MARKERS FOR THE EARLY DETECTION OF COLON CELL PROLIFERATIVE DISORDERS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 63/128,545, filed December 21, 2020, which is incorporated by reference herein in its entirety.

FIELD

[0002] This disclosure is related to biomarkers and methods for the early identification of colon cell proliferative disorders including advanced adenoma and colorectal cancer.

BACKGROUND

[0003] Colorectal cancer (CRC) is the leading cause of cancer-related mortality in the western world. Although CRC is one of the best characterized solid tumors, CRC continues to be one of the main causes of death in developed countries because of late diagnosis. Among other reasons, late diagnosis of patients is due to the fact that diagnostic tests, such as colonoscopy, are performed too late. Deaths from CRC can be prevented through effective and early screening.

SUMMARY

[0004] The present disclosure provides methods and systems directed to protein profiling of biological samples associated with CRC detection and disease progression. Studies described herein enable determination of the presence of a specific protein signature of CRC showing the presence of specific biomarkers of the colon cell proliferative disorders with potential for detecting colon cell proliferative disorders, stratifying patient populations, and classifying populations using plasma from subjects with a colon cell proliferative disorder with high specificity and high sensitivity.

[0004a] In one aspect, provided herein is a method for detecting a colon cell proliferative disorder in a subject, comprising:

a) measuring, in a biological sample containing proteins obtained from the subject, an amount of a protein from a predetermined protein panel comprising at least 6 of proteins selected from the group consisting of: EGF, FGF-2, FLT3L, Fractalkine, IL-1a, IL-2, IL-6, IL-8, GROa, MIP-3a, Complement component C2, Complement component C9, Factor D, Factor I, MBL, MMP-2, GDF-15, Osteonectin, Periostin, ANGPTL4, FGF-21, FGF-23, HGF, Angiopoietin-2, BMP-9, IL-1RII, HGFR, IL-6ra, OPN, Tenascin-C, Thrombospondin-2, uPAR, CD44, Kallikrein-6, Mesothelin, EpCAM, Apo A1, AGP, A2MB, Fetuin A, HP, L-Selectin, Complement component C1q, Complement component C3, Complement component C3b, Factor B, Factor H, Properdin, AGRP, MMP-12,

CYFRA21-1, HE4, total PSA, MIF, AFP, CA125, CA19-9, CA15-3 (MUC-1), and CEA, to provide a protein profile of the subject; and

b) processing the protein profile using a machine learning model trained to be capable of distinguishing between subjects with the colon cell proliferative disorder and subjects without the colon cell proliferative disorder to provide an output value associated with presence or absence of the colon cell proliferative disorder, thereby indicating the presence or absence of the colon cell proliferative disorder in the subject.

[0004b] In another aspect, provided herein is a method for determining a protein profile of a biological sample from a subject comprising:

measuring, in a biological sample containing proteins obtained from the subject, an amount of a protein from a predetermined panel of proteins comprising at least 6 proteins selected from the group consisting of: EGF, FGF-2, FLT3L, Fractalkine, IL-1a, IL-2, IL-6, IL-8, GROa, MIP-3a, Complement component C2, Complement component C9, Factor D, Factor I, MBL, MMP-2, GDF-15, Osteonectin, Periostin, ANGPTL4, FGF-21, FGF-23, HGF, Angiopoietin-2, BMP-9, IL-1RII, HGFR, IL-6ra, OPN, Tenascin-C, Thrombospondin-2, uPAR, CD44, Kallikrein-6, Mesothelin, EpCAM, Apo A1, AGP, A2MB, Fetuin A, HP, L-Selectin, Complement component C1q, Complement component C3, Complement component C3b, Factor B, Factor H, Properdin, AGRP, MMP-12, CYFRA21-1, HE4, total PSA, MIF, AFP, CA125, CA19-9, CA15-3 (MUC-1), and CEA to provide the protein profile of the subject.

[0004c] In another aspect, provided herein is a classifier configured to distinguish a population of subjects having a colon cell proliferative disorder from subjects not having the colon cell proliferative disorder, comprising: sets of measured values representative of proteins from a predetermined protein panel characteristic of the colon cell proliferative disorder, wherein the sets of measured values are obtained from protein expression data from subjects having the colon cell proliferative disorder and subjects not having the colon cell proliferative disorder, wherein the measured values are used to generate a set of features corresponding to properties of the protein expression data, wherein the set of features is computer-processed using a machine learning or statistical model, and wherein the machine learning or statistical model provides a feature vector useful as a classifier capable of distinguishing a population of subjects having the colon cell proliferative disorder from subjects not having the colon cell proliferative disorder.

[0005] In an aspect, the present disclosure provides a predetermined protein panel characteristic of a colon cell proliferative disorder comprising at least 6 proteins selected from the group consisting of: EGF, FGF-2, FLT3L, Fractalkine, IL-1a, IL-2, IL-6, IL-8, GROa, MIP-3a,

Complement component C2, Complement component C9, Factor D, Factor I, MBL, MMP-2, GDF-15, Osteonectin, Periostin, ANGPTL4, FGF-21, FGF-23, HGF, Angiopoietin-2, BMP-9, IL-1RII (IL1R2), HGFR, IL-6ra, OPN, Tenascin-C (TNC), Thrombospondin-2 (THBS2), uPAR, CD44, Kallikrein-6, Mesothelin, EpCAM, Apo A1, AGP, A2MB, Fetuin A, HP, L-Selectin, Complement component C1q, Complement component C3, Complement component C3b, Factor B, Factor H, Properdin, AGRP, MMP-12, CYFRA21-1, HE4, total PSA, MIF, AFP, CA125 (MUC-16), CA19-9, CA15-3 (MUC-1), and CEA.