

DOCUMENT MADE AVAILABLE UNDER THE PATENT COOPERATION TREATY (PCT)

International application number:	PCT/US2022/018313
International filing date:	01 March 2022 (01.03.2022)
Document type:	Certified copy of priority document
Document details:	Country/Office: US
	Number: 63/154,821
	Filing date: 01 March 2021 (01.03.2021)
Date of receipt at the International Bureau:	13 March 2022 (13.03.2022)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a),(b) or (b-bis)

1194843

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

March 12, 2022

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 63/154,821

FILING DATE: *March 01, 2021*

RELATED PCT APPLICATION NUMBER: *PCT/US22/18313*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US63/154,821*



Certified by

Anders Iversen

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 1 of 2

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Priority Mail Express® Label No. N/A

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
S. Laura	Chang	Verona, WI
Lori J.	Pierce	Ann Arbor, MI
Corey	Speers	Dexter, MI
Felix	Feng	Hillsborough, CA
Per	Malmstrom	Skanor, SE
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto.		
TITLE OF THE INVENTION (500 characters max):		
METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		
Direct all correspondence to: CORRESPONDENCE ADDRESS		
☒ The address corresponding to Customer Number: 119008		
OR		
☐ Firm or ☐ Individual Name		
Address		
City	State	Zip
Country	Telephone	Email
ENCLOSED APPLICATION PARTS (check all that apply)		
☒ Application Data Sheet. See 37 CFR 1.76. ☐ CD(s), Number of CDs _____		
☒ Drawing(s) Number of Sheets <u>4</u> ☐ Other (specify) _____		
☒ Specification (e.g., description of the invention) Number of Pages <u>48</u>		
Fees Due: Filing Fee of \$300 (\$150 for small entity) (\$75 for micro entity). If the specification and drawings exceed 100 sheets of paper, an application size fee is also due, which is \$420 (\$210 for small entity) (\$105 for micro entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).		
METHOD OF PAYMENT OF THE FILING FEE AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT		
☒ Applicant asserts small entity status. See 37 CFR 1.27.		
☐ Applicant certifies micro entity status. See 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.		
☐ A check or money order made payable to the <i>Director of the United States Patent and Trademark Office</i> is enclosed to cover the filing fee and application size fee (if applicable).		
☒ Payment by credit card. Form PTO-2038 is attached.		
☐ The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: _____		
150 TOTAL FEE AMOUNT (\$)		

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 10 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 2 of 2

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government. (NOTE: Providing this information on a provisional cover sheet, such as this Provisional Application for Patent Cover Sheet (Form PTO/SB/16), does not satisfy the requirement of 35 U.S.C. 202(c)(6), which requires that the *specification* contain a statement specifying that the invention was made with Government support and that the Government has certain rights in the invention.)



No.



Yes, the invention was made by an agency of the U.S. Government. The U.S. Government agency name is:



Yes, the invention was made under a contract with an agency of the U.S. Government.

The contract number is: _____

The U.S. Government agency name is: _____

In accordance with 35 U.S.C. 202(c)(6) and 37 CFR 401.14(f)(4), the specifications of any United States patent applications and any patent issuing thereon covering the invention, including the enclosed provisional application, must state the following:

“This invention was made with government support under [IDENTIFY THE CONTRACT] awarded by [IDENTIFY THE FEDERAL AGENCY]. The government has certain rights in the invention.”

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

SIGNATURE /Christopher T. Jacob/ DATE February 28, 2021TYPED OR PRINTED NAME Christopher T. Jacob REGISTRATION NO. 59938
(if appropriate)TELEPHONE 858-245-0139 DOCKET NUMBER PF103 PRV

Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Inventors listing continued

Inventor(s)		
Given Name	Family Name	Residence
Marten	Ferno	Limhamn, SE
Erik	Holmberg	Aspero, SE
Per O.	Karlsson	Gothenburg, SE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PF103 PRV
		Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		
The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76. This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.			

Secrecy Order 37 CFR 5.2:

☐	Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)
--------------------------	---

Inventor Information:

Inventor	1				Remove	
Legal Name						
Prefix	Given Name	Middle Name	Family Name	Suffix		
	S.	Laura	Chang			
Residence Information (Select One) ☒ US Residency ☐ Non US Residency ☐ Active US Military Service						
City	Verona	State/Province	WI	Country of Residence	US	
Mailing Address of Inventor:						
Address 1	818 Pine Hill Drive					
Address 2						
City	Verona	State/Province	WI			
Postal Code	53593	Country i	US			
Inventor	2				Remove	
Legal Name						
Prefix	Given Name	Middle Name	Family Name	Suffix		
	Lori	J	Pierce			
Residence Information (Select One) ☒ US Residency ☐ Non US Residency ☐ Active US Military Service						
City	Ann Arbor	State/Province	MI	Country of Residence	US	
Mailing Address of Inventor:						
Address 1	1601 Newport Creek Drive					
Address 2						
City	Ann Arbor	State/Province	MI			
Postal Code	48103	Country i	US			
Inventor	3				Remove	
Legal Name						
Prefix	Given Name	Middle Name	Family Name	Suffix		
	Corey		Speers			
Residence Information (Select One) ☒ US Residency ☐ Non US Residency ☐ Active US Military Service						

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number		PF103 PRV	
		Application Number			
Title of Invention		METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY			

City	Dexter	State/Province	MI	Country of Residence ⁱ	US
------	--------	----------------	----	-----------------------------------	----

Mailing Address of Inventor:

Address 1	3000 Walsh Rd				
Address 2					
City	Dexter	State/Province	MI		
Postal Code	48130	Country ⁱ	US		
Inventor	4				Remove
Legal Name					
Prefix	Given Name	Middle Name	Family Name	Suffix	
	Felix		Feng		
Residence Information (Select One) ☒ US Residency ☐ Non US Residency ☐ Active US Military Service					
City	Hillsborough	State/Province	CA	Country of Residence ⁱ	US

Mailing Address of Inventor:

Address 1	332 Darrell Rd				
Address 2					
City	Hillsborough	State/Province	CA		
Postal Code	94010	Country ⁱ	US		
Inventor	5				Remove
Legal Name					
Prefix	Given Name	Middle Name	Family Name	Suffix	
	Per		Malmström		
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service					
City	Skånör	Country of Residence ⁱ			SE

Mailing Address of Inventor:

Address 1	Östergatan 1				
Address 2					
City	Skånör	State/Province	SE		
Postal Code	239 30	Country ⁱ	SE		
Inventor	6				Remove
Legal Name					
Prefix	Given Name	Middle Name	Family Name	Suffix	
	Mårten		Fernö		
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service					

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PF103 PRV
		Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		

City	Limhamn	Country of Residence ⁱ	SE
------	---------	-----------------------------------	----

Mailing Address of Inventor:

Address 1	Sallysgata 33		
Address 2			
City	Limhamn	State/Province	SE
Postal Code	216 44	Country ⁱ	SE
Inventor	7	Remove	

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Erik		Holmberg	

Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service

City	Asperö	Country of Residence ⁱ	SE
------	--------	-----------------------------------	----

Mailing Address of Inventor:

Address 1	Körevägen 12		
Address 2			
City	Asperö	State/Province	SE
Postal Code	430 80	Country ⁱ	SE
Inventor	8	Remove	

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Per	O.	Karlsson	

Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service

City	Gothenburg	Country of Residence ⁱ	SE
------	------------	-----------------------------------	----

Mailing Address of Inventor:

Address 1	Welandergatan 12b		
Address 2			
City	Gothenburg	State/Province	SE
Postal Code	416 56	Country ⁱ	SE

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button.[Add](#)

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PF103 PRV
		Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below. For further information see 37 CFR 1.33(a).			
☐ An Address is being provided for the correspondence information of this application.			
Customer Number	119008		
Email Address	chris@jacoblawoffice.com	Add Email	Remove Email
Email Address	cljacob@gmail.com	Add Email	Remove Email

Application Information:

Title of the Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		
Attorney Docket Number	PF103 PRV	Small Entity Status Claimed ☒	
Application Type	Provisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	4	Suggested Figure for Publication (if any)	

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)
☐ Request Not to Publish. I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application has not and will not be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.
--

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PF103 PRV
		Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		
Please Select One:	☒ Customer Number	☐ US Patent Practitioner	☐ Limited Recognition (37 CFR 11.9)
Customer Number	119008		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.

When referring to the current application, please leave the "Application Number" field blank.

Prior Application Status		Remove
Application Number	Continuity Type	Prior Application Number
Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the Add button.		Add

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX)ⁱ the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Application Number	Country ⁱ	Filing Date (YYYY-MM-DD)	Access Code ^j (if applicable)
Additional Foreign Priority Data may be generated within this form by selecting the Add button.			Add

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

- ☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.
- NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PF103 PRV
		Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PF103 PRV
		Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Applicant	1	Remove	
If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section. Clear			
Assignee	Legal Representative under 35 U.S.C. 117	Joint Inventor	
☒ Person to whom the inventor is obligated to assign.		☐ Person who shows sufficient proprietary interest	
If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:			
Name of the Deceased or Legally Incapacitated Inventor:			
If the Applicant is an Organization check here. ☒			
Organization Name	PFS Genomics Inc		
Mailing Address Information For Applicant:			
Address 1	1152 Mainland St, Suite 430		
Address 2			
City	Vancouver	State/Province	BC
Country	CA	Postal Code	V6B 4X2
Phone Number		Fax Number	
Email Address			
Additional Applicant Data may be generated within this form by selecting the Add button. Add			

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PF103 PRV
	Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY	

Assignee	1
Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.	

Remove

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

Prefix	Given Name	Middle Name	Family Name	Suffix

Mailing Address Information For Assignee including Non-Applicant Assignee:

Address 1				
Address 2				
City		State/Province		
Country i		Postal Code		
Phone Number		Fax Number		
Email Address				

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.

Add

Signature:

Remove

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). However, if this Application Data Sheet is submitted with the **INITIAL** filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Christopher Jacob/		Date (YYYY-MM-DD)	2021-02-28
First Name	Christopher	Last Name	Jacob	Registration Number
				59938

Additional Signature may be generated within this form by selecting the Add button.

Add

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PF103 PRV
		Application Number	
Title of Invention	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY		

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

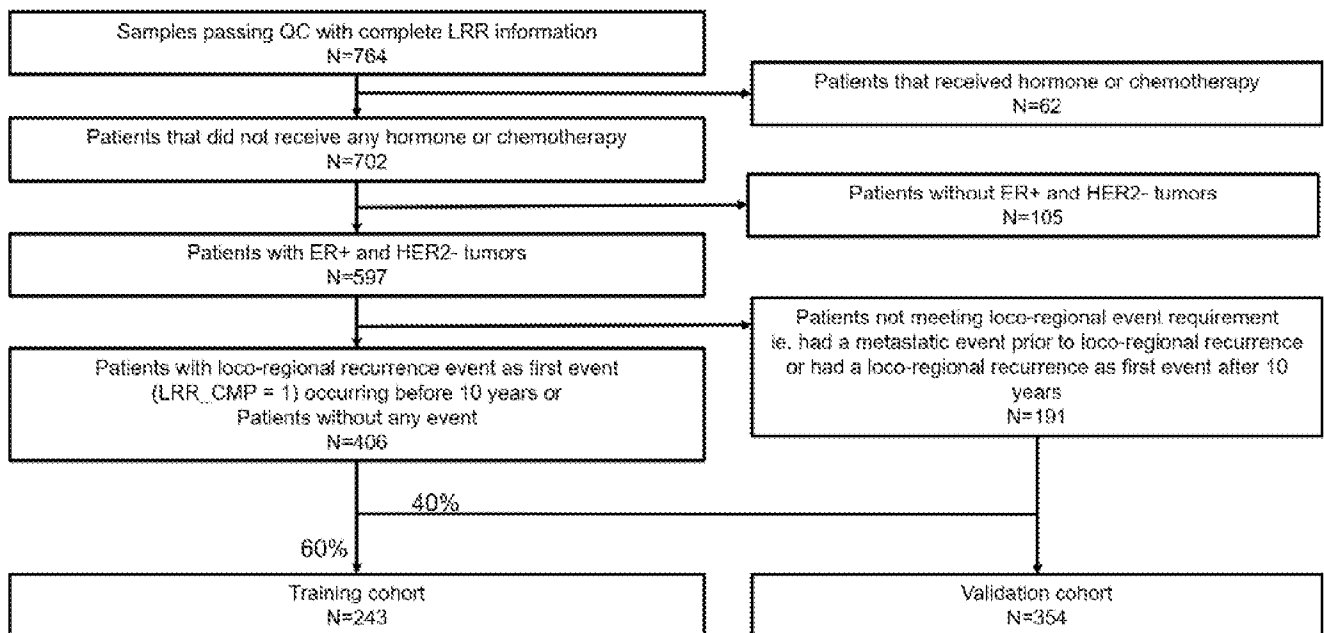


FIG. 1

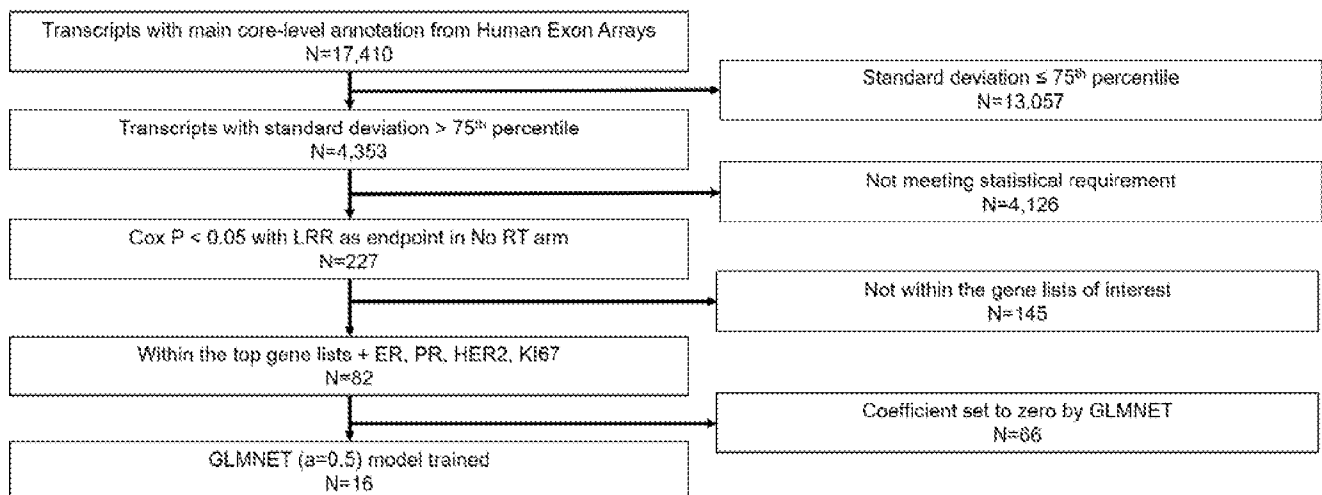


FIG. 2

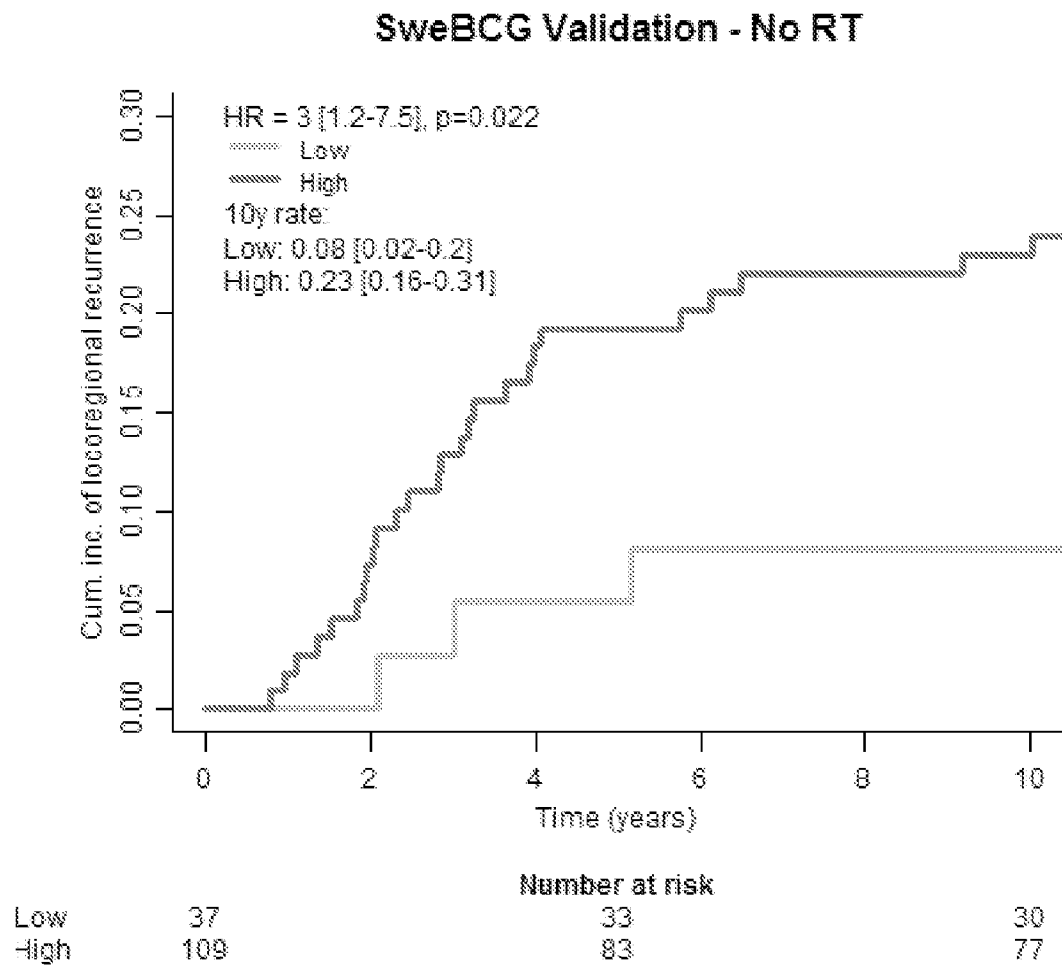
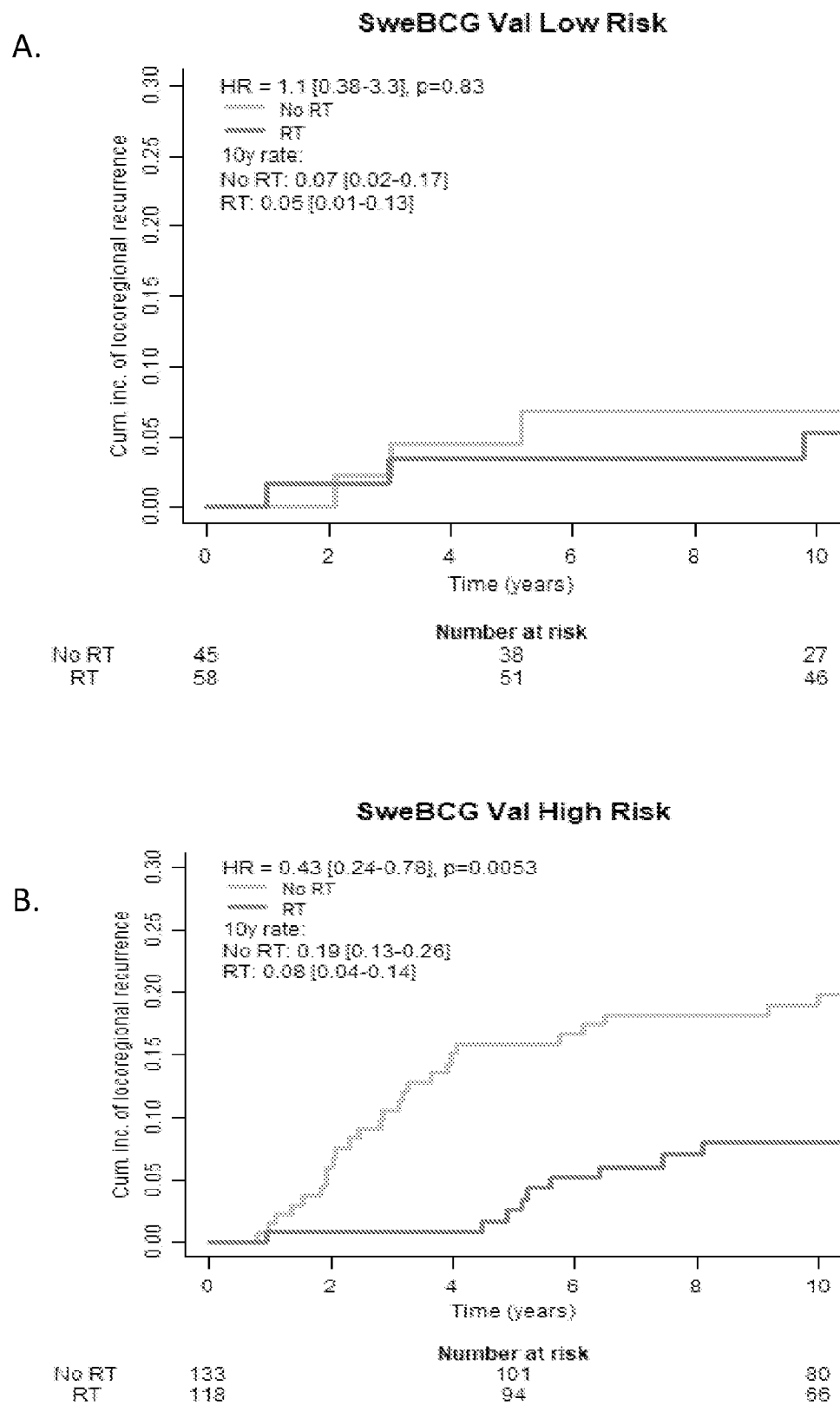


FIG. 3



FIGS. 4A-4B

Electronic Acknowledgement Receipt

EFS ID:	42046017
Application Number:	63154821
International Application Number:	
Confirmation Number:	9554
Title of Invention:	METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY
First Named Inventor/Applicant Name:	S. Laura Chang
Customer Number:	119008
Filer:	Christopher T. Jacob
Filer Authorized By:	
Attorney Docket Number:	PF103 PRV
Receipt Date:	01-MAR-2021
Filing Date:	
Time Stamp:	01:54:46
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	\$ 150
RAM confirmation Number	E202131255320816
Deposit Account	
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Provisional Cover Sheet (SB16)	PF103PRV_coversheet.pdf	367833	no	4
			2504279b706919714661f07d2c3999a40e188b16		
Warnings:					
This is not a USPTO supplied Provisional Cover Sheet SB16 form.					
Information:					
2	Application Data Sheet	PF103PRV_ADS.pdf	1256890	no	10
			1149ef1add9c961a65dab144e4368a300dd11b28		
Warnings:					
Information:					
3	Specification	PF103PRV_SPECIFICATION_03_01_21.pdf	498688	no	48
			28686b70933dc83edb14b22d4ed5164467f32539		
Warnings:					
Information:					
4	Drawings-only black and white line drawings	PF103PRV_FIGS_03_01_21.pdf	130686	no	4
			e5e40e56058860d7e2f166627671954c5dc0bf90		
Warnings:					
Information:					
5	Fee Worksheet (SB06)	fee-info.pdf	29974	no	2
			5f8a88bc933c96dba8cf4c15dccac30cae24d65d		
Warnings:					
Information:					
Total Files Size (in bytes):			2284071		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

METHODS AND GENOMIC CLASSIFIERS FOR PROGNOSIS OF BREAST CANCER AND IDENTIFYING SUBJECTS NOT LIKELY TO BENEFIT FROM RADIOTHERAPY

FIELD OF THE INVENTION

[0001] The present disclosure pertains to the field of personalized medicine and methods for treating breast cancer. In particular, the disclosure relates to the use of transcriptomic profiling for prognosis of cancer recurrence (e.g., locoregional recurrence) in subjects with breast cancer and methods for identifying subjects who are not likely to benefit from radiotherapy.

BACKGROUND OF THE INVENTION

[0002] Cancer is the uncontrolled growth of abnormal cells anywhere in a body. The abnormal cells are termed cancer cells, malignant cells, or tumor cells. Many cancers and the abnormal cells that compose the cancer tissue are further identified by the name of the tissue that the abnormal cells originated from (for example, breast cancer). Cancer cells can proliferate uncontrollably and form a mass of cancer cells. Cancer cells can break away from this original mass of cells, travel through the blood and lymph systems, and lodge in other organs where they can again repeat the uncontrolled growth cycle. This process of cancer cells leaving an area and growing in another body area is often termed metastatic spread or metastatic disease. For example, if breast cancer cells spread to a bone (or anywhere else), it can mean that the individual has metastatic breast cancer.

Most women with early-stage invasive breast cancer are operated with breast-conserving surgery (BCS) and adjuvant post-operative radiation therapy (RT). While RT provides a significant reduction in local recurrence risk and increased breast cancer-specific survival, better treatment selection is needed. Meta-analyses from the Early Breast Cancer Trialists' Collaborative Group have shown that at 10 years post-surgery, 70% of women with stage I-II node-negative breast cancer are free from local recurrence despite not receiving RT, and 10% experience a local recurrence even with the receipt of RT, thus suggesting both over- and under treatment. (Darby S et al. *Lancet* 2011; **378**(9804): 1707-16.)

[0003] Several clinico-pathologic risk factors have been described for local recurrence, where young age is one of the strongest, and younger patients are commonly given a radiation boost. (Fredriksson I et al. *The British journal of surgery* 2003; **90**(9): 1093-102 and Fredriksson I et al. *European journal of cancer (Oxford, England: 1990)* 2001; **37**(12): 1537-44.) However, with more modern surgical and oncological treatment, the rate of local recurrence has considerably decreased, attributed in part to better surgery and better adjuvant treatment. (Bouganin N et al. *Breast cancer research and*

treatment 2013; **139**(2): 603-6 and Aalders KC et al. *European journal of cancer (Oxford, England: 1990)* 2016; **63**: 118-26.) This has led to several trials aimed at de-escalating RT in presumed low-risk women, including the PRIMETIME, LUMINA, and NCT02653755 trials. (Wickberg A et al. *European journal of surgical oncology: the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 2018; **44**(7): 951-6 and Kirwan CC et al. *Clinical oncology (Royal College of Radiologists (Great Britain))* 2016; **28**(9): 594-6.) However, there remains a need for improved methods for identifying which breast cancer subjects would not likely benefit from radiotherapy. Further, there is a need for novel gene signatures that are prognostic for loco-regional recurrence and radiation sensitivity.

[0004] This background information is provided for the purpose of making known information believed by the applicant to be of possible relevance to the present disclosure. No admission is necessarily intended, nor should be construed, that any of the preceding information constitutes prior art against the present disclosure.

SUMMARY OF THE INVENTION

[0005] The present disclosure relates to methods, systems, and kits for the diagnosis, prognosis, and treatment of breast cancer in a subject. The disclosure also provides biomarkers and classifiers for identifying subjects at low risk of breast cancer recurrence and not likely to benefit from adjuvant radiotherapy. Further disclosed herein, in certain instances, are probe sets for use in detecting such biomarkers for determining the risk of breast cancer recurrence in a subject. The disclosure further provides biomarkers and classifiers for identifying subjects at risk for locoregional recurrence (LRR) and predicting response to radiotherapy. Methods of treating breast cancer based on expression profiling and/or age to determine the risk of breast cancer recurrence are also provided.

[0006] In some embodiments, the disclosure provides a method for prognosing and/or predicting benefit from adjuvant radiotherapy in a subject having breast cancer, the method comprising: a) obtaining or having obtained an expression level in a sample from a subject for a plurality of genes, wherein the plurality of genes are selected from Table 5; and b) determining that the subject is at low risk or high risk of cancer recurrence based on the expression level, and/or likely to benefit or not likely benefit from adjuvant radiotherapy based on the expression level, thereby prognosing and/or predicting benefit from adjuvant radiotherapy in the subject. In some embodiments, the method further comprises withholding adjuvant radiotherapy therapy if the subject is identified as not likely to benefit from adjuvant radiotherapy and/or administering a cancer treatment other than adjuvant radiotherapy. In other embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis). In some embodiments, the levels of expression of one or more of the genes selected from Table 5 may be increased or decreased compared to a control. In some embodiments, the expression levels of all of the genes selected from Table 5 are measured in the

biological sample. In some embodiments, the method further comprises treating the subject with adjuvant radiotherapy. In some embodiments, the method further comprises treating the subject with a cancer therapy other than adjuvant radiotherapy. In still other embodiments, the method further comprises treating the subject with mastectomy, radiation boost, or adjuvant systemic therapy. In some embodiments, radiotherapy is withheld from the subject following breast conserving surgery (BCS). In some embodiments, the method further comprises determining that the subject is at low risk of cancer recurrence based on the age of the subject, or determining that the subject is not at low risk of cancer recurrence based on the age of the subject. In some embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis). In other embodiments, the subject is at high risk for cancer recurrence and treated with radiotherapy.

[0007] In some embodiments, the disclosure provides a method comprising: a) obtaining or having obtained an expression level in a sample from a subject for a plurality of genes, wherein the plurality of genes are selected from Table 5; and b) determining that the subject is at low risk of cancer recurrence and not likely to benefit from treatment with adjuvant radiotherapy based on the expression level, or determining that the subject is not at high risk of cancer recurrence and likely to benefit from treatment with adjuvant radiotherapy based on the expression level. In other embodiments, the method further comprises administering a cancer therapy other than adjuvant radiotherapy therapy if the subject is identified as not likely to benefit from adjuvant radiotherapy. In other embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis). In some embodiments, the levels of expression of one or more of the genes selected from Table 5 may be increased or decreased compared to a control. In some embodiments, the expression levels of all of the genes selected from Table 5 are measured in the biological sample. In still other embodiments, the method further comprises treating the subject with mastectomy, radiation boost, or adjuvant systemic therapy. In some embodiments, radiotherapy is withheld from the subject following breast conserving surgery (BCS). In some embodiments, the method further comprises determining that the subject is at low risk of cancer recurrence based on the age of the subject, or determining that the subject is not at low risk of cancer recurrence based on the age of the subject. In some embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis).

[0008] In some embodiments, the disclosure provides a method of treating breast cancer in a subject, comprising: a) obtaining or having obtained an expression level in a sample from a subject for a plurality of genes, wherein the plurality of genes are selected from Table 5; b) determining that the subject is at low risk of cancer recurrence and not likely to benefit from treatment with adjuvant radiotherapy based on the expression level; and c) administering a cancer treatment other than adjuvant radiotherapy therapy if the subject is identified as not likely to benefit from adjuvant radiotherapy based on the expression level. In other embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis). In some embodiments, the levels of

expression of one or more of the genes selected from Table 5 may be increased or decreased compared to a control. In some embodiments, the expression levels of all of the genes selected from Table 5 are measured in the biological sample. In still other embodiments, the method further comprises treating the subject with mastectomy, radiation boost, or adjuvant systemic therapy. In some embodiments, radiotherapy is withheld from the subject following breast conserving surgery (BCS). In some embodiments, the method further comprises determining that the subject is at low risk of cancer recurrence based on the age of the subject, or determining that the subject is not at low risk of cancer recurrence based on the age of the subject. In some embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis).

[0009] In some embodiments, the plurality of genes used in the methods and genomic classifiers of the present disclosure are selected from the group consisting of Anterior Gradient 2, Protein Disulphide Isomerase (AGR2), Beta-1,4-Galactosyltransferase 1 (B4GALT1), Claudin 7 (CLDN7), Ezrin (EZR), G Protein Subunit Gamma 11 (GNG11), Jun Proto Oncogene (JUN), Matrix Metalloproteinase 11 (MMP11), CAMP-Dependent Protein Kinase Inhibitor Beta (PKIB), Phosphoribosyl Pyrophosphate Synthetase 1 (PRPS1), Proteasome 26S Subunit, Non ATPase 10 (PSMD10), SH3 Domain Binding Protein 5 (SH3BP5), Solute Carrier Family 16 Member 3 (SLC16A3), Solute Carrier Family 7 Member 11 (SLC7A11), Secreted Phosphoprotein 1 (SPP1), Troponin T1, Slow Skeletal Type (TNNT1), and Ubiquitin Conjugating Enzyme E2 E1 (UBE2E1).

[0010] In certain embodiments, the subject has estrogen receptor positive (ER+) breast cancer, human epidermal growth factor receptor 2 negative (HER2-) breast cancer, Stage I-II breast cancer, or node-negative breast cancer and/or is post-menopausal.

[0011] In other embodiments, the biological sample obtained from the subject is a breast biopsy or tumor sample. In yet other embodiments, biological sample is a bodily fluid or tissue of the subject that contains breast cancer cells. In certain embodiments, nucleic acids (e.g., RNA transcripts) comprising sequences from genes selected from Table 5, or complements thereof, are further isolated from the biological sample, and/or purified, and/or amplified prior to analysis.

[0012] In other embodiments, the expression levels of biomarkers are determined by *in situ* hybridization, PCR-based methods, array-based methods, immunohistochemical methods, RNA assay methods, or immunoassay methods. In other embodiments, the levels of gene expression are determined using one or more reagents. In certain embodiments, the one or more reagents are nucleic acid probes, nucleic acid primers, and/or antibodies. In other embodiments, determining the level of expression of a biomarker comprises measuring the level of a nucleic acid (e.g., RNA transcript).

[0013] In some embodiments, the level of expression of at least one gene is reduced compared to a control. In other embodiments, the level of expression of at least one gene is increased compared to a control.

[0014] In certain embodiments, the methods described herein are performed prior to treatment of the subject with adjuvant radiotherapy. In certain embodiments, the methods described herein are

performed prior to treatment of the subject with mastectomy, radiation boost, or adjuvant systemic therapy.

[0015] In other embodiments, the method further comprises calculating a risk score for the subject, wherein adjuvant radiotherapy is withheld from the subject if the subject is identified as being at low risk of cancer recurrence and not likely to benefit from adjuvant radiotherapy based on both the risk score and the expression levels of the genes selected from Table 5 in the biological sample, and administering a cancer therapy other than adjuvant radiotherapy to the subject. In some embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis).

[0016] The significance of the expression levels of one or more biomarker genes may be evaluated using, for example, a T-test, P-value, KS (Kolmogorov Smirnov) P-value, accuracy, accuracy P-value, positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, AUC, AUC P-value (Auc.pvalue), Wilcoxon Test P-value, Median Fold Difference (MFD), Kaplan Meier (KM) curves, survival AUC (survAUC), Kaplan Meier P-value (KM P-value), Univariable Analysis Odds Ratio P-value (uvaORPval), multivariable analysis Odds Ratio P-value (mvaORPval), Univariable Analysis Hazard Ratio P-value (uvaHRPval) and Multivariable Analysis Hazard Ratio P-value (mvaHRPval). The significance of the expression level of the one or more targets may be based on two or more metrics selected from the group comprising AUC, AUC P-value (Auc.pvalue), Wilcoxon Test P-value, Median Fold Difference (MFD), Kaplan Meier (KM) curves, survival AUC (survAUC), Univariable Analysis Odds Ratio P-value (uvaORPval), multivariable analysis Odds Ratio P-value (mvaORPval), Kaplan Meier P-value (KM P-value), Univariable Analysis Hazard Ratio P-value (uvaHRPval) or Multivariable Analysis Hazard Ratio P-value (mvaHRPval).

[0017] In another aspect, the disclosure includes a probe set for determining a prognosis of a subject having breast cancer and whether or not to treat the subject with radiotherapy, the probe set comprising a plurality of probes for detecting a plurality of target nucleic acids, wherein the plurality of target nucleic acids comprises one or more gene sequences, or complements thereof, of genes selected from Table 5. Probes may be detectably labeled to facilitate detection. In some embodiments, the prognosis comprises cancer recurrence prognosis. In further embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis).

[0018] In another aspect, the disclosure includes a system for determining a prognosis of a subject who has breast cancer and whether or not to treat the subject with radiotherapy, the system comprising: a) a probe set described herein; and b) a computer model or algorithm for analyzing an expression level or expression profile of the plurality of target nucleic acids hybridized to the plurality of probes in a biological sample from a subject who has breast cancer and determining if the subject is at low risk of cancer recurrence based on the expression level or expression profile and should be treated with radiotherapy. In some embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis).

[0019] In some embodiments, the disclosure includes a kit for determining a prognosis of a subject having breast cancer and whether or not to treat the subject with adjuvant radiotherapy, the kit comprising agents for measuring levels of expression of a plurality of genes selected from Table 5. In other embodiments, the kit may include one or more agents (e.g., hybridization probes, PCR primers, or microarray) for measuring levels of expression of a plurality of genes, wherein said plurality of genes comprises one or more genes selected from Table 5, a container for holding a biological sample comprising breast cancer cells isolated from a human subject for testing, and printed instructions for reacting the agents with the biological sample or a portion of the biological sample to determine if the subject is at low risk of cancer recurrence of the breast cancer and likely to benefit from treatment with adjuvant radiotherapy. In some embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis). In other embodiments, the agents are packaged in separate containers. In yet other embodiments, the kit further comprises one or more control reference samples or other reagents for measuring gene expression (e.g., reagents for performing PCR, RT-PCR, microarray analysis, a Northern blot, an immunoassay, or immunohistochemistry). In yet other embodiments, the kit comprises agents for measuring the levels of expression of a plurality of genes listed in Table 5. In another embodiment, the kit comprises agents for measuring the levels of expression of all the genes listed in Table 5. In certain embodiments, the kit comprises a probe set, as described herein, for detecting a plurality of target nucleic acids, wherein the plurality of target nucleic acids comprises one or more gene sequences, or complements thereof, of genes selected from Table 5, or any combination thereof.

[0020] In other embodiments, the kit further comprises a system, wherein the system comprises: a) a probe set comprising a plurality of probes for detecting a plurality of target nucleic acids, wherein the plurality of target nucleic acids comprises one or more gene sequences, or complements thereof, of genes selected from Table 5; and b) a computer model or algorithm for analyzing an expression level or expression profile of the plurality of target nucleic acids hybridized to the plurality of probes in a biological sample from a subject who has breast cancer and determining if the subject is at low risk of cancer recurrence based on the expression level or expression profile and not likely to benefit from treatment with adjuvant radiotherapy. In some embodiments, the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis).

[0021] These and other embodiments of the subject disclosure will readily occur to those of skill in the art in view of the disclosure herein.

INCORPORATION BY REFERENCE

[0022] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference in their entireties to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023] FIG. 1 sets forth data showing a consort diagram for selection of training and validation cohorts in SweBCG.

[0024] FIG. 2 sets forth data showing the selection of genes for inclusion in a model of the disclosure.

[0025] FIG. 3 sets forth data showing prognostic performance of a novel genomic classifier in breast cancer subject without radio therapy and rates of locoregional recurrence.

[0026] FIGS. 4A-4B set forth data showing prognostic performance of a novel genomic classifier in low risk breast cancer subjects (4A) and high risk breast cancer subjects (4B) with and without radio therapy and rates of locoregional recurrence.

DETAILED DESCRIPTION OF THE INVENTION

[0027] The present disclosure discloses systems and methods for diagnosing, predicting, and/or monitoring the status or outcome of a breast cancer in a subject using expression-based analysis of a plurality of gene targets. Generally, the method comprises (a) obtaining or having obtained an expression level in a sample from a subject for a plurality of genes; (b) determining that the subject is at risk of cancer recurrence based on the expression level of the plurality of genes; and c) administering or withholding adjuvant radiotherapy therapy if the subject is identified as being at risk of cancer recurrence based on the expression level of the plurality of gene targets.

[0028] Assaying the expression level for a plurality of gene targets in the sample may comprise applying the sample to a microarray. In some instances, assaying the expression level may comprise the use of an algorithm. The algorithm may be used to produce a genomic classifier. Alternatively, the classifier may comprise a probe selection region. In some instances, assaying the expression level for a plurality of targets comprises detecting and/or quantifying the plurality of targets. In some embodiments, assaying the expression level for a plurality of targets comprises sequencing the plurality of targets. In some embodiments, assaying the expression level for a plurality of targets comprises amplifying the plurality of targets. In some embodiments, assaying the expression level for a plurality of targets comprises quantifying the plurality of targets. In some embodiments, assaying the expression level for a plurality of targets comprises conducting a multiplexed reaction on the plurality of targets.

[0029] Further disclosed herein are methods for determining if the subject is at low risk of recurrence of the breast cancer. Generally, the method comprises: (a) providing a sample comprising breast cancer cells from a subject; (b) assaying the expression level for a plurality of targets in the sample; and (c) determining if the subject is at low risk of recurrence of the breast cancer based on the expression level of the plurality of targets and whether or not to treat the subject with adjuvant radiotherapy, chemotherapy, or endocrine therapy. For example, a subject identified as being at low

risk of recurrence of the breast cancer according to the methods of the present disclosure, may be less likely or not likely to respond to adjuvant radiotherapy, whereas a subject identified as being at higher risk of recurrence of the breast cancer may be more likely to respond to adjuvant radiotherapy.

[0030] Before the present disclosure is described in further detail, it is to be understood that this disclosure is not limited to the particular methodology, compositions, articles or machines described, as such methods, compositions, articles or machines can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the disclosure.

Targets

[0031] In some instances, assaying the expression level of a plurality of genes comprises detecting and/or quantifying a plurality of target analytes. In some embodiments, assaying the expression level of a plurality of genes comprises sequencing a plurality of target nucleic acids. In some embodiments, assaying the expression level of a plurality of biomarker genes comprises amplifying a plurality of target nucleic acids. In some embodiments, assaying the expression level of a plurality of biomarker genes comprises conducting a multiplexed reaction on a plurality of target analytes.

[0032] The methods disclosed herein often comprise assaying the expression level of a plurality of targets. The plurality of targets may comprise coding targets and/or non-coding targets of a protein-coding gene or a non-protein-coding gene. A protein-coding gene structure may comprise an exon and an intron. The exon may further comprise a coding sequence (CDS) and an untranslated region (UTR). The protein-coding gene may be transcribed to produce a pre-mRNA and the pre-mRNA may be processed to produce a mature mRNA. The mature mRNA may be translated to produce a protein.

[0033] A non-protein-coding gene structure may comprise an exon and intron. Usually, the exon region of a non-protein-coding gene primarily contains a UTR. The non-protein-coding gene may be transcribed to produce a pre-mRNA and the pre-mRNA may be processed to produce a non-coding RNA (ncRNA).

[0034] A coding target may comprise a coding sequence of an exon. A non-coding target may comprise a UTR sequence of an exon, intron sequence, intergenic sequence, promoter sequence, non-coding transcript, CDS antisense, intronic antisense, UTR antisense, or non-coding transcript antisense. A non-coding transcript may comprise a non-coding RNA (ncRNA).

[0035] In some instances, the plurality of targets comprises one or more targets selected from Table 5. In some instances, the plurality of targets comprises at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 15, or 16 targets selected from Table 5.

[0036] In some embodiments, the plurality of targets are selected from the group consisting of Anterior Gradient 2, Protein Disulphide Isomerase (AGR2), Beta-1,4-Galactosyltransferase 1 (B4GALT1), Claudin 7 (CLDN7), Ezrin (EZR), G Protein Subunit Gamma 11 (GNG11), Jun Proto

Oncogene (JUN), Matrix Metalloproteinase 11 (MMP11), CAMP-Dependent Protein Kinase Inhibitor Beta (PKIB), Phosphoribosyl Pyrophosphate Synthetase 1 (PRPS1), Proteasome 26S Subunit, Non ATPase 10 (PSMD10), SH3 Domain Binding Protein 5 (SH3BP5), Solute Carrier Family 16 Member 3 (SLC16A3), Solute Carrier Family 7 Member 11 (SLC7A11), Secreted Phosphoprotein 1 (SPP1), Troponin T1, Slow Skeletal Type (TNNT1), and Ubiquitin Conjugating Enzyme E2 E1 (UBE2E1).

[0037] In some instances, the plurality of targets comprises a coding target, non-coding target, or any combination thereof. In some instances, the coding target comprises an exonic sequence. In other instances, the non-coding target comprises a non-exonic or exonic sequence. Alternatively, a non-coding target comprises a UTR sequence, an intronic sequence, antisense, or a non-coding RNA transcript. In some instances, a non-coding target comprises sequences which partially overlap with a UTR sequence or an intronic sequence. A non-coding target also includes non-exonic and/or exonic transcripts. Exonic sequences may comprise regions on a protein-coding gene, such as an exon, UTR, or a portion thereof. Non-exonic sequences may comprise regions on a protein-coding, non-protein-coding gene, or a portion thereof. For example, non-exonic sequences may comprise intronic regions, promoter regions, intergenic regions, a non-coding transcript, an exon anti-sense region, an intronic anti-sense region, UTR anti-sense region, non-coding transcript anti-sense region, or a portion thereof. In other instances, the plurality of targets comprises a non-coding RNA transcript.

[0038] The plurality of targets may comprise one or more targets selected from a classifier disclosed herein. The classifier may be generated from one or more models or algorithms. The one or more models or algorithms may be Naïve Bayes (NB), recursive Partitioning (Rpart), random forest (RF), support vector machine (SVM), k-nearest neighbor (KNN), high dimensional discriminant analysis (HDDA), linear model, or a combination thereof. The classifier may have an AUC of equal to or greater than 0.60. The classifier may have an AUC of equal to or greater than 0.61. The classifier may have an AUC of equal to or greater than 0.62. The classifier may have an AUC of equal to or greater than 0.63. The classifier may have an AUC of equal to or greater than 0.64. The classifier may have an AUC of equal to or greater than 0.65. The classifier may have an AUC of equal to or greater than 0.66. The classifier may have an AUC of equal to or greater than 0.67. The classifier may have an AUC of equal to or greater than 0.68. The classifier may have an AUC of equal to or greater than 0.69. The classifier may have an AUC of equal to or greater than 0.70. The classifier may have an AUC of equal to or greater than 0.75. The classifier may have an AUC of equal to or greater than 0.77. The classifier may have an AUC of equal to or greater than 0.78. The classifier may have an AUC of equal to or greater than 0.79. The classifier may have an AUC of equal to or greater than 0.80. The AUC may be clinically significant based on its 95% confidence interval (CI). The accuracy of the classifier may be at least about 70%. The accuracy of the classifier may be at least about 73%. The accuracy of the classifier may be at least about 75%. The accuracy of the classifier may be at least about 77%. The accuracy of the classifier may be at least about 80%. The accuracy of the classifier may be at least about 83%. The accuracy of the classifier may be at least about 84%. The

accuracy of the classifier may be at least about 86%. The accuracy of the classifier may be at least about 88%. The accuracy of the classifier may be at least about 90%. The p-value of the classifier may be less than or equal to 0.05. The p-value of the classifier may be less than or equal to 0.04. The p-value of the classifier may be less than or equal to 0.03. The p-value of the classifier may be less than or equal to 0.02. The p-value of the classifier may be less than or equal to 0.01. The p-value of the classifier may be less than or equal to 0.008. The p-value of the classifier may be less than or equal to 0.006. The p-value of the classifier may be less than or equal to 0.004. The p-value of the classifier may be less than or equal to 0.002. The p-value of the classifier may be less than or equal to 0.001.

[0039] The plurality of targets may comprise one or more targets selected from a linear model classifier. The plurality of targets may comprise two or more targets selected from a linear model classifier. The plurality of targets may comprise three or more targets selected from a linear model classifier. The plurality of targets may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 27 or more targets selected from a linear model classifier. The linear model classifier may be an LM2, and LM3, or an LM4 classifier. The linear model classifier may be an LM16 classifier (e.g., a linear model classifier with 16 targets). For example, a linear model classifier of the present disclosure may comprise two or more targets selected from Table 5.

[0040] The plurality of targets may comprise one or more targets selected from an SVM classifier. The plurality of targets may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10 or more targets selected from an SVM classifier. The plurality of targets may comprise 12, 13, 14, 15, 17, 20, 22, 25 or more targets selected from an SVM classifier. The SVM classifier may be an SVM2 classifier. An SVM classifier of the present disclosure may comprise two or more targets selected from Table 5.

[0041] The plurality of targets may comprise one or more targets selected from a KNN classifier. The plurality of targets may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10 or more targets selected from a KNN classifier. The plurality of targets may comprise 12, 13, 14, 15, 17, 20, 22, 25 or more targets selected from a KNN classifier. For example, a KNN classifier of the present disclosure may comprise two or more targets selected from Table 5.

[0042] The plurality of targets may comprise one or more targets selected from a Naïve Bayes (NB) classifier. The plurality of targets may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10 or more targets selected from an NB classifier. The plurality of targets may comprise 12, 13, 14, 15, 17, 20, 22, 25 or more targets selected from an NB classifier. For example, a NB classifier of the present disclosure may comprise two or more targets selected from Table 5.

[0043] The plurality of targets may comprise one or more targets selected from a recursive partitioning (Rpart) classifier. The plurality of targets may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10 or more targets selected from an Rpart classifier. The plurality of targets may comprise 12, 13, 14, 15, 17, 20, 22, 25 or more targets selected from an Rpart classifier. For example, an Rpart classifier of the present disclosure may comprise two or more targets selected from Table 5.

[0044] The plurality of targets may comprise one or more targets selected from a high dimensional discriminate analysis (HDDA) classifier. The plurality of targets may comprise two or more targets selected from a high dimensional discriminate analysis (HDDA) classifier. The plurality of targets may comprise three or more targets selected from a high dimensional discriminate analysis (HDDA) classifier. The plurality of targets may comprise 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 20, 22, 25 or more targets selected from a high dimensional discriminate analysis (HDDA) classifier. For example, an Rpart classifier of the present disclosure may comprise two or more targets selected from Table 5.

Probes/Primers

[0045] The present disclosure provides for a probe set for diagnosing, monitoring and/or predicting a status or outcome of breast cancer in a subject comprising a plurality of probes, wherein (i) the probes in the set are capable of detecting an expression level of at least one target; and (ii) the expression level determines the cancer status (e.g., risk of recurrence) of the subject with at least about 40% specificity.

[0046] The probe set may comprise one or more polynucleotide probes. Individual polynucleotide probes comprise a nucleotide sequence derived from the nucleotide sequence of the target sequences or complementary sequences thereof. The nucleotide sequence of the polynucleotide probe is designed such that it corresponds to, or is complementary to the target sequences. The polynucleotide probe can specifically hybridize under either stringent or lowered stringency hybridization conditions to a region of the target sequences, to the complement thereof, or to a nucleic acid sequence (such as a cDNA) derived therefrom.

[0047] The selection of the polynucleotide probe sequences and determination of their uniqueness may be carried out *in silico* using techniques known in the art, for example, based on a BLASTN search of the polynucleotide sequence in question against gene sequence databases, such as the Human Genome Sequence, UniGene, dbEST or the non-redundant database at NCBI. In one embodiment of the disclosure, the polynucleotide probe is complementary to a region of a target mRNA derived from a target sequence in the probe set. Computer programs can also be employed to select probe sequences that may not cross hybridize or may not hybridize non-specifically.

[0048] In some instances, microarray hybridization of RNA, extracted from breast cancer tissue samples and amplified, may yield a dataset that is then summarized and normalized by the fRMA technique. After removal (or filtration) of cross-hybridizing PSRs, and PSRs containing less than 4 probes, the remaining PSRs can be used in further analysis. Following fRMA and filtration, the data can be decomposed into its principal components and an analysis of variance model is used to determine the extent to which a batch effect remains present in the first 10 principal components.

[0049] These remaining PSRs can then be subjected to filtration by a T-test between CR (clinical recurrence) and non-CR samples. Using a p-value cut-off of 0.01, the remaining features (e.g., PSRs) can be further refined. Feature selection can be performed by regularized logistic regression using the

elastic-net penalty. The regularized regression may be bootstrapped over 1000 times using all training data; with each iteration of bootstrapping, features that have non-zero co-efficient following 3-fold cross validation can be tabulated. In some instances, features that were selected in at least 25% of the total runs were used for model building.

[0050] The polynucleotide probes of the present disclosure may range in length from about 15 nucleotides to the full length of the coding target or non-coding target. In one embodiment of the disclosure, the polynucleotide probes are at least about 15 nucleotides in length. In another embodiment, the polynucleotide probes are at least about 20 nucleotides in length. In a further embodiment, the polynucleotide probes are at least about 25 nucleotides in length. In another embodiment, the polynucleotide probes are between about 15 nucleotides and about 500 nucleotides in length. In other embodiments, the polynucleotide probes are between about 15 nucleotides and about 450 nucleotides, about 15 nucleotides and about 400 nucleotides, about 15 nucleotides and about 350 nucleotides, about 15 nucleotides and about 300 nucleotides, about 15 nucleotides and about 250 nucleotides, about 15 nucleotides and about 200 nucleotides in length. In some embodiments, the probes are at least 15 nucleotides in length. In some embodiments, the probes are at least 15 nucleotides in length. In some embodiments, the probes are at least 20 nucleotides, at least 25 nucleotides, at least 50 nucleotides, at least 75 nucleotides, at least 100 nucleotides, at least 125 nucleotides, at least 150 nucleotides, at least 200 nucleotides, at least 225 nucleotides, at least 250 nucleotides, at least 275 nucleotides, at least 300 nucleotides, at least 325 nucleotides, at least 350 nucleotides, at least 375 nucleotides in length.

[0051] The polynucleotide probes of a probe set can comprise RNA, DNA, RNA or DNA mimetics, or combinations thereof, and can be single-stranded or double-stranded. Thus the polynucleotide probes can be composed of naturally-occurring nucleobases, sugars and covalent internucleoside (backbone) linkages as well as polynucleotide probes having non-naturally-occurring portions which function similarly. Such modified or substituted polynucleotide probes may provide desirable properties such as, for example, enhanced affinity for a target gene and increased stability. The probe set may comprise a coding target and/or a non-coding target.

[0052] In some embodiments, the probe set comprise a plurality of target sequences that hybridize to at least about 5 coding targets and/or non-coding targets. Alternatively, the probe set comprise a plurality of target sequences that hybridize to at least about 10 coding targets and/or non-coding targets. In some embodiments, the probe set comprise a plurality of target sequences that hybridize to at least about 15 coding targets and/or non-coding targets. In some embodiments, the probe set comprise a plurality of target sequences that hybridize to at least about 20 coding targets and/or non-coding targets. In some embodiments, the probe set comprise a plurality of target sequences that hybridize to at least about 30 coding targets and/or non-coding targets.

[0053] The system of the present disclosure further provides for primers and primer pairs capable of amplifying target sequences defined by the probe set, or fragments or subsequences or complements

thereof. The nucleotide sequences of the probe set may be provided in computer-readable media for *in silico* applications and as a basis for the design of appropriate primers for amplification of one or more target sequences of the probe set.

[0054] Primers based on the nucleotide sequences of target sequences can be designed for use in amplification of the target sequences. For use in amplification reactions such as PCR, a pair of primers can be used. The exact composition of the primer sequences is not critical to the disclosure, but for most applications the primers may hybridize to specific sequences of the probe set under stringent conditions, particularly under conditions of high stringency, as known in the art. The pairs of primers are usually chosen so as to generate an amplification product of at least about 50 nucleotides, more usually at least about 100 nucleotides. Algorithms for the selection of primer sequences are generally known, and are available in commercial software packages. These primers may be used in standard quantitative or qualitative PCR-based assays to assess transcript expression levels of RNAs defined by the probe set. Alternatively, these primers may be used in combination with probes, such as molecular beacons in amplifications using real-time PCR.

[0055] In one embodiment, the primers or primer pairs, when used in an amplification reaction, specifically amplify at least a portion of a nucleic acid sequence of a target (or subgroups thereof as set forth herein), an RNA form thereof, or a complement to either thereof.

[0056] A label can optionally be attached to or incorporated into a probe or primer polynucleotide to allow detection and/or quantitation of a target polynucleotide representing the target sequence of interest. The target polynucleotide may be the expressed target sequence RNA itself, a cDNA copy thereof, or an amplification product derived therefrom, and may be the positive or negative strand, so long as it can be specifically detected in the assay being used. Similarly, an antibody may be labeled.

[0057] In certain multiplex formats, labels used for detecting different targets may be distinguishable. The label can be attached directly (e.g., via covalent linkage) or indirectly, e.g., via a bridging molecule or series of molecules (e.g., a molecule or complex that can bind to an assay component, or via members of a binding pair that can be incorporated into assay components, e.g. biotin-avidin or streptavidin). Many labels are commercially available in activated forms which can readily be used for such conjugation (for example through amine acylation), or labels may be attached through known or determinable conjugation schemes, many of which are known in the art.

[0058] Labels useful in the disclosure described herein include any substance which can be detected when bound to or incorporated into the biomolecule of interest. Any effective detection method can be used, including optical, spectroscopic, electrical, piezoelectrical, magnetic, Raman scattering, surface plasmon resonance, colorimetric, calorimetric, etc. A label is typically selected from a chromophore, a lumiphore, a fluorophore, one member of a quenching system, a chromogen, a hapten, an antigen, a magnetic particle, a material exhibiting nonlinear optics, a semiconductor nanocrystal, a metal nanoparticle, an enzyme, an antibody or binding portion or equivalent thereof, an aptamer, and one member of a binding pair, and combinations thereof. Quenching schemes may be

used, wherein a quencher and a fluorophore as members of a quenching pair may be used on a probe, such that a change in optical parameters occurs upon binding to the target introduce or quench the signal from the fluorophore. One example of such a system is a molecular beacon. Suitable quencher/fluorophore systems are known in the art. The label may be bound through a variety of intermediate linkages. For example, a polynucleotide may comprise a biotin-binding species, and an optically detectable label may be conjugated to biotin and then bound to the labeled polynucleotide. Similarly, a polynucleotide sensor may comprise an immunological species such as an antibody or fragment, and a secondary antibody containing an optically detectable label may be added.

[0059] Chromophores useful in the methods described herein include any substance which can absorb energy and emit light. For multiplexed assays, a plurality of different signaling chromophores can be used with detectably different emission spectra. The chromophore can be a lumophore or a fluorophore. Typical fluorophores include fluorescent dyes, semiconductor nanocrystals, lanthanide chelates, polynucleotide-specific dyes and green fluorescent protein.

[0060] In some embodiments, polynucleotides of the disclosure comprise at least 20 consecutive bases of the nucleic acid sequence of a target or a complement thereto. The polynucleotides may comprise at least 21, 22, 23, 24, 25, 27, 30, 32, 35, 40, 45, 50, or more consecutive bases of the nucleic acids sequence of a target.

[0061] The polynucleotides may be provided in a variety of formats, including as solids, in solution, or in an array. The polynucleotides may optionally comprise one or more labels, which may be chemically and/or enzymatically incorporated into the polynucleotide.

[0062] In some embodiments, one or more polynucleotides provided herein can be provided on a substrate. The substrate can comprise a wide range of material, either biological, nonbiological, organic, inorganic, or a combination of any of these. For example, the substrate may be a polymerized Langmuir Blodgett film, functionalized glass, Si, Ge, GaAs, GaP, SiO₂, SiN₄, modified silicon, or any one of a wide variety of gels or polymers such as (poly)tetrafluoroethylene, (poly)vinylidenedifluoride, polystyrene, cross-linked polystyrene, polyacrylic, polylactic acid, polyglycolic acid, poly(lactide coglycolide), polyanhydrides, poly(methyl methacrylate), poly(ethylene-co-vinyl acetate), polysiloxanes, polymeric silica, latexes, dextran polymers, epoxies, polycarbonates, or combinations thereof. Conducting polymers and photoconductive materials can be used.

[0063] The substrate can take the form of an array, a photodiode, an optoelectronic sensor such as an optoelectronic semiconductor chip or optoelectronic thin-film semiconductor, or a biochip. The location(s) of probe(s) on the substrate can be addressable; this can be done in highly dense formats, and the location(s) can be microaddressable or nanoaddressable.

Diagnostic Samples

[0064] A biological sample containing breast cancer cells is collected from a subject in need of treatment for cancer to evaluate if the subject is at low risk of cancer recurrence based on an expression level or expression profile and not likely to benefit from adjuvant radiotherapy.

Diagnostic samples for use with the systems and in the methods of the present disclosure comprise nucleic acids suitable for providing RNA expression information. In principle, the biological sample from which the expressed RNA is obtained and analyzed for target gene expression can be any material suspected of comprising cancerous breast tissue or cells. The diagnostic sample can be a biological sample used directly in a method of the disclosure. Alternatively, the diagnostic sample can be a sample prepared from a biological sample.

[0065] In one embodiment, the sample or portion of the sample comprising or suspected of comprising cancerous tissue or cells can be any source of biological material, including cells, tissue or fluid, including bodily fluids. Non-limiting examples of the source of the sample include an aspirate, a needle biopsy, a cytology pellet, a bulk tissue preparation or a section thereof obtained for example by surgery or autopsy, lymph fluid, blood, plasma, serum, tumors, and organs. In some embodiments, the sample is from a breast tumor biopsy.

[0066] The samples may be archival samples, having a known and documented medical outcome, or may be samples from current subjects whose ultimate medical outcome is not yet known.

[0067] In some embodiments, the sample may be dissected prior to molecular analysis. The sample may be prepared via macrodissection of a bulk tumor specimen or portion thereof, or may be treated via microdissection, for example via Laser Capture Microdissection (LCM).

[0068] The sample may initially be provided in a variety of states, as fresh tissue, fresh frozen tissue, fine needle aspirates, and may be fixed or unfixed. Frequently, medical laboratories routinely prepare medical samples in a fixed state, which facilitates tissue storage. A variety of fixatives can be used to fix tissue to stabilize the morphology of cells, and may be used alone or in combination with other agents. Exemplary fixatives include crosslinking agents, alcohols, acetone, Bouin's solution, Zenker solution, Hely solution, osmic acid solution and Carnoy solution.

[0069] Crosslinking fixatives can comprise any agent suitable for forming two or more covalent bonds, for example an aldehyde. Sources of aldehydes typically used for fixation include formaldehyde, paraformaldehyde, glutaraldehyde or formalin. Preferably, the crosslinking agent comprises formaldehyde, which may be included in its native form or in the form of paraformaldehyde or formalin. One of skill in the art would appreciate that for samples in which crosslinking fixatives have been used special preparatory steps may be necessary including for example heating steps and proteinase-k digestion; see methods.

[0070] One or more alcohols may be used to fix tissue, alone or in combination with other fixatives. Exemplary alcohols used for fixation include methanol, ethanol and isopropanol.

[0071] Formalin fixation is frequently used in medical laboratories. Formalin comprises both an alcohol, typically methanol, and formaldehyde, both of which can act to fix a biological sample.

[0072] Whether fixed or unfixed, the biological sample may optionally be embedded in an embedding medium. Exemplary embedding media used in histology including paraffin, Tissue-Tek® V.I.P, Paramat, Paramat Extra, Paraplast, Paraplast X-tra, Paraplast Plus, Peel Away Paraffin Embedding Wax, Polyester Wax, Carbowax Polyethylene Glycol, Polyfin, Tissue Freezing Medium TFMFM, Cryo-Gef, and OCT Compound (Electron Microscopy Sciences, Hatfield, PA). Prior to molecular analysis, the embedding material may be removed via any suitable techniques, as known in the art. For example, where the sample is embedded in wax, the embedding material may be removed by extraction with organic solvent(s), for example xylenes. Kits are commercially available for removing embedding media from tissues. Samples or sections thereof may be subjected to further processing steps as needed, for example serial hydration or dehydration steps.

[0073] In some embodiments, the sample is a fixed, wax-embedded biological sample. Frequently, samples from medical laboratories are provided as fixed, wax-embedded samples, most commonly as formalin-fixed, paraffin embedded (FFPE) tissues.

[0074] Whatever the source of the biological sample, the target polynucleotide that is ultimately assayed can be prepared synthetically (in the case of control sequences), but typically is purified from the biological source and subjected to one or more preparative steps. The RNA may be purified to remove or diminish one or more undesired components from the biological sample or to concentrate it. Conversely, where the RNA is too concentrated for the particular assay, it may be diluted.

RNA Extraction

[0075] RNA can be extracted and purified from biological samples using any suitable technique. A number of techniques are known in the art, and several are commercially available (e.g., FormaPure nucleic acid extraction kit, Agencourt Biosciences, Beverly MA, High Pure FFPE RNA Micro Kit, Roche Applied Science, Indianapolis, IN). RNA can be extracted from frozen tissue sections using TRIzol (Invitrogen, Carlsbad, CA) and purified using RNeasy Protect kit (Qiagen, Valencia, CA). RNA can be further purified using DNase I treatment (Ambion, Austin, TX) to eliminate any contaminating DNA. RNA concentrations can be made using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Rockland, DE). RNA can be further purified to eliminate contaminants that interfere with cDNA synthesis by cold sodium acetate precipitation. RNA integrity can be evaluated by running electropherograms, and RNA integrity number (RIN, a correlative measure that indicates intactness of mRNA) can be determined using the RNA 6000 PicoAssay for the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA).

Kits

[0076] Kits for performing the desired method(s) are also provided, and comprise a container or housing for holding the components of the kit, one or more vessels containing one or more nucleic acid(s), and optionally one or more vessels containing one or more reagents. The reagents include

those described in the composition of matter section above, and those reagents useful for performing the methods described, including amplification reagents, and may include one or more probes, primers or primer pairs, enzymes (including polymerases and ligases), intercalating dyes, labeled probes, and labels that can be incorporated into amplification products.

[0077] In some embodiments, the kit comprises primers or primer pairs specific for those subsets and combinations of target sequences described herein. The primers or pairs of primers are suitable for selectively amplifying the target sequences. The kit may comprise at least two, three, four or five primers or pairs of primers suitable for selectively amplifying one or more targets. The kit may comprise at least 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, or more primers or pairs of primers suitable for selectively amplifying one or more targets.

[0078] In some embodiments, the primers or primer pairs of the kit, when used in an amplification reaction, specifically amplify a non-coding target, coding target, exonic, or non-exonic target described herein, a nucleic acid sequence corresponding to a target selected from Table 5, an RNA form thereof, or a complement to either thereof. The kit may include a plurality of such primers or primer pairs which can specifically amplify a corresponding plurality of different amplify a non-coding target, coding target, exonic, or non-exonic transcript described herein, a nucleic acid sequence corresponding to a target selected from Table 5, RNA forms thereof, or complements thereto. At least two, three, four or five primers or pairs of primers suitable for selectively amplifying the one or more targets can be provided in kit form. In some embodiments, the kit comprises from five to fifty primers or pairs of primers suitable for amplifying the one or more targets.

[0079] The reagents may independently be in liquid or solid form. The reagents may be provided in mixtures. Control samples and/or nucleic acids may optionally be provided in the kit. Control samples may include tissue and/or nucleic acids obtained from or representative of tumor samples from subjects showing no evidence of disease, as well as tissue and/or nucleic acids obtained from or representative of tumor samples from subjects that develop systemic cancer.

[0080] The nucleic acids may be provided in an array format, and thus an array or microarray may be included in the kit. The kit optionally may be certified by a government agency for use in prognosing the disease outcome of cancer subjects and/or for designating a treatment modality.

[0081] Instructions for using the kit to perform one or more methods of the disclosure can be provided with the container, and can be provided in any fixed medium. The instructions may be located inside or outside the container or housing, and/or may be printed on the interior or exterior of any surface thereof. A kit may be in multiplex form for concurrently detecting and/or quantitating one or more different target polynucleotides representing the expressed target genes.

Amplification and Hybridization

[0082] Following sample collection and nucleic acid extraction, the nucleic acid portion of the sample comprising RNA that is or can be used to prepare the target polynucleotide(s) of interest can

be subjected to one or more preparative reactions. These preparative reactions can include *in vitro* transcription (IVT), labeling, fragmentation, amplification and other reactions. mRNA can first be treated with reverse transcriptase and a primer to create cDNA prior to detection, quantitation and/or amplification; this can be done *in vitro* with purified mRNA or *in situ*, e.g., in cells or tissues affixed to a slide.

[0083] By "amplification" is meant any process of producing at least one copy of a nucleic acid, in this case an expressed RNA, and in many cases produces multiple copies. An amplification product can be RNA or DNA, and may include a complementary strand to the expressed target sequence. DNA amplification products can be produced initially through reverse translation and then optionally from further amplification reactions. The amplification product may include all or a portion of a target sequence, and may optionally be labeled. A variety of amplification methods are suitable for use, including polymerase-based methods and ligation-based methods. Exemplary amplification techniques include the polymerase chain reaction method (PCR), the ligase chain reaction (LCR), ribozyme-based methods, self-sustained sequence replication (3SR), nucleic acid sequence-based amplification (NASBA), the use of Q Beta replicase, reverse transcription, nick translation, and the like.

[0084] Asymmetric amplification reactions may be used to preferentially amplify one strand representing the target sequence that is used for detection as the target polynucleotide. In some cases, the presence and/or amount of the amplification product itself may be used to determine the expression level of a given target sequence. In other instances, the amplification product may be used to hybridize to an array or other substrate comprising sensor polynucleotides which are used to detect and/or quantitate target sequence expression.

[0085] The first cycle of amplification in polymerase-based methods typically forms a primer extension product complementary to the template strand. If the template is single-stranded RNA, a polymerase with reverse transcriptase activity is used in the first amplification to reverse transcribe the RNA to DNA, and additional amplification cycles can be performed to copy the primer extension products. The primers for a PCR must, of course, be designed to hybridize to regions in their corresponding template that can produce an amplifiable segment; thus, each primer must hybridize so that its 3' nucleotide is paired to a nucleotide in its complementary template strand that is located 3' from the 3' nucleotide of the primer used to replicate that complementary template strand in the PCR.

[0086] The target polynucleotide can be amplified by contacting one or more strands of the target polynucleotide with a primer and a polymerase having suitable activity to extend the primer and copy the target polynucleotide to produce a full-length complementary polynucleotide or a smaller portion thereof. Any enzyme having a polymerase activity that can copy the target polynucleotide can be used, including DNA polymerases, RNA polymerases, reverse transcriptases, enzymes having more than one type of polymerase or enzyme activity. The enzyme can be thermolabile or thermostable. Mixtures of enzymes can also be used. Exemplary enzymes include: DNA polymerases such as DNA

Polymerase I ("Pol I"), the Klenow fragment of Pol I, T4, T7, Sequenase® T7, Sequenase® Version 2.0 T7, *Tub*, *Taq*, *Tth*, *Pfic*, *Pfu*, *Tsp*, *Tfl*, *Tli* and *Pyrococcus sp* GB-D DNA polymerases; RNA polymerases such as *E. coil*, SP6, T3 and T7 RNA polymerases; and reverse transcriptases such as AMV, M-MuLV, MMLV, RNase H MMLV (SuperScript®), SuperScript® II, ThermoScript®, HIV-1, and RAV2 reverse transcriptases. All of these enzymes are commercially available. Exemplary polymerases with multiple specificities include RAV2 and *Tli* (exo-) polymerases. Exemplary thermostable polymerases include *Tub*, *Taq*, *Tth*, *Pfic*, *Pfu*, *Tsp*, *Tfl*, *Tli* and *Pyrococcus sp*. GB-D DNA polymerases.

[0087] Suitable reaction conditions are chosen to permit amplification of the target polynucleotide, including pH, buffer, ionic strength, presence and concentration of one or more salts, presence and concentration of reactants and cofactors such as nucleotides and magnesium and/or other metal ions (e.g., manganese), optional cosolvents, temperature, thermal cycling profile for amplification schemes comprising a polymerase chain reaction, and may depend in part on the polymerase being used as well as the nature of the sample. Cosolvents include formamide (typically at from about 2 to about 10 %), glycerol (typically at from about 5 to about 10 %), and DMSO (typically at from about 0.9 to about 10 %). Techniques may be used in the amplification scheme in order to minimize the production of false positives or artifacts produced during amplification. These include "touchdown" PCR, hot-start techniques, use of nested primers, or designing PCR primers so that they form stem-loop structures in the event of primer-dimer formation and thus are not amplified. Techniques to accelerate PCR can be used, for example centrifugal PCR, which allows for greater convection within the sample, and comprising infrared heating steps for rapid heating and cooling of the sample. One or more cycles of amplification can be performed. An excess of one primer can be used to produce an excess of one primer extension product during PCR; preferably, the primer extension product produced in excess is the amplification product to be detected. A plurality of different primers may be used to amplify different target polynucleotides or different regions of a particular target polynucleotide within the sample.

[0088] An amplification reaction can be performed under conditions which allow an optionally labeled sensor polynucleotide to hybridize to the amplification product during at least part of an amplification cycle. When the assay is performed in this manner, real-time detection of this hybridization event can take place by monitoring for light emission or fluorescence during amplification, as known in the art.

[0089] Where the amplification product is to be used for hybridization to an array or microarray, a number of suitable commercially available amplification products are available. These include amplification kits available from NuGEN, Inc. (San Carlos, CA), including the WT-Ovation™ System, WT-Ovation™ System v2, WT-Ovation™ Pico System, WT-Ovation™ FFPE Exon Module, WT-Ovation™ FFPE Exon Module RiboAmp and RiboAmp^{Plus} RNA Amplification Kits (MDS Analytical Technologies (formerly Arcturus) (Mountain View, CA), Genisphere, Inc. (Hatfield,

PA), including the RampUp Plus and SenceAmp RNA Amplification kits, alone or in combination. Amplified nucleic acids may be subjected to one or more purification reactions after amplification and labeling, for example using magnetic beads (e.g., RNAClean magnetic beads, Agencourt Biosciences).

[0090] Multiple RNA biomarkers can be analyzed using real-time quantitative multiplex RT-PCR platforms and other multiplexing technologies such as GenomeLab GeXP Genetic Analysis System (Beckman Coulter, Foster City, CA), SmartCycler® 9600 or GeneXpert® Systems (Cepheid, Sunnyvale, CA), ABI 7900 HT Fast Real Time PCR system (Applied Biosystems, Foster City, CA), LightCycler® 480 System (Roche Molecular Systems, Pleasanton, CA), xMAP 100 System (Luminex, Austin, TX) Solexa Genome Analysis System (Illumina, Hayward, CA), OpenArray Real Time qPCR (BioTrove, Woburn, MA) and BeadXpress System (Illumina, Hayward, CA).

Detection and/or Quantification of Target Genes

[0091] Any method of detecting and/or quantitating the expression of the encoded target genes can in principle be used in the disclosure. The expressed target genes can be directly detected and/or quantitated, or may be copied and/or amplified to allow detection of amplified copies of the expressed target genes.

[0092] Methods for detecting and/or quantifying a target gene can include Northern blotting, sequencing, array or microarray hybridization, by enzymatic cleavage of specific structures (e.g., a Clariom S assay, ThermoFisher Scientific, an Invader® assay, Third Wave Technologies, e.g. as described in U.S. Pat. Nos. 5,846,717, 6,090,543; 6,001,567; 5,985,557; and 5,994,069) and amplification methods, e.g. RT-PCR, including in a TaqMan® assay (PE Biosystems, Foster City, Calif., e.g. as described in U.S. Pat. Nos. 5,962,233 and 5,538,848), and may be quantitative or semi-quantitative, and may vary depending on the origin, amount and condition of the available biological sample. Combinations of these methods may also be used. For example, nucleic acids may be amplified, labeled and subjected to microarray analysis. Methods for detecting and/or quantifying a target gene can include gene-level expression analysis of annotated genes using microarray hybridization (e.g., GeneChip Human Exon 1.0 ST assay or Clariom S assay, ThermoFisher Scientific).

[0093] In some instances, target genes may be detected by sequencing. Sequencing methods may comprise whole genome sequencing or exome sequencing. Sequencing methods such as Maxim-Gilbert, chain-termination, or high-throughput systems may also be used. Additional, suitable sequencing techniques include classic dideoxy sequencing reactions (Sanger method) using labeled terminators or primers and gel separation in slab or capillary, sequencing by synthesis using reversibly terminated labeled nucleotides, pyrosequencing, 454 sequencing, allele specific hybridization to a library of labeled oligonucleotide probes, sequencing by synthesis using allele specific hybridization

to a library of labeled clones that is followed by ligation, real time monitoring of the incorporation of labeled nucleotides during a polymerization step, and SOLiD sequencing.

[0094] Additional methods for detecting and/or quantifying a target gene include single-molecule sequencing (e.g., Helicos, PacBio), sequencing by synthesis (e.g., Illumina, Ion Torrent), sequencing by ligation (e.g., ABI SOLiD), sequencing by hybridization (e.g., Complete Genomics), in situ hybridization, bead-array technologies (e.g., Luminex xMAP, Illumina BeadChips), branched DNA technology (e.g., Panomics, Genisphere). Sequencing methods may use fluorescent (e.g., Illumina) or electronic (e.g., Ion Torrent, Oxford Nanopore) methods of detecting nucleotides.

Reverse Transcription for QRT-PCR Analysis

[0095] Reverse transcription can be performed by any method known in the art. For example, reverse transcription may be performed using the Omniscript kit (Qiagen, Valencia, CA), Superscript III kit (Invitrogen, Carlsbad, CA), for RT-PCR. Target-specific priming can be performed in order to increase the sensitivity of detection of target genes and generate target-specific cDNA.

TaqMan® Gene Expression Analysis

[0096] TaqMan® RT-PCR can be performed using Applied Biosystems Prism (ABI) 7900 HT instruments in a 5 1.11 volume with target gene-specific cDNA equivalent to 1 ng total RNA.

[0097] Primers and probes concentrations for TaqMan analysis are added to amplify fluorescent amplicons using PCR cycling conditions such as 95°C for 10 minutes for one cycle, 95°C for 20 seconds, and 60°C for 45 seconds for 40 cycles. A reference sample can be assayed to ensure reagent and process stability. Negative controls (e.g., no template) should be assayed to monitor any exogenous nucleic acid contamination.

Classification Arrays

[0098] The present disclosure contemplates that a probe set or probes derived therefrom may be provided in an array format. In the context of the present disclosure, an "array" is a spatially or logically organized collection of polynucleotide probes. An array comprising probes specific for a coding target, non-coding target, or a combination thereof may be used. Alternatively, an array comprising probes specific for two or more of transcripts of a target, or a product derived thereof, can be used. Desirably, an array may be specific for 5, 10, 15, 20, 25, 30 or more of transcripts of a target gene. Expression of these genes may be detected alone or in combination with other transcripts. In some embodiments, an array is used which comprises a wide range of sensor probes for breast-specific expression products, along with appropriate control sequences. In some instances, the array may comprise the Human Exon 1.0 ST Array (HuEx 1.0 ST, Thermo Fisher Scientific, Santa Clara, CA.).

[0099] Typically the polynucleotide probes are attached to a solid substrate and are ordered so that the location (on the substrate) and the identity of each are known. The polynucleotide probes can be attached to one of a variety of solid substrates capable of withstanding the reagents and conditions necessary for use of the array. Examples include, but are not limited to, polymers, such as (poly)tetrafluoroethylene, (poly)vinylidene difluoride, polystyrene, polycarbonate, polypropylene and polystyrene; ceramic; silicon; silicon dioxide; modified silicon; (fused) silica, quartz or glass; functionalized glass; paper, such as filter paper; diazotized cellulose; nitrocellulose filter; nylon membrane; and polyacrylamide gel pad. Substrates that are transparent to light are useful for arrays that may be used in an assay that involves optical detection.

[00100] Examples of array formats include membrane or filter arrays (for example, nitrocellulose, nylon arrays), plate arrays (for example, multiwell, such as a 24-, 96-, 256-, 384-, 864- or 1536-well, microtitre plate arrays), pin arrays, and bead arrays (for example, in a liquid "slurry"). Arrays on substrates such as glass or ceramic slides are often referred to as chip arrays or "chips." Such arrays are well known in the art. In one embodiment of the present disclosure, the Cancer Prognostic array is a chip.

Data Analysis

[00101] In some embodiments, one or more pattern recognition methods can be used in analyzing the expression level of target genes. The pattern recognition method can comprise a linear combination of expression levels, or a nonlinear combination of expression levels. In some embodiments, expression measurements for RNA transcripts or combinations of RNA transcript levels are formulated into linear or non-linear models or algorithms (e.g., an 'expression signature') and converted into a likelihood score. This likelihood score indicates the probability that a biological sample is from a subject who may exhibit no evidence of disease, who may exhibit systemic cancer, or who may exhibit biochemical recurrence or locoregional recurrence. The likelihood score can be used to distinguish these disease states. The models and/or algorithms can be provided in machine readable format, and may be used to correlate expression levels or an expression profile with a disease state, and/or to designate a treatment modality for a subject or class of subjects.

[00102] Assaying the expression level for a plurality of target genes may comprise the use of an algorithm or classifier. Array data can be managed, classified, and analyzed using techniques known in the art. Assaying the expression level for a plurality of gene targets may comprise probe set modeling and data pre-processing. Probe set modeling and data pre-processing can be derived using the Robust Multi-Array (RMA) algorithm or variants GC-RMA, *f*RMA, Probe Logarithmic Intensity Error (PLIER) algorithm or variant iterPLIER. Variance or intensity filters can be applied to pre-process data using the RMA algorithm, for example by removing target genes with a standard deviation of < 10 or a mean intensity of < 100 intensity units of a normalized data range, respectively.

[00103] Alternatively, assaying the expression level for a plurality of gene targets may comprise the use of a machine learning algorithm. The machine learning algorithm may comprise a supervised learning algorithm. Examples of supervised learning algorithms may include Average One-Dependence Estimators (AODE), Artificial neural network (e.g., Backpropagation), Bayesian statistics (e.g., Naive Bayes classifier, Bayesian network, Bayesian knowledge base), Case-based reasoning, Decision trees, Inductive logic programming, Gaussian process regression, Group method of data handling (GMDH), Learning Automata, Learning Vector Quantization, Minimum message length (decision trees, decision graphs, etc.), Lazy learning, Instance-based learning Nearest Neighbor Algorithm, Analogical modeling, Probably approximately correct learning (PAC) learning, Ripple down rules, a knowledge acquisition methodology, Symbolic machine learning algorithms, Subsymbolic machine learning algorithms, Support vector machines, Random Forests, Ensembles of classifiers, Bootstrap aggregating (bagging), and Boosting. Supervised learning may comprise ordinal classification such as regression analysis and Information fuzzy networks (IFN). Alternatively, supervised learning methods may comprise statistical classification, such as AODE, Linear classifiers (e.g., Fisher's linear discriminant, Logistic regression, Naive Bayes classifier, Perceptron, and Support vector machine, Lasso and Elastic-Net Regularized General Linear models), quadratic classifiers, k-nearest neighbor, Boosting, Decision trees (e.g., C4.5, Random forests), Bayesian networks, and Hidden Markov models.

[00104] The machine learning algorithms may also comprise an unsupervised learning algorithm. Examples of unsupervised learning algorithms may include artificial neural network, Data clustering, Expectation-maximization algorithm, Self-organizing map, Radial basis function network, Vector Quantization, Generative topographic map, Information bottleneck method, and IBSEAD. Unsupervised learning may also comprise association rule learning algorithms such as Apriori algorithm, Eclat algorithm and FP-growth algorithm. Hierarchical clustering, such as Single-linkage clustering and Conceptual clustering, may also be used. Alternatively, unsupervised learning may comprise partitional clustering such as K-means algorithm and Fuzzy clustering.

[00105] In some instances, the machine learning algorithms comprise a reinforcement learning algorithm. Examples of reinforcement learning algorithms include, but are not limited to, temporal difference learning, Q-learning and Learning Automata. Alternatively, the machine learning algorithm may comprise Data Pre-processing.

[00106] Preferably, the machine learning algorithms may include, but are not limited to, Average One-Dependence Estimators (AODE), Fisher's linear discriminant, Logistic regression, Perceptron, Multilayer Perceptron, Artificial Neural Networks, Support vector machines, Quadratic classifiers, Boosting, Decision trees, C4.5, Bayesian networks, Hidden Markov models, High-Dimensional Discriminant Analysis, and Gaussian Mixture Models. The machine learning algorithm may comprise support vector machines, Naïve Bayes classifier, k-nearest neighbor, high-dimensional discriminant

analysis, or Gaussian mixture models. In some instances, the machine learning algorithm comprises Random Forests.

Subtyping

[00107] Molecular subtyping is a method of classifying breast cancer into one of multiple genetically-distinct categories, or subtypes. Each subtype responds differently to different kinds of treatments, and the presence of a particular subtype is predictive of, for example, radioresistance or chemoresistance, higher or lower risk of recurrence, or good or poor prognosis for an individual. Differential expression analysis of a plurality of the gene targets listed in Table 5 allows for the identification of subjects at low risk of cancer recurrence who, for example, may benefit least from adjuvant radiotherapy. In some instances, the molecular subtyping methods of the present disclosure are used in combination with other biomarkers, like tumor grade and hormone levels, for analyzing the breast cancer. For example, a subject with estrogen receptor positive (ER+), human epidermal growth factor receptor 2 negative (HER2-) breast cancer, node-negative breast cancer, who is post-menopausal, may be more likely to have a lower risk of recurrence (e.g., locoregional recurrence).

Cancer Recurrence

[00108] Cancer recurrence is the return of cancer after a period when no cancer cells are detected in the body. Following surgery for operable breast cancer, disease can recur locally, regionally, and/or at distant metastatic sites. A local recurrence is the reappearance of cancer on the ipsilateral chest wall. In contrast, a regional recurrence denotes tumor involving the regional lymph nodes, usually ipsilateral axillary or supraclavicular, and less commonly infraclavicular and/or internal mammary nodes. Some breast cancer patients will have local or locoregional recurrence after breast-conserving surgery and radiotherapy within ten years of first being diagnosed with breast cancer. If the breast was removed in the course of initial treatment, these women will have a local recurrence in the armpit or the chest wall within ten years. In some embodiments, the subject treated in the methods of the present disclosure has a node-negative breast cancer.

Therapeutic regimens

[00109] Diagnosing, predicting, or monitoring a status or outcome of a cancer may comprise treating a cancer or preventing a cancer progression. In addition, diagnosing, predicting, or monitoring a status or outcome of a cancer may comprise identifying or predicting that a subject is at low or high risk of a recurrence (e.g., locoregional recurrence). In some instances, diagnosing, predicting, or monitoring may comprise determining a therapeutic regimen. Determining a therapeutic regimen may comprise administering an anti-cancer therapy. Alternatively, determining a therapeutic regimen may comprise modifying, recommending, continuing or discontinuing an anti-cancer regimen. For example, a subject determined to be at low risk of recurrence of breast cancer based on expression

profiling, as described herein, may be spared adjuvant radiotherapy. In some instances, a subject determined to be at high risk of recurrence of breast cancer based on expression profiling, as described herein, may be treated with mastectomy, radiation boost, or adjuvant systemic therapy. In some instances, if the sample expression patterns are consistent with the expression pattern for a known disease or disease outcome, the expression patterns can be used to designate one or more treatment modalities (e.g., therapeutic regimens, anti-cancer regimen). An anti-cancer regimen may comprise one or more anti-cancer therapies. Examples of anti-cancer therapies include hormonal/endocrine therapy, surgery, chemotherapy, radiation therapy, immunotherapy/biological therapy, and photodynamic therapy.

[00110] Hormonal therapy or endocrine therapy may involve administration of hormones, such as steroid hormones or hormone antagonists to modulate the levels of certain hormones in order to arrest growth or induce apoptosis of hormone-responsive cancer cells. For example, breast cancer may be treated with a selective estrogen receptor modulator (SERM) such as, but not limited to, tamoxifen, raloxifene, and toremifene. Alternately or additionally, inhibitors of hormone synthesis such as aromatase inhibitors, including, but not limited to, letrozole, anastrozole, exemestane, and aminoglutethimide may be used to treat breast cancer. In some cases, hormone supplementation with progestins such as, but not limited to, megestrol acetate and medroxyprogesterone acetate may be used for the treatment of hormone-responsive, advanced breast cancer. In particular, ER+ breast cancer can be treated with either an estrogen receptor antagonist (e.g. tamoxifen) or a drug that blocks the production of estrogen such as an aromatase inhibitor (e.g. anastrozole or letrozole). Hormonal therapy may also include surgical removal of endocrine organs (e.g., orchiectomy or oophorectomy).

[00111] Surgical oncology uses surgical methods to diagnose, stage, and treat cancer, and to relieve certain cancer-related symptoms. Surgery may be used to remove the tumor (e.g., excisions, resections, debulking surgery), reconstruct a part of the body (e.g., restorative surgery), and/or to relieve symptoms such as pain (e.g., palliative surgery). Surgery may also include cryosurgery. Cryosurgery (also called cryotherapy) may use extreme cold produced by liquid nitrogen (or argon gas) to destroy abnormal tissue. Cryosurgery can be used to treat external tumors, such as those on the skin. For external tumors, liquid nitrogen can be applied directly to the cancer cells with a cotton swab or spraying device. Cryosurgery may also be used to treat tumors inside the body (internal tumors and tumors in the bone). For internal tumors, liquid nitrogen or argon gas may be circulated through a hollow instrument called a cryoprobe, which is placed in contact with the tumor. An ultrasound or MRI may be used to guide the cryoprobe and monitor the freezing of the cells, thus limiting damage to nearby healthy tissue. A ball of ice crystals may form around the probe, freezing nearby cells. Sometimes more than one probe is used to deliver the liquid nitrogen to various parts of the tumor. The probes may be put into the tumor during surgery or through the skin (percutaneously). After cryosurgery, the frozen tissue thaws and may be naturally absorbed by the body (for internal tumors), or may dissolve and form a scab (for external tumors).

[00112] Chemotherapeutic agents may also be used for the treatment of cancer. Examples of chemotherapeutic agents include alkylating agents, anti-metabolites, plant alkaloids and terpenoids, vinca alkaloids, podophyllotoxin, taxanes, topoisomerase inhibitors, and cytotoxic antibiotics. Cisplatin, carboplatin, and oxaliplatin are examples of alkylating agents. Other alkylating agents include mechlorethamine, cyclophosphamide, chlorambucil, ifosfamide. Alkylating agents may impair cell function by forming covalent bonds with the amino, carboxyl, sulfhydryl, and phosphate groups in biologically important molecules. Alternatively, alkylating agents may chemically modify a cell's DNA.

[00113] Anti-metabolites are another example of chemotherapeutic agents. Anti-metabolites may masquerade as purines or pyrimidines and may prevent purines and pyrimidines from becoming incorporated into DNA during the "S" phase (of the cell cycle), thereby stopping normal development and division. Antimetabolites may also affect RNA synthesis. Examples of metabolites include azathioprine and mercaptopurine.

[00114] Alkaloids may be derived from plants and block cell division may also be used for the treatment of cancer. Alkaloids may prevent microtubule function. Examples of alkaloids are vinca alkaloids and taxanes. Vinca alkaloids may bind to specific sites on tubulin and inhibit the assembly of tubulin into microtubules (M phase of the cell cycle). The vinca alkaloids may be derived from the Madagascar periwinkle, *Catharanthus roseus* (formerly known as *Vinca rosea*). Examples of vinca alkaloids include, but are not limited to, vincristine, vinblastine, vinorelbine, or vindesine. Taxanes are diterpenes produced by the plants of the genus *Taxus* (yews). Taxanes may be derived from natural sources or synthesized artificially. Taxanes include paclitaxel (Taxol) and docetaxel (Taxotere). Taxanes may disrupt microtubule function. Microtubules are essential to cell division, and taxanes may stabilize GDP-bound tubulin in the microtubule, thereby inhibiting the process of cell division. Thus, in essence, taxanes may be mitotic inhibitors. Taxanes may also be radiosensitizing and often contain numerous chiral centers.

[00115] Alternative chemotherapeutic agents include podophyllotoxin. Podophyllotoxin is a plant-derived compound that may help with digestion and may be used to produce cytostatic drugs such as etoposide and teniposide. They may prevent the cell from entering the G1 phase (the start of DNA replication) and the replication of DNA (the S phase).

[00116] Topoisomerases are essential enzymes that maintain the topology of DNA. Inhibition of type I or type II topoisomerases may interfere with both transcription and replication of DNA by upsetting proper DNA supercoiling. Some chemotherapeutic agents may inhibit topoisomerases. For example, some type I topoisomerase inhibitors include *camptothecins*: irinotecan and topotecan. Examples of type II inhibitors include amsacrine, etoposide, etoposide phosphate, and teniposide. Kinase inhibitors may be used to treat breast cancer.

[00117] Another example of chemotherapeutic agents is cytotoxic antibiotics. Cytotoxic antibiotics are a group of antibiotics that are used for the treatment of cancer because they may interfere with

DNA replication and/or protein synthesis. Cytotoxic antibiotics include, but are not limited to, actinomycin, anthracyclines, doxorubicin, daunorubicin, valrubicin, idarubicin, epirubicin, bleomycin, plicamycin, and mitomycin.

[00118] In some instances, the anti-cancer treatment may comprise radiation therapy. Radiation can come from a machine outside the body (external-beam radiation therapy) or from radioactive material placed in the body near cancer cells (internal radiation therapy, more commonly called brachytherapy). Systemic radiation therapy uses a radioactive substance, given by mouth or into a vein that travels in the blood to tissues throughout the body.

[00119] External-beam radiation therapy may be delivered in the form of photon beams (either x-rays or gamma rays). A photon is the basic unit of light and other forms of electromagnetic radiation. An example of external-beam radiation therapy is called 3-dimensional conformal radiation therapy (3D-CRT). 3D-CRT may use computer software and advanced treatment machines to deliver radiation to very precisely shaped target areas. Many other methods of external-beam radiation therapy are currently being tested and used in cancer treatment. These methods include, but are not limited to, intensity-modulated radiation therapy (IMRT), image-guided radiation therapy (IGRT), Stereotactic radiosurgery (SRS), Stereotactic body radiation therapy (SBRT), and proton therapy.

[00120] Intensity-modulated radiation therapy (IMRT) is an example of external-beam radiation and may use hundreds of tiny radiation beam-shaping devices, called collimators, to deliver a single dose of radiation. The collimators can be stationary or can move during treatment, allowing the intensity of the radiation beams to change during treatment sessions. This kind of dose modulation allows different areas of a tumor or nearby tissues to receive different doses of radiation. IMRT is planned in reverse (called inverse treatment planning). In inverse treatment planning, the radiation doses to different areas of the tumor and surrounding tissue are planned in advance, and then a high-powered computer program calculates the required number of beams and angles of the radiation treatment. In contrast, during traditional (forward) treatment planning, the number and angles of the radiation beams are chosen in advance and computers calculate how much dose may be delivered from each of the planned beams. The goal of IMRT is to increase the radiation dose to the areas that need it and reduce radiation exposure to specific sensitive areas of surrounding normal tissue.

[00121] Another example of external-beam radiation is image-guided radiation therapy (IGRT). In IGRT, repeated imaging scans (CT, MRI, or PET) may be performed during treatment. These imaging scans may be processed by computers to identify changes in a tumor's size and location due to treatment and to allow the position of the subject or the planned radiation dose to be adjusted during treatment as needed. Repeated imaging can increase the accuracy of radiation treatment and may allow reductions in the planned volume of tissue to be treated, thereby decreasing the total radiation dose to normal tissue.

[00122] Tomotherapy is a type of image-guided IMRT. A tomotherapy machine is a hybrid between a CT imaging scanner and an external-beam radiation therapy machine. The part of the tomotherapy

machine that delivers radiation for both imaging and treatment can rotate completely around the subject in the same manner as a normal CT scanner. Tomotherapy machines can capture CT images of the subject's tumor immediately before treatment sessions, to allow for very precise tumor targeting and sparing of normal tissue.

[00123] Stereotactic radiosurgery (SRS) can deliver one or more high doses of radiation to a small tumor. SRS uses extremely accurate image-guided tumor targeting and subject positioning. Therefore, a high dose of radiation can be given without excess damage to normal tissue. SRS can be used to treat small tumors with well-defined edges. It is most commonly used in the treatment of brain or spinal tumors and brain metastases from other cancer types. For the treatment of some brain metastases, subjects may receive radiation therapy to the entire brain (called whole-brain radiation therapy) in addition to SRS. SRS requires the use of a head frame or other device to immobilize the subject during treatment to ensure that the high dose of radiation is delivered accurately.

[00124] Stereotactic body radiation therapy (SBRT) delivers radiation therapy in fewer sessions, using smaller radiation fields and higher doses than 3D-CRT in most cases. SBRT may treat tumors that lie outside the brain and spinal cord. Because these tumors are more likely to move with the normal motion of the body, and therefore cannot be targeted as accurately as tumors within the brain or spine, SBRT is usually given in more than one dose. SBRT can be used to treat small, isolated tumors, including cancers in the lung and liver. SBRT systems may be known by their brand names, such as the CyberKnife®.

[00125] In proton therapy, external-beam radiation therapy may be delivered by proton. Protons are a type of charged particle. Proton beams differ from photon beams mainly in the way they deposit energy in living tissue. Whereas photons deposit energy in small packets all along their path through tissue, protons deposit much of their energy at the end of their path (called the Bragg peak) and deposit less energy along the way. Use of protons may reduce the exposure of normal tissue to radiation, possibly allowing the delivery of higher doses of radiation to a tumor.

[00126] Other charged particle beams such as electron beams may be used to irradiate superficial tumors, such as skin cancer or tumors near the surface of the body, but they cannot travel very far through tissue.

[00127] Internal radiation therapy (brachytherapy) is radiation delivered from radiation sources (radioactive materials) placed inside or on the body. Several brachytherapy techniques are used in cancer treatment. Interstitial brachytherapy may use a radiation source placed within tumor tissue, such as within a breast tumor. Intracavitary brachytherapy may use a source placed within a surgical cavity or a body cavity, such as the chest cavity, near a tumor. Episcleral brachytherapy, which may be used to treat melanoma inside the eye, may use a source that is attached to the eye. In brachytherapy, radioactive isotopes can be sealed in tiny pellets or "seeds." These seeds may be placed in subjects using delivery devices, such as needles, catheters, or some other type of carrier. As the isotopes decay naturally, they give off radiation that may damage nearby cancer cells.

Brachytherapy may be able to deliver higher doses of radiation to some cancers than external-beam radiation therapy while causing less damage to normal tissue.

[00128] Brachytherapy can be given as a low-dose-rate or a high-dose-rate treatment. In low-dose-rate treatment, cancer cells receive continuous low-dose radiation from the source over a period of several days. In high-dose-rate treatment, a robotic machine attached to delivery tubes placed inside the body may guide one or more radioactive sources into or near a tumor, and then removes the sources at the end of each treatment session. High-dose-rate treatment can be given in one or more treatment sessions. An example of a high-dose-rate treatment is the MammoSite® system.

Brachytherapy may be used to treat subjects with breast cancer who have undergone breast-conserving surgery.

[00129] The placement of brachytherapy sources can be temporary or permanent. For permanent brachytherapy, the sources may be surgically sealed within the body and left there, even after all of the radiation has been given off. In some instances, the remaining material (in which the radioactive isotopes were sealed) does not cause any discomfort or harm to the subject. Permanent brachytherapy is a type of low-dose-rate brachytherapy. For temporary brachytherapy, tubes (catheters) or other carriers are used to deliver the radiation sources, and both the carriers and the radiation sources are removed after treatment. Temporary brachytherapy can be either low-dose-rate or high-dose-rate treatment. Brachytherapy may be used alone or in addition to external-beam radiation therapy to provide a “boost” of radiation to a tumor while sparing surrounding normal tissue.

[00130] In systemic radiation therapy, a subject may swallow or receive an injection of a radioactive substance, such as radioactive iodine or a radioactive substance bound to a monoclonal antibody. Radioactive iodine (¹³¹I) is a type of systemic radiation therapy commonly used to help treat cancer, such as thyroid cancer. Thyroid cells naturally take up radioactive iodine. For systemic radiation therapy for some other types of cancer, a monoclonal antibody may help target the radioactive substance to the right place. The antibody joined to the radioactive substance travels through the blood, locating and killing tumor cells. For example, the drug ibritumomab tiuxetan (Zevalin®) may be used for the treatment of certain types of B-cell non-Hodgkin lymphoma (NHL). The antibody part of this drug recognizes and binds to a protein found on the surface of B lymphocytes. The combination drug regimen of tositumomab and iodine I ¹³¹ tositumomab (Bexxar®) may be used for the treatment of certain types of cancer, such as NHL. In this regimen, nonradioactive tositumomab antibodies may be given to subjects first, followed by treatment with tositumomab antibodies that have ¹³¹I attached. Tositumomab may recognize and bind to the same protein on B lymphocytes as ibritumomab. The nonradioactive form of the antibody may help protect normal B lymphocytes from being damaged by radiation from ¹³¹I.

[00131] Some systemic radiation therapy drugs relieve pain from cancer that has spread to the bone (bone metastases). This is a type of palliative radiation therapy. The radioactive drugs samarium-153-

lexidronam (Quadramet®) and strontium-89 chloride (Metastron®) are examples of radiopharmaceuticals may be used to treat pain from bone metastases.

[00132] Biological therapy (sometimes called immunotherapy, biotherapy, biologic therapy, or biological response modifier (BRM) therapy) uses the body's immune system, either directly or indirectly, to fight cancer or to lessen the side effects that may be caused by some cancer treatments. Biological therapies include interferons, interleukins, colony-stimulating factors, monoclonal antibodies, vaccines, gene therapy, and nonspecific immunomodulating agents.

[00133] Interferons (IFNs) are types of cytokines that occur naturally in the body. Interferon alpha, interferon beta, and interferon gamma are examples of interferons that may be used in cancer treatment.

[00134] Like interferons, interleukins (ILs) are cytokines that occur naturally in the body and can be made in the laboratory. Many interleukins have been identified for the treatment of cancer. For example, interleukin-2 (IL-2 or aldesleukin), interleukin 7, and interleukin 12 have may be used as an anti-cancer treatment. IL-2 may stimulate the growth and activity of many immune cells, such as lymphocytes, that can destroy cancer cells. Interleukins may be used to treat a number of cancers, including leukemia, lymphoma, and brain, colorectal, ovarian, breast, kidney and prostate cancers.

[00135] Colony-stimulating factors (CSFs) (sometimes called hematopoietic growth factors) may also be used for the treatment of cancer. Some examples of CSFs include, but are not limited to, G-CSF (filgrastim) and GM-CSF (sargramostim). CSFs may promote the division of bone marrow stem cells and their development into white blood cells, platelets, and red blood cells. Bone marrow is critical to the body's immune system because it is the source of all blood cells. Because anticancer drugs can damage the body's ability to make white blood cells, red blood cells, and platelets, stimulation of the immune system by CSFs may benefit subjects undergoing other anti-cancer treatment, thus CSFs may be combined with other anti-cancer therapies, such as chemotherapy. CSFs may be used to treat a large variety of cancers, including lymphoma, leukemia, multiple myeloma, melanoma, and cancers of the brain, lung, esophagus, breast, uterus, ovary, prostate, kidney, colon, and rectum.

[00136] Another type of biological therapy includes monoclonal antibodies (MOABs or MoABs). These antibodies may be produced by a single type of cell and may be specific for a particular antigen. To create MOABs, a human cancer cells may be injected into mice. In response, the mouse immune system can make antibodies against these cancer cells. The mouse plasma cells that produce antibodies may be isolated and fused with laboratory-grown cells to create “hybrid” cells called hybridomas. Hybridomas can indefinitely produce large quantities of these pure antibodies, or MOABs. MOABs may be used in cancer treatment in a number of ways. For instance, MOABs that react with specific types of cancer may enhance a subject's immune response to the cancer. MOABs can be programmed to act against cell growth factors, thus interfering with the growth of cancer cells.

[00137] MOABs may be linked to other anti-cancer therapies such as chemotherapeutics, radioisotopes (radioactive substances), other biological therapies, or other toxins. When the antibodies latch onto cancer cells, they deliver these anti-cancer therapies directly to the tumor, helping to destroy it. MOABs carrying radioisotopes may also prove useful in diagnosing certain cancers, such as colorectal, ovarian, prostate and breast.

[00138] Rituxan® (rituximab) and Herceptin® (trastuzumab) are examples of MOABs that may be used as a biological therapy. Rituxan may be used for the treatment of non-Hodgkin lymphoma. Herceptin can be used to treat metastatic breast cancer in subjects with tumors that produce excess amounts of a protein called HER2. Alternatively, MOABs may be used to treat lymphoma, leukemia, melanoma, and cancers of the brain, breast, lung, kidney, colon, rectum, ovary, prostate, and other areas.

[00139] Cancer vaccines are another form of biological therapy. Cancer vaccines may be designed to encourage the subject's immune system to recognize cancer cells. Cancer vaccines may be designed to treat existing cancers (therapeutic vaccines) or to prevent the development of cancer (prophylactic vaccines). Therapeutic vaccines may be injected in a person after cancer is diagnosed. These vaccines may stop the growth of existing tumors, prevent cancer from recurring, or eliminate cancer cells not killed by prior treatments. Cancer vaccines given when the tumor is small may be able to eradicate the cancer. On the other hand, prophylactic vaccines are given to healthy individuals before cancer develops. These vaccines are designed to stimulate the immune system to attack viruses that can cause cancer. By targeting these cancer-causing viruses, development of certain cancers may be prevented. For example, cervarix and gardasil are vaccines to treat human papilloma virus and may prevent cervical cancer. Therapeutic vaccines may be used to treat melanoma, lymphoma, leukemia, and cancers of the brain, breast, lung, kidney, ovary, prostate, pancreas, colon, and rectum. Cancer vaccines can be used in combination with other anti-cancer therapies.

[00140] Gene therapy is another example of a biological therapy. Gene therapy may involve introducing genetic material into a person's cells to fight disease. Gene therapy methods may improve a subject's immune response to cancer. For example, a gene may be inserted into an immune cell to enhance its ability to recognize and attack cancer cells. In another approach, cancer cells may be injected with genes that cause the cancer cells to produce cytokines and stimulate the immune system.

[00141] In some instances, biological therapy includes nonspecific immunomodulating agents. Nonspecific immunomodulating agents are substances that stimulate or indirectly augment the immune system. Often, these agents target key immune system cells and may cause secondary responses such as increased production of cytokines and immunoglobulins. Two nonspecific immunomodulating agents used in cancer treatment are bacillus Calmette-Guerin (BCG) and levamisole. BCG may be used in the treatment of superficial bladder cancer following surgery. BCG may work by stimulating an inflammatory, and possibly an immune, response. A solution of BCG may be instilled in the bladder. Levamisole is sometimes used along with fluorouracil (5-FU)

chemotherapy in the treatment of stage III (Dukes' C) colon cancer following surgery. Levamisole may act to restore depressed immune function.

[00142] Photodynamic therapy (PDT) is an anti-cancer treatment that may use a drug, called a photosensitizer or photosensitizing agent, and a particular type of light. When photosensitizers are exposed to a specific wavelength of light, they may produce a form of oxygen that kills nearby cells. A photosensitizer may be activated by light of a specific wavelength. This wavelength determines how far the light can travel into the body. Thus, photosensitizers and wavelengths of light may be used to treat different areas of the body with PDT.

[00143] In the first step of PDT for cancer treatment, a photosensitizing agent may be injected into the bloodstream. The agent may be absorbed by cells all over the body but may stay in cancer cells longer than it does in normal cells. Approximately 24 to 72 hours after injection, when most of the agent has left normal cells but remains in cancer cells, the tumor can be exposed to light. The photosensitizer in the tumor can absorb the light and produces an active form of oxygen that destroys nearby cancer cells. In addition to directly killing cancer cells, PDT may shrink or destroy tumors in two other ways. The photosensitizer can damage blood vessels in the tumor, thereby preventing the cancer from receiving necessary nutrients. PDT may also activate the immune system to attack the tumor cells.

[00144] The light used for PDT can come from a laser or other sources. Laser light can be directed through fiber optic cables (thin fibers that transmit light) to deliver light to areas inside the body. For example, a fiber optic cable can be inserted through an endoscope (a thin, lighted tube used to look at tissues inside the body) into the lungs or esophagus to treat cancer in these organs. Other light sources include light-emitting diodes (LEDs), which may be used for surface tumors, such as skin cancer. PDT is usually performed as an outsubject procedure. PDT may also be repeated and may be used with other therapies, such as surgery, radiation, or chemotherapy.

[00145] Extracorporeal photopheresis (ECP) is a type of PDT in which a machine may be used to collect the subject's blood cells. The subject's blood cells may be treated outside the body with a photosensitizing agent, exposed to light, and then returned to the subject. ECP may be used to help lessen the severity of skin symptoms of cutaneous T-cell lymphoma that has not responded to other therapies. ECP may be used to treat other blood cancers, and may also help reduce rejection after transplants.

[00146] Additionally, photosensitizing agent, such as porfimer sodium or Photofrin®, may be used in PDT to treat or relieve the symptoms of esophageal cancer and non-small cell lung cancer. Porfimer sodium may relieve symptoms of esophageal cancer when the cancer obstructs the esophagus or when the cancer cannot be satisfactorily treated with laser therapy alone. Porfimer sodium may be used to treat non-small cell lung cancer in subjects for whom the usual treatments are not appropriate, and to relieve symptoms in subjects with non-small cell lung cancer that obstructs the

airways. Porfimer sodium may also be used for the treatment of precancerous lesions in subjects with Barrett esophagus, a condition that can lead to esophageal cancer.

[00147] Laser therapy may use high-intensity light to treat cancer and other illnesses. Lasers can be used to shrink or destroy tumors or precancerous growths. Lasers are most commonly used to treat superficial cancers (cancers on the surface of the body or the lining of internal organs) such as basal cell skin cancer and the very early stages of some cancers, such as cervical, penile, vaginal, vulvar, and non-small cell lung cancer.

[00148] Lasers may also be used to relieve certain symptoms of cancer, such as bleeding or obstruction. For example, lasers can be used to shrink or destroy a tumor that is blocking a subject's trachea (windpipe) or esophagus. Lasers also can be used to remove colon polyps or tumors that are blocking the colon or stomach.

[00149] Laser therapy is often given through a flexible endoscope (a thin, lighted tube used to look at tissues inside the body). The endoscope is fitted with optical fibers (thin fibers that transmit light). It is inserted through an opening in the body, such as the mouth, nose, anus, or vagina. Laser light is then precisely aimed to cut or destroy a tumor.

[00150] Laser-induced interstitial thermotherapy (LITT), or interstitial laser photocoagulation, also uses lasers to treat some cancers. LITT is similar to a cancer treatment called hyperthermia, which uses heat to shrink tumors by damaging or killing cancer cells. During LITT, an optical fiber is inserted into a tumor. Laser light at the tip of the fiber raises the temperature of the tumor cells and damages or destroys them. LITT is sometimes used to shrink tumors in the liver.

[00151] Laser therapy can be used alone, but most often it is combined with other treatments, such as surgery, chemotherapy, or radiation therapy. In addition, lasers can seal nerve endings to reduce pain after surgery and seal lymph vessels to reduce swelling and limit the spread of tumor cells.

[00152] Lasers used to treat cancer may include carbon dioxide (CO₂) lasers, argon lasers, and neodymium:yttrium-aluminum-garnet (Nd:YAG) lasers. Each of these can shrink or destroy tumors and can be used with endoscopes. CO₂ and argon lasers can cut the skin's surface without going into deeper layers. Thus, they can be used to remove superficial cancers, such as skin cancer. In contrast, the Nd:YAG laser is more commonly applied through an endoscope to treat internal organs, such as the uterus, esophagus, and colon. Nd:YAG laser light can also travel through optical fibers into specific areas of the body during LITT. Argon lasers are often used to activate the drugs used in PDT.

[00153] For subjects with systemic disease, additional treatment modalities such as adjuvant chemotherapy (e.g., docetaxel, mitoxantrone and prednisone) and systemic radiation therapy (e.g., samarium or strontium) can be designated. Such subjects would likely be treated immediately with radiation therapy in order to eliminate presumed micro-metastatic disease, which cannot be detected clinically but can be revealed by the target gene expression signature. Such subjects can also be more closely monitored for signs of disease progression.

[00154] For subjects that do not have systemic disease, localized adjuvant radiation therapy (e.g., localized to breast tissue) or endocrine therapy or chemotherapy may be administered. For subjects with no evidence of disease (NED), expression profiling and/or calculation of a risk score, as described herein, may be used to determine the risk of recurrence of the breast cancer. For subjects identified as having a low risk of recurrence of breast cancer and identified as having a high benefit of radiotherapy using the methods described herein, adjuvant radiation therapy would be recommended.

[00155] Target genes can be grouped so that information obtained about the set of target genes in the group can be used to make or assist in making a clinically relevant judgment such as a diagnosis, prognosis, or treatment choice.

[00156] A subject report is also provided comprising a representation of measured expression levels of a plurality of target genes in a biological sample from the subject, wherein the representation comprises expression levels of target genes corresponding to any one, two, three, four, five, six, eight, ten, twenty, thirty or more of the target genes, the subsets described herein, or a combination thereof. In some embodiments, the representation of the measured expression level(s) may take the form of a linear or nonlinear combination of expression levels of the target genes of interest. The subject report may be provided in a machine (e.g., a computer) readable format and/or in a hard (paper) copy. The report can also include standard measurements of expression levels of said plurality of target genes from one or more sets of subjects with known disease status and/or outcome. The report can be used to inform the subject and/or treating physician of the expression levels of the expressed target genes, the likely medical diagnosis and/or implications, and optionally may recommend a treatment modality for the subject.

[00157] Also provided are representations of the gene expression profiles useful for treating, diagnosing, prognosticating, and otherwise assessing disease. In some embodiments, these profile representations are reduced to a medium that can be automatically read by a machine such as computer readable media (magnetic, optical, and the like). The articles can also include instructions for assessing the gene expression profiles in such media. For example, the articles may comprise a readable storage form having computer instructions for comparing gene expression profiles of the portfolios of genes described above. The articles may also have gene expression profiles digitally recorded therein so that they may be compared with gene expression data from subject samples. Alternatively, the profiles can be recorded in different representational format. A graphical recordation is one such format. Clustering algorithms can assist in the visualization of such data.

EXAMPLES

Example 1: Development and Validation of a Genomic Classifier for Prognosis of Local Recurrence of Breast Cancer and Prediction of Response to Radiation Therapy.

[00158] A genomic classifier for the prognosis of local recurrence of breast cancer and prediction of response to radiation therapy (RT) was developed as follows. The SweBCG91-RT was a large breast cancer trial conducted in Sweden. Briefly, the trial randomized 1,178 patients with node-negative, stage I-IIa breast cancer to adjuvant whole-breast RT or no RT following breast conserving surgery. All patients had negative surgical margins. Hormone therapy and chemotherapy were provided according to local guidelines at the time and was only administered to 8% of patients in the original trial. Subtyping by immunohistochemical staining of ER, PgR, HER2 and KI67 was performed. Gene expression data (Gene Expression Omnibus, GSE119295) from primary tumors of SweBCG91-RT were obtained for 764 of the patients of the trial. The expression data were generated from GeneChip Human Exon 1.0 ST Arrays (Thermo Fisher Scientific, South San Francisco) in a CLIA/CAP-certified laboratory (Decipher Biosciences, San Diego, CA).

Validation dataset

[00159] The SweBCG 91-RT trial was a multi-institutional trial randomizing women with stage I-II lymph node-negative breast cancer to postoperative whole breast RT or no RT after BCS. (Sjostrom M et al. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 2017; **35**(28): 3222-9 and Malmström P et al. *European journal of cancer (Oxford, England: 1990)* 2003; **39**(12): 1690-7.) In the trial, all subsets of patients benefitted from radiation therapy and no clinical subgroup analyzed in the trial could be spared radiation, including a low-risk group of patients above 64 years of age with hormone receptor positive tumors smaller than 21 mm. (Killander F et al. *European journal of cancer (Oxford, England : 1990)* 2016; **67**: 57-65.) The SweBCG 91-RT trial randomized stage I-IIA node-negative breast cancer patients to postoperative whole breast RT or to no RT following BCS. The original study found a benefit of radiotherapy for the local and regional recurrence endpoints, but not to metastasis or breast cancer death. Out of the original 1,178 patients, primary tumor tissue derived from formalin-fixed paraffin-embedded (FFPE) samples from 765 patients were successfully profiled on the GeneChip Human Exon 1.0 ST array (ThermoFisher). Complete locoregional recurrence information was available for further analysis in 764 patients. Data is available at Gene Expression Omnibus with accession number GSE119295.

Development of Training and Validation Sets

[00160] In this analysis, only patients with ER+, HER2- tumors were included. This was to ensure that the resulting signature could distinguish risk among clinically low-risk patients who would be candidates for radiotherapy omission. Furthermore, as systemic therapy can influence locoregional recurrence, patients treated with chemotherapy and/or hormone therapy were removed from analysis. After these considerations, 597 patients of the SweBCG91-RT trial were included in this study.

[00161] For training, to improve the signal to noise ratio for identification of a locoregional recurrence event, only patients with a recorded locoregional event as first event prior to 10 years or patients

without any recorded event with a follow-up time greater than 10 years were included. 40% of these patients were retained for the training cohort and all remaining patients used for validation (Figure 1). There were no significant differences in the distribution of clinicopathologic between the No RT and RT arms of the training and validation cohorts (See Table 1 and Table 2). The lower rate of ipsilateral breast recurrence in the RT arm compared to No RT was preserved in both the training and validation cohorts (See Table 3).

Table 1. Training cohort demographics

	No RT (N=131)	RT (N=112)	Total (N=243)	P-value
Age at Surgery (years)				
Median (range)	60 (33 - 76)	55 (39 - 74)	57 (33 - 76)	0.15
≤39	5 (0.038)	1 (0.009)	6 (0.025)	
40-49	26 (0.198)	26 (0.232)	52 (0.214)	
50-59	34 (0.26)	47 (0.42)	81 (0.333)	
60-69	57 (0.435)	34 (0.304)	91 (0.374)	
≥70	9 (0.069)	4 (0.036)	13 (0.053)	
Menopausal Status				
Pre-Menopausal	33 (0.252)	24 (0.214)	57 (0.235)	0.63
Post-Menopausal	95 (0.725)	84 (0.75)	179 (0.737)	
NA	3 (0.023)	4 (0.036)	7 (0.029)	
Histologic Grade				
Grade 1	14 (0.107)	24 (0.214)	38 (0.156)	0.07
Grade 2	99 (0.756)	73 (0.652)	172 (0.708)	
Grade 3	18 (0.137)	14 (0.125)	32 (0.132)	
NA	0 (0)	1 (0.009)	1 (0.004)	
Tumor Size (mm)				
Median (range)	12 (2 - 30)	12 (2 - 23)	12 (2 - 30)	0.75
≤10	55 (0.42)	43 (0.384)	98 (0.403)	
11 - 20	71 (0.542)	64 (0.571)	135 (0.556)	
21-30	4 (0.031)	3 (0.027)	7 (0.029)	
NA	1 (0.008)	2 (0.018)	3 (0.012)	
Subtype by IHC				
Luminal A	93 (0.71)	73 (0.652)	166 (0.683)	0.41
Luminal B (HER2-)	38 (0.29)	39 (0.348)	77 (0.317)	
Adjuvant endocrine therapy				
No	131 (1)	112 (1)	243 (1)	
Adjuvant chemotherapy				
No	131 (1)	112 (1)	243 (1)	

Table 2. Validation cohort demographics

	No RT (N=178)	RT (N=176)	Total (N=354)	P-value
Age at Surgery (years)				
Median (range)	59 (33 - 78)	62.5 (38 - 78)	60 (33 - 78)	

≤39	4 (0.022)	3 (0.017)	7 (0.02)	0.69
40-49	34 (0.191)	26 (0.148)	60 (0.169)	
50-59	54 (0.303)	49 (0.278)	103 (0.291)	
60-69	56 (0.315)	64 (0.364)	120 (0.339)	
≥70	30 (0.169)	34 (0.193)	64 (0.181)	
Menopausal Status				
Pre-Menopausal	36 (0.202)	29 (0.165)	65 (0.184)	0.37
Post-Menopausal	135 (0.758)	145 (0.824)	280 (0.791)	
NA	7 (0.039)	2 (0.011)	9 (0.025)	
Histologic Grade				
Grade 1	32 (0.18)	28 (0.159)	60 (0.169)	0.36
Grade 2	108 (0.607)	118 (0.67)	226 (0.638)	
Grade 3	34 (0.191)	25 (0.142)	59 (0.167)	
NA	4 (0.022)	5 (0.028)	9 (0.025)	
Tumor Size (mm)				
Median (range)	12 (1 - 29)	12 (4 - 30)	12 (1 - 30)	
≤10	70 (0.393)	69 (0.392)	139 (0.393)	0.78
11-20	97 (0.545)	99 (0.562)	196 (0.554)	
21-30	11 (0.062)	8 (0.045)	19 (0.054)	
Subtype by IHC				
Luminal A	115 (0.646)	113 (0.642)	228 (0.644)	1.00
Luminal B (HER2-)	63 (0.354)	63 (0.358)	126 (0.356)	
Adjuvant endocrine therapy				
No	178 (1)	176 (1)	354 (1)	
Adjuvant chemotherapy				
No	178 (1)	176 (1)	354 (1)	

Table 3. Distribution of events in training and validation cohorts

Training cohort (N=243)				
	No RT (N=131)	RT (N=112)	Total	P-value
Ipsilateral breast recurrence				
No	91 (0.695)	97 (0.866)	188 (0.774)	0.0025
Yes	40 (0.305)	15 (0.134)	55 (0.226)	
Regional recurrence				
No	125 (0.954)	109 (0.973)	234 (0.963)	0.66
Yes	6 (0.046)	3 (0.027)	9 (0.037)	
Distant metastasis				
No	120 (0.916)	107 (0.955)	227 (0.934)	0.33
Yes	11 (0.084)	5 (0.045)	16 (0.066)	
Breast cancer specific death				
No	116 (0.885)	102 (0.911)	218 (0.897)	0.66
Yes	15 (0.115)	10 (0.089)	25 (0.103)	

Death				
No	89 (0.679)	84 (0.75)	173 (0.712)	0.28
Yes	42 (0.321)	28 (0.25)	70 (0.288)	
Validation cohort (N=354)				
	No RT (N=178)	RT (N=176)	Total (N=354)	P-value
Ipsilateral breast recurrence				
No	135 (0.758)	153 (0.869)	288 (0.814)	0.01
Yes	43 (0.242)	23 (0.131)	66 (0.186)	
Regional recurrence				
No	170 (0.955)	172 (0.977)	342 (0.966)	0.39
Yes	8 (0.045)	4 (0.023)	12 (0.034)	
Distant metastasis				
No	151 (0.848)	146 (0.83)	297 (0.839)	0.74
Yes	27 (0.152)	30 (0.17)	57 (0.161)	
Breast cancer specific death				
No	141 (0.792)	136 (0.773)	277 (0.782)	0.75
Yes	37 (0.208)	40 (0.227)	77 (0.218)	
Death				
No	79 (0.444)	77 (0.438)	156 (0.441)	0.99
Yes	99 (0.556)	99 (0.562)	198 (0.559)	

Gene Set Enrichment Analysis

[00162] To identify gene sets related to LRR, Gene Set Enrichment Analysis (GSEA) was performed within the 131 patients of the training cohort not treated with RT. First, for each annotated gene transcript, a Cox model to locoregional recurrence was calculated. A pre-ranked gene list was then entered into GSEA, where for each gene, a statistic was created with the following equation: $-\log(p\text{-value}) \times (\text{if } HR > 1, 1, \text{ else } -1)$, where p-value and hazard ratio (HR) are from a Cox model using gene expression. Thus, the Cox model p-value affects the magnitude of the statistic, where the smaller the p-value, the greater the value, and the sign of the value is dependent on the hazard ratio, where genes with $HR > 1$ have positive values and genes with $HR < 1$ have negative values. Individual genes highly prognostic for increased locoregional recurrence in patients will have large positive values, and individual genes highly prognostic for decreased locoregional recurrence in patients will have large negative values. The ranked list was entered into GSEA (version 4.0.3) and the Molecular Signatures Database hallmarks (H), curated gene sets (C2), and ontology gene sets (C5) collections were specified (version 7.2). Collections were analyzed separately to identify trends across collections. Genes sets with positive enrichment scores are interpreted as associated with worse prognosis (increase of LRR), and gene sets with negative enrichment scores are associated with better prognosis (decrease of LRR).

Model Development

[00163] From the GSEA analysis, genes from gene lists with a false discovery rate $q\text{-value} < 0.05$ and in the top ten of positive and negative gene lists of the hallmarks, curated gene sets, and ontology gene sets (C5) collections (Table 4) were identified as potential candidates for the model. These nominated genes, along with ER, PR, HER2 and KI67, were intersected with genes that met the following requirements: genes with a standard deviation in the top quartile of standard deviation values and genes with a $p\text{-value} < 0.05$ when evaluated in a Cox model to the LRR endpoint in patients not treated with RT in the training set (Figure 2).

[00164] The resulting 82 genes were fed into an elastic net regression model (R glmnet, $\alpha = 0.5$) for training with the 131 patients not treated with radiation in the training cohort, using LRR as the endpoint. Elastic net regression modeling resulted in a final set of 16 genes for the model.

Table 4. Gene lists identified by GSEA as highly positively and negatively enriched.

	Positive (associated with worse prognosis)	Negative (associated with better prognosis)
H	HALLMARK_E2F_TARGETS HALLMARK_MYC_TARGETS_V2 HALLMARK_G2M_CHECKPOINT HALLMARK_MYC_TARGETS_V1 HALLMARK_ESTROGEN_RESPONSE_LATE HALLMARK_MYOGENESIS HALLMARK_OXIDATIVE_PHOSPHORYLATION HALLMARK_ESTROGEN_RESPONSE_EARLY HALLMARK_MTORC1_SIGNALING HALLMARK_HEDGEHOG_SIGNALING	HALLMARK_TNFA_SIGNALING_VIA_NFKB HALLMARK_KRAS_SIGNALING_UP HALLMARK_ALLOGRAFT_REJECTION
C2	SOTIRIOU_BREAST_CANCER_GRADE_1_VS_3_UP ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER NIKOLSKY_BREAST_CANCER_16P13_AMPLICON ZHOU_CELL_CYCLE_GENES_IN_IR_RESPONSE_24HR BLANCO_MELO_BRONCHIAL_EPITHELIAL_CELLS_INFLUENZA_A_DEL_NS1_INFECTION_DN DUTERTRE ESTRADIOL_RESPONSE_24HR_UP WHITEFORD_PEDIATRIC_CANCER_MARKERS ZHAN_MULTIPLE_MYELOMA_PR_UP LEE_EARLY_T_LYMPHOCYTE_UP CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP	SMID_BREAST_CANCER_NORMAL_LIKE_UP WINTER_HYPOXIA_DN PHONG_TNF_TARGETS_UP ZHENG_FOXP3_TARGETS_IN_T_LYMPHOCYTE_DN KEGG_GRAFT_VERSUS_HOST_DISEASE BOQUEST_STEM_CELL_CULTURED_VS_FRESH_DN BERTUCCI_INVASIVE_CARCINOMA_DUCTAL_VS_LOBULAR_DN NAKAYAMA_SOFT_TISSUE_TUMORS_PCA2_DN REACTOME_CHEMOKINE_RECEPTORS_BIND_CHEMOKINES KEGG_AUTOIMMUNE_THYROID_DISEASE
C5	GO_ANAPHASE_PROMOTING_COMPLEX_DEPENDENT_CATABOLIC_PROCESS GO_DNA_REPLICATION_INDEPENDENT_NUCLEOSOME_ORGANIZATION GO_CENTROMERE_COMPLEX_ASSEMBLY GO_CHROMATIN_REMODELING_AT_CENTROMERE	GO_B_CELL_RECEPTOR_SIGNALING_PATHWAY GO_POSITIVE_REGULATION_OF_INTERLEUKIN_2_PRODUCTION GO_INTERLEUKIN_1_BETA_PRODUCTION GO_REGULATION_OF_ANTIGEN_RECEPTOR_MEDIATED_SIGNALING_PATHWAY GO_NEGATIVE_REGULATION_OF_SMOOTH_MUSCLE_CELL_PROLIFERATION GO_INTERLEUKIN_1_PRODUCTION GO_REGULATION_OF_B_CELL_RECEPTOR_SIGNALING_PATHWAY GO_RESPONSE_TO_CHEMOKINE GO_LYMPHOCYTE_CHEMOTAXIS GO_DENDRITIC_CELL_CHEMOTAXIS

Statistical Methods

[00165] Statistical analyses were performed using R, version 4.0.2 (<https://cran.r-project.org/bin/windows/base/old/4.0.2/>). The primary end point was cumulative incidence of LRR using time to LRR as first event. Cumulative incidences were computed using a competing risk approach (R cmprsk package). Distant metastasis and death without recurrence were considered competing events. Eight patients had synchronous distant metastasis and LRR, defined as LRR registered at the same time or within 3 months as the metastasis, and they were regarded as having an LRR. Cause-specific Cox proportional hazards regression models were used to infer the relationship of genomic and clinical variables on LRR, and P values for differences between groups were calculated with the Wald test or the equivalent log-rank test. Hazard ratios are reported with 95% CIs. Classifier scores were tested as continuous scores.

Gene Associated with Locoregional Recurrence

[00166] We used GSEA to identify biological concepts related to locoregional recurrence in breast cancer. The top positively enriched gene lists across the H, C2 and C5 collections included many involving cell cycle and proliferation (Table 4), whereas top negatively enriched gene lists included many related to the immune system.

[00167] Top gene lists were intersected with genes individually prognostic for LRR in the patients of the training set that were not given RT. The resulting 16-gene signature includes genes that incorporate the concepts that are broadly shown from the GSEA lists (Table 5).

Table 5. 16 genes in the model

AGR2	Anterior Gradient 2, Protein Disulphide Isomerase
B4GALT1	Beta-1,4-Galactosyltransferase 1
CLDN7	Claudin 7
EZR	Ezrin
GNG11	G Protein Subunit Gamma 11
JUN	Jun Proto-Oncogene
MMP11	Matrix Metalloproteinase 11
PKIB	CAMP-Dependent Protein Kinase Inhibitor Beta
PRPS1	Phosphoribosyl Pyrophosphate Synthetase 1
PSMD10	Proteasome 26S Subunit, Non-ATPase 10
SH3BP5	SH3 Domain Binding Protein 5
SLC16A3	Solute Carrier Family 16 Member 3
SLC7A11	Solute Carrier Family 7 Member 11
SPP1	Secreted Phosphoprotein 1
TNNT1	Troponin T1, Slow Skeletal Type

UBE2E1	Ubiquitin Conjugating Enzyme E2 E1
--------	------------------------------------

Signature is prognostic for LRR in patients not treated with radiotherapy

[00168] In the validation cohort, the signature was prognostic for locoregional recurrence in patients not treated with radiotherapy in a univariable Cox model to locoregional recurrence as the endpoint (HR=1.7 [1.3,2.2], $p<0.001$). The signature remained prognostic in a multivariable model including age, grade, tumor size, and luminal A vs luminal B (HR=1.7 [1.2,2.3], $p<0.001$) (Table 6). Patients categorized as low-risk by the signature had a 10-year locoregional recurrence rate of 7% even in the absence of radiotherapy (Figure 3).

Table 6

Univariable Analysis		HR [95%CI]	P-value
RT Omission Model		1.7 [1.3,2.2]	<0.001
Multivariable Analysis		HR [95%CI]	P-value
RT Omission Model		1.7 [1.2,2.3]	<0.001
Age		0.98 [0.95, 1]	0.1
Grade	1	Reference	
	2	1.1 [0.43, 2.9]	0.82
	3	0.74 [0.23, 2.4]	0.61
Tumor size	≤2cm	Reference	
	>2cm, ≤5cm	0.41 [0.055, 3.1]	0.39
Subtype	Luminal A	Reference	
	Luminal B	1.5 [0.75, 2.8]	0.27

Differential effect of radiotherapy on patients with low and high signature scores

[00169] Notably, there was no significant benefit from RT for these low-risk patients (HR=1.1 [0.38,3.3], $p=0.83$) (Figure 4A), whereas patients categorized as high-risk by the signature had a significant decreased risk of locoregional recurrence when treated with RT (HR=0.43 [0.24,0.78], $p=0.0053$) (Figure 4B).

[00170] These results showed that methods and classifiers of the present disclosure are prognostic for locoregional recurrence and identify breast cancer patients who would not benefit from adjuvant radiotherapy (radiation treatment). The results also showed that the methods and classifiers of the disclosure can identify breast cancer patients who have a low risk of locoregional recurrence. These

results further showed that the methods and classifiers of the present are useful for treating breast cancer.

[00171] The genomic classifier of the present disclosure identified patients at low risk of LRR without adjuvant radiotherapy post-breast conserving surgery, thus identifying them as candidates for omission of radiotherapy after breast conserving surgery. Further, the genomic classifier of the present disclosure showed that breast cancer patients of lowest risk of LRR have no significant benefit from RT.

[00172] All references cited herein are hereby incorporated by reference herein in their entirety.

CLAIMS

What is claimed is:

1. A method for prognosing and/or predicting benefit from adjuvant radiotherapy in a subject having breast cancer, the method comprising:
 - a) obtaining or having obtained an expression level in a sample from a subject for a plurality of genes, wherein the plurality of genes are selected from Table 5; and
 - b) determining that the subject is at low risk or high risk of cancer recurrence based on the expression level, and likely to benefit or not likely benefit from adjuvant radiotherapy based on the expression level, thereby prognosing and predicting benefit from adjuvant radiotherapy in the subject.
2. The method of claim 1, further comprising: administering adjuvant radiotherapy therapy if the subject is identified as likely to benefit from adjuvant radiotherapy, and recommending treatment intensification if the subject is not identified as likely to benefit from adjuvant radiotherapy.
3. The method of claim 1, wherein the cancer recurrence is local or locoregional recurrence or distant recurrence (metastasis).
4. The method of claim 1, further comprising determining that the subject is at low risk of cancer recurrence based on the age of the subject, or determining that the subject is not at low risk of cancer recurrence based on the age of the subject.
5. The method of claim 1, further comprising calculating a risk score for the subject, wherein adjuvant radiotherapy is administered to the subject if the subject is identified as being as likely to benefit from adjuvant radiotherapy based on both the risk score and the expression levels of the genes selected from Table 5 in the biological sample, and recommending treatment omission if the subject is not identified as being likely to benefit from adjuvant radiotherapy based on both the risk score and the expression levels of the genes selected from Table 5 in the biological sample.
6. The method of claim 1, wherein the plurality of genes are selected from the group consisting of Anterior Gradient 2, Protein Disulphide Isomerase (AGR2), Beta-1,4-Galactosyltransferase 1 (B4GALT1), Claudin 7 (CLDN7), Ezrin (EZR), G Protein Subunit Gamma 11 (GNG11), Jun Proto Oncogene (JUN), Matrix Metalloproteinase 11 (MMP11), CAMP-Dependent Protein Kinase Inhibitor Beta (PKIB), Phosphoribosyl Pyrophosphate Synthetase 1 (PRPS1), Proteasome 26S Subunit, Non ATPase 10 (PSMD10), SH3 Domain Binding Protein 5 (SH3BP5), Solute Carrier Family 16 Member

3 (SLC16A3), Solute Carrier Family 7 Member 11 (SLC7A11), Secreted Phosphoprotein 1 (SPP1), Troponin T1, Slow Skeletal Type (TNNT1), and Ubiquitin Conjugating Enzyme E2 E1 (UBE2E1).

7. The method of claim 1, wherein the biological sample is a biopsy or a tumor sample.
8. The method of claim 1, wherein the subject is a human being.
9. The method of claim 1, further comprising treating the subject with mastectomy, radiation boost, or adjuvant systemic therapy.
10. The method of claim 1, wherein the subject is treated with radiotherapy following breast conserving surgery (BCS).
11. The method of claim 1, wherein the subject has estrogen receptor positive (ER+) breast cancer, human epidermal growth factor receptor 2 negative (HER2-) breast cancer, Stage I-II breast cancer, or node negative breast cancer and/or is post-menopausal.
12. The method of claim 1, wherein the expression levels of all of the genes selected from Table 5 are measured in the biological sample.
13. The method of claim 1, wherein the levels of expression of one or more of the genes selected from Table 5 are increased or reduced compared to a control.
14. The method of claim 1, wherein said measuring the levels of expression comprises performing *in situ* hybridization, a PCR-based method, an array-based method, an immunohistochemical method, an RNA assay method, or an immunoassay method.
15. The method of claim 13, wherein said measuring the levels of expression comprises using a reagent selected from the group consisting of a nucleic acid probe, one or more nucleic acid primers, and an antibody.
16. The method of claim 1, wherein said measuring the level of expression comprises measuring the level of an RNA transcript.
17. The method of claim 1, wherein the method is performed prior to treatment of the subject with adjuvant chemotherapy, endocrine therapy, or radiotherapy.

18. A kit for determining a prognosis of a subject having breast cancer and whether or not to treat the subject with adjuvant radiotherapy, the kit comprising agents for measuring levels of expression of a plurality of genes selected from Table 5.
19. The kit of claim 17, wherein the kit comprises agents for measuring the levels of expression of all of the genes listed in Table 5.
20. The kit of claim 17, wherein said agents comprise reagents for performing *in situ* hybridization, a PCR-based method, an array-based method, an immunohistochemical method, an RNA assay method, or an immunoassay method.
21. The kit of claim 17, wherein said agents comprise one or more of a microarray, a nucleic acid probe, a nucleic acid primer, or an antibody.
22. The kit of claim 17, wherein the kit comprises at least one set of PCR primers capable of amplifying a nucleic acid comprising a sequence of a gene selected from Table 5 or its complement.
23. The kit of claim 17, wherein the kit comprises at least one probe capable of hybridizing to a nucleic acid comprising a sequence of a gene selected from Table 5 or its complement.
24. The kit of claim 17, further comprising information, in electronic or paper form, comprising instructions on how to determine the prognosis of a subject having breast cancer and whether or not to treat the subject with adjuvant radiotherapy.
25. The kit of claim 17, further comprising one or more control reference samples.
26. A probe set for determining a prognosis of a subject having breast cancer and whether or not to treat the subject with adjuvant radiotherapy, the probe set comprising a plurality of probes for detecting a plurality of target nucleic acids, wherein the plurality of target nucleic acids comprises one or more gene sequences, or complements thereof, of genes selected from Table 5.
27. The probe set of claim 25, wherein the probe set comprises a plurality of probes for detecting a plurality of target nucleic acids comprising gene sequences, or complements thereof, of the genes selected from Table 5.
28. The probe set of claim 26, wherein at least one probe is detectably labeled.

29. A kit for determining a prognosis of a subject having breast cancer and whether or not to treat the subject with adjuvant radiotherapy, chemotherapy, or endocrine therapy, the kit comprising the probe set of claim 27.
30. A system comprising:
 - a) the probe set of claim 25; and
 - b) a computer model or algorithm for analyzing an expression level or expression profile of the plurality of target nucleic acids hybridized to the plurality of probes in a biological sample from the subject who has breast cancer and determining if the subject is at low risk of cancer recurrence based on the expression level or expression profile and predicting whether the subject will benefit from adjuvant radiotherapy.
31. A kit comprising the system of claim 29.
32. The kit of claim 30, further comprising a computer model or algorithm for designating a treatment modality for the subject.
33. The kit of claim 30, further comprising a computer model or algorithm for normalizing the expression level or expression profile of the plurality of target nucleic acids.

ABSTRACT

[00173] The present disclosure pertains to the field of personalized medicine and methods for treating breast cancer. In particular, the disclosure relates to the use of transcriptomic profiling for prognosis of cancer recurrence (e.g., locoregional recurrence) in subjects with breast cancer and methods for identifying subjects who are not likely to benefit from radiotherapy.