

36. **Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, et al.** (2009) ABySS: a parallel assembler for short read sequence data. Genome research 19: 1117–1123.

37. **Li H, Ruan J, Durbin R** (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome research 18: 1851–1858.

38. **Bao H, Guo H, Wang J, Zhou R, Lu X, et al.** (2009) MapView: visualization of short reads alignment on a desktop computer. Bioinformatics 25: 1554–1555.

39. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, et al.** (2008) The RAST Server: rapid annotations using subsystems technology. BMC genomics 9: 75.

40. **Jones P, Binns D, Chang HY, Fraser M, Li W, et al.** (2014) InterProScan 5: genome-scale protein function classification. Bioinformatics 30: 1236–1240.

41. **Petkau A, Stuart-Edwards M, Stothard P, Van Domselaar G** (2010) Interactive microbial genome visualization with GView. Bioinformatics 26: 3125–3126.

42. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** (1990) Basic local alignment search tool. Journal of molecular biology 215: 403–410.

43. **Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, et al.** (2005) ACT: the Artemis Comparison Tool. Bioinformatics 21: 3422–3423.

44. **Rice P, Longden I, Bleasby A** (2000) EMBOSS: the European Molecular Biology Open Software Suite. Trends in genetics: TIG 16: 276–277.

45. **Earl D, Bradnam K, St John J, Darling A, Lin D, et al.** (2011) Assemblathon 1: a competitive assessment of de novo short read assembly methods. Genome research 21: 2224–2241.

46. **Li H, Durbin R** (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26: 589–595.

47. **Li H** (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987–2993.

48. **Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, et al.** (2009) VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25: 2283–2285.

49. **Stamatakis A** (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.

50. **Katoh K, Toh H** (2010) Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics 26: 1899–1900.

51. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al.** (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948.