

What is claimed is:

1. A method for distinguishing between active and latent *Mycobacterium tuberculosis* infection in a patient suspected of being infected with *Mycobacterium tuberculosis*, the method comprising:
obtaining a gene expression dataset from a whole blood sample from the patient;
- 5 determining the differential expression of one or more transcriptional gene expression modules that distinguish between infected patients and non-infected individuals, wherein the dataset demonstrates an aggregate change in the levels of polynucleotides in the one or more transcriptional gene expression modules as compared to matched non-infected individuals, and
distinguishing between active and latent *Mycobacterium tuberculosis* (TB) infection based on the one or
10 more transcriptional gene expression modules that differentiate between active and latent infection.
2. The method of claim 1, further comprising the step of using the determined comparative gene product information to formulate a diagnosis.
3. The method of claim 1, further comprising the step of using the determined comparative gene product information to formulate a prognosis.
- 15 4. The method of claim 1, further comprising the step of using the determined comparative gene product information to formulate a treatment plan.
5. The method of claim 1, further comprising the step of distinguishing patients with latent TB from active TB patients.
6. The method of claim 1, wherein the module comprises a dataset of the genes in modules M1.2, M1.3,
20 M1.4, M1.5, M1.8, M2.1, M2.4, M2.8, M3.1, M3.2, M3.3, M3.4, M3.6, M3.7, M3.8 or M3.9 to detect active pulmonary infection.
7. The method of claim 1, wherein the module comprises a dataset of the genes in modules M1.5, M2.1, M2.6, M2.10, M3.2 or M3.3 to detect a latent infection.
8. The method of claim 1, wherein the following genes are down-regulated in active pulmonary
25 infection CD3, CTLA-4, CD28, ZAP-70, IL-7R, CD2, SLAM, CCR7 and GATA-3.
9. The method of claim 1, wherein the expression profile of Figure 9 is indicative of active pulmonary infection.
10. The method of claim 1, wherein the expression profile of Figure 10 is indicative of latent infection.
11. The method of claim 1, wherein the underexpression of genes in modules M3.4, M3.6, M3.7, M3.8
30 and M3.9 is indicative of active infection.

12. The method of claim 1, wherein the overexpression of genes in modules M3.1 is indicative of active infection.

13. The method of claim 1, further comprising the step of distinguishing TB infection from other bacterial infections by determining the gene expression in modules M2.2, M2.3 and M3.5, which are overexpressed by the peripheral blood mononuclear cells or whole blood in infection other than *Mycobacterium*.

14. The method of claim 1, further comprising the step of distinguishing the differential and reciprocal transcriptional signatures in the blood of latent and active TB patients using two or more of the following modules: M1.3, M1.4, M1.5, M1.8, M2.1, M2.4, M2.8, M3.1, M3.2, M3.3, M3.4, M3.6, M3.7, M3.8 or M3.9 for active pulmonary infection and modules M1.5, M2.1, M2.6, M2.10, M3.2 or M3.3 for a latent infection.

15. The method of claim 1, wherein the genes that are upregulated in active pulmonary TB infection versus a healthy patient are selected from Tables 7A, 7D, 7I, 7J and 7K.

16. The method of claim 1, wherein the genes that are downregulated in active pulmonary TB infection versus a healthy patient are selected from Tables 7B, 7C, 7E, 7F, 7G, 7H, 7L, 7M, 7N, 7O and 7P.

17. The method of claim 1, wherein the genes that are upregulated in latent TB infection versus a healthy patient are selected from Table 8B.

18. The method of claim 1, wherein the genes that are downregulated in latent TB infection versus a healthy patient are selected from Tables 8A, 8C, 8D, 8E and 8F.

19. A method for distinguishing between active and latent *Mycobacterium tuberculosis* infection in a patient suspected of being infected with *Mycobacterium tuberculosis*, the method comprising:

obtaining a first gene expression dataset obtained from a first clinical group with active *Mycobacterium tuberculosis* infection, a second gene expression dataset obtained from a second clinical group with a latent *Mycobacterium tuberculosis* infection patient and a third gene expression dataset obtained from a clinical group of non-infected individuals;

generating a gene cluster dataset comprising the differential expression of genes between any two of the first, second and third datasets; and

determining a unique pattern of expression/representation that is indicative of latent infection, active infection or being healthy.

20. The method of claim 19, wherein each clinical group is separated into a unique pattern of expression/representation for each of the 119 genes of Table 6.

21. The method of claim 19, wherein values for the first and third datasets are compared and the values for the dataset from the third dataset are subtracted therefrom.
22. The method of claim 19, wherein values for the second and third datasets are compared and the values for the dataset from the third dataset are subtracted therefrom.
- 5 23. The method of claim 19, further comprising the step of comparing values for two different datasets and subtracting the values for the remaining dataset to distinguish between a patient with a latent infection, a patient with an active infection and a non-infected individual.
24. The method of claim 19, further comprising the step of using the determined comparative gene product information to formulate a diagnosis or a prognosis.
- 10 25. The method of claim 19, further comprising the step of using the determined comparative gene product information to formulate a treatment plan.
26. The method of claim 19, further comprising the step of distinguishing patients with latent TB from active TB patients.
27. The method of claim 19, further comprising of determining the expression levels of the genes:
15 ST3GAL6, PAD14, TNFRSF12A, VAMP3, BR13, RGS19, PILRA, NCF1, LOC652616, PLAUR(CD87), SIGLEC5, B3GALT7, IBRD3(NKLM), ALOX5AP(FLAP), MMP9, ANPEP(APN), NALP12, CSF2RA, IL6R(CD126), RASGRP4, TNFSF14(CD258), NCF4, HK2, ARID3A, PGLYRP1(PGRP), which are underexpressed/underrepresented in the blood of Latent TB patients but not in the blood of Healthy individuals or Active TB patients.
- 20 28. The method of claim 19, further comprising of determining the expression levels of the genes: ABCG1, SREBF1, RBP7(CRBP4), C22orf5, FAM101B, S100P, LOC649377, UBTD1, PSTPIP-1, RENBP, PGM2, SULF2, FAM7A1, HOM-TES-103, NDUFAF1, CES1, CYP27A1, FLJ33641, GPR177, MID1IP1(MIG-12), PSD4, SF3A1, NOV(CCN3), SGK(SGK1), CDK5R1, LOC642035, which are
25 overexpressed/overrepresented in the blood of Healthy control individuals but were underexpressed/underrepresented in the blood of Latent TB patients, and underexpressed/underrepresented in the blood of Active TB patients.
29. The method of claim 19, further comprising of determining the expression levels of the genes: ARSG, LOC284757, MDM4, CRNKL1, IL8, LOC389541, CD300LB, NIN, PHKG2, HIP1, which are
30 overexpressed/overrepresented in the blood of Healthy individuals, are underexpressed/underrepresented in the blood of both Latent and Active TB patients.
30. The method of claim 19, further comprising of determining the expression levels of the genes: PSMB8(LMP7), APOL6, GBP2, GBP5, GBP4, ATF3, GCH1, VAMP5, WARS, LIMK1, NPC2, IL-15,

LMTK2, STX11(FHL4), which are overexpressed/overrepresented in the blood of Active TB, and underexpressed/underrepresented in the blood of Latent TB patients and Healthy control individuals.

31. The method of claim 19, further comprising of determining the expression levels of the genes: FLJ11259(DRAM), JAK2, GSDMDC1(DF5L)(FKSG10), SIPAIL1, [2680400](KIAA1632),
5 ACTA2(ACTSA), KCNMB1(SLO-BETA), which are overexpressed/overrepresented in blood from Active TB patients, and underexpressed/underrepresented in the blood from Latent TB patients and Healthy control individuals.

32. The method of claim 19, further comprising of determining the expression levels of the genes: SPTANI, KIAAD179(Nnp1)(RRP1), FAM84B(NSE2), SELM, IL27RA, MRPS34, [6940246](IL23A),
10 PRKCA(PKCA), CCDC41, CD52(CDW52), [3890241](ZN404), MCCC1(MCCA/B), SOX8, SYNJ2, FLJ21127, FHIT, which are underexpressed/underrepresented in the blood of Active TB patients but not in the blood of Latent TB patients or Healthy Control individuals.

33. The method of claim 19, further comprising of determining the expression levels of the genes: CDKL1(p42), MICALCL, MBNL3, RHD, ST7(RAY1), PPR3R1, [360739](PIP5K2A), AMFR, FLJ22471,
15 CRAT(CAT1), PLA2G4C, ACOT7(ACT)(ACH1), RNF182, KLRC3(NKG2E), HLA-DPB1, which are underexpressed/underrepresented in the blood of Healthy Control individuals, overexpressed/overrepresented in the blood of the Latent TB patients, and overexpressed/overrepresented in the blood of Active TB patients.

34. A method for distinguishing between active and latent mycobacterium tuberculosis infection in a patient suspected of being infected with *Mycobacterium tuberculosis*, the method comprising:

20 obtaining a gene expression dataset from a whole blood sample;

sorting the gene expression dataset into one or more transcriptional gene expression modules; and

mapping the differential expression of the one or more transcriptional gene expression modules that distinguish between active and latent *Mycobacterium tuberculosis* infection, thereby distinguishing between active and latent *Mycobacterium tuberculosis* infection.

25 35. The method of claim 34, wherein the dataset comprises TRIM genes.

36. The method of claim 34, wherein the dataset comprises TRIM genes, and TRIM 5, 6, 19(PML), 21, 22, 25, 68 are overrepresented/expressed in active pulmonary TB.

37. The method of claim 34, wherein the dataset comprises TRIM genes, and TRIM 28, 32, 51, 52, 68, are underrepresented/expressed in active pulmonary TB.

30 38. A method of diagnosing a patient with active and latent *Mycobacterium tuberculosis* infection in a patient suspected of being infected with mycobacterium tuberculosis, the method comprising detecting differential expression of one or more transcriptional gene expression modules that distinguish between

infected and non-infected patients obtained from whole blood, wherein whole blood demonstrates an aggregate change in the levels of polynucleotides in the one or more transcriptional gene expression modules as compared to matched non-infected patients, thereby distinguishing between active and latent mycobacterium tuberculosis infection.

5 39. The method of claim 38, further comprising the step of using the determined comparative gene product information to formulate a diagnosis.

40. The method of claim 38, further comprising the step of using the determined comparative gene product information to formulate a prognosis.

10 41. The method of claim 38, further comprising the step of using the determined comparative gene product information to formulate a treatment plan.

42. The method of claim 38, wherein the module comprises a dataset of the genes in modules M1.2, M1.3, M1.4, M1.5, M1.8, M2.1, M2.4, M2.8, M3.1, M3.2, M3.3, M3.4, M3.6, M3.7, M3.8 or M3.9 to detect active pulmonary infection.

15 43. The method of claim 38, wherein the module comprises a dataset of the genes in modules M1.5, M2.1, M2.6, M2.10, M3.2 or M3.3 to detect a latent infection.

44. The method of claim 38, wherein the following genes are down-regulated in active pulmonary infection CD3, CTLA-4, CD28, ZAP-70, IL-7R, CD2, SLAM, CCR7 and GATA-3.

45. The method of claim 38, wherein the expression profile of modules of Figure 9 is diagnostic of active pulmonary infection.

20 46. The method of claim 38, wherein the expression profile of modules of Figure 10 is diagnostic of latent infection.

47. The method of claim 38, wherein the underexpression of genes in modules M3.4, M3.6, M3.7, M3.8 and M3.9 is indicative of active infection.

25 48. The method of claim 38, wherein the overexpression of genes in modules M3.1 is indicative of active infection.

49. The method of claim 38, further comprising the step of distinguishing TB infection from other bacterial infections by determining the gene expression in modules M2.2, M2.3 and M3.5, which are overexpressed by the peripheral blood mononuclear cells or whole blood in infection other than Mycobacterium.

30 50. The method of claim 38, further comprising the step of distinguishing the differential and reciprocal transcriptional signatures in the blood of latent and active TB patients using two or more of the following

modules: M1.3, M1.4, M1.5, M1.8, M2.1, M2.4, M2.8, M3.1, M3.2, M3.3, M3.4, M3.6, M3.7, M3.8 or M3.9 for active pulmonary infection and modules M1.5, M2.1, M2.6, M2.10, M3.2 or M3.3 for a latent infection.

51. A kit for diagnosing a patient with active and latent mycobacterium tuberculosis infection in a patient suspected of being infected with *Mycobacterium tuberculosis*, the kit comprising:

- 5 a gene expression detector for obtaining a gene expression dataset from the patient; and
a processor capable of comparing the gene expression to pre-defined gene module dataset that distinguish between infected and non-infected patients obtained from whole blood, wherein whole blood demonstrates an aggregate change in the levels of polynucleotides in the one or more transcriptional gene expression modules as compared to matched non-infected patients, thereby distinguishing between active and latent
10 *Mycobacterium tuberculosis* infection.

52. A system of diagnosing a patient with active and latent *Mycobacterium tuberculosis* infection comprising:

- a gene expression dataset from the patient; and
a processor capable of comparing the gene expression to pre-defined gene module dataset that distinguish
15 between infected and non-infected patients obtained from whole blood, wherein whole blood demonstrates an aggregate change in the levels of polynucleotides in the one or more transcriptional gene expression modules as compared to matched non-infected patients, thereby distinguishing between active and latent
Mycobacterium tuberculosis infection, wherein the modules are selected from M1.3, M1.4, M1.5, M1.8, M2.1, M2.4, M2.8, M3.1, M3.2, M3.3, M3.4, M3.6, M3.7, M3.8 or M3.9 for active pulmonary infection and
20 modules M1.5, M2.1, M2.6, M2.10, M3.2 or M3.3 for a latent infection.

Claims

1. A method for distinguishing between active and latent *Mycobacterium tuberculosis* (TB) infection in a patient suspected of being infected with TB, the method comprising: obtaining a gene expression dataset from a whole blood sample from the patient; determining the expression pattern of two or more genes selected from the genes listed in Table 6 that distinguish between active and latent infection; and distinguishing between active and latent TB infection based on the expression pattern of the two or more genes.

2. The method of claim 1, comprising determining the expression pattern of five or more, ten or more, or 15 or more genes selected from the genes listed in Table 6.

3. The method of claim 1, comprising determining the expression pattern of the genes: ST3GAL6, PAD14, TNFRSF12A, VAMP3, BR13, RGS19, PILRA, NCF1, LOC652616, PLAUR(CD87), SIGLEC5, B3GALT7, IBRDC3(NKLAM), ALOX5AP(FLAP), MMP9, ANPEP(APN), NALP12, CSF2RA, IL6R(CD126), RASGRP4, TNFSF14(CD258), NCF4, HK2, ARID3A and PGLYRP1(PGRP), which are underexpressed/underrepresented in the blood of latent TB patients but not in the blood of healthy individuals or active TB patients.

4. The method of claim 1, comprising determining the expression pattern of the genes: PSMB8(LMP7), APOL6, GBP2, GBP5, GBP4, ATF3, GCH1, VAMP5, WARS, LIMK1, NPC2, IL-15, LMTK2 and STX11(FHL4), which are overexpressed/overrepresented in the blood of active TB, and underexpressed/underrepresented in the blood of latent TB patients and healthy control individuals.

5. The method of claim 1, comprising determining the expression pattern of the genes: FLJ11259(DRAM), JAK2, GSDMDC1(DF5L)(FKSG10), SIPAIL1, [2680400](KIAA1632), ACTA2(ACTSA) and KCNMB1(SLO-BETA), which are overexpressed/overrepresented in blood from active TB patients, and underexpressed/underrepresented in the blood from latent TB patients and healthy control individuals.

6. The method of claim 1, comprising determining the expression pattern of the genes: SPTANI, KIAAD179(Nnp1)(RRP1), FAM84B(NSE2), SELM, IL27RA, MRPS34, [6940246](IL23A), PRKCA(PKCA), CCDC41, CD52(CDW52), [3890241](ZN404), MCCC1(MCCA/B), SOX8, SYNJ2, FLJ21127 and FHIT, which are underexpressed/underrepresented in the blood of active TB patients but not in the blood of latent TB patients or healthy control individuals.