

## **Supplemental Material**

Supplemental Table 1. Protein names and respective Olink® panel sorted in alphabetical order

| Protein full name  | Entry name        | Olink® Panel* | Uniprot ID** | >25% undetectable values |
|--|-------------------|---------------|--------------|--------------------------|
| Angiotensin-converting enzyme 2                                  | ACE2              | CVD II        | Q9BYF1       |                          |
| Adisintegrin and metalloproteinase with thrombospondin motifs 13 | ADAMTS13          | CVD II        | Q76LX8       |                          |
| ADM  | ADM               | CVD II        | P35318       |                          |
| Agouti-related protein   | AGRP              | CVD II        | O00253       |                          |
| Protein AMBP   | AMBP              | CVD II        | P02760       |                          |
| Angiopoietin-1   | ANG1              | CVD II        | Q15389       |                          |
| Bone morphogenetic protein 6                                     | BMP6              | CVD II        | P22004       |                          |
| Natriuretic peptides B   | BNP               | CVD II        | P16860       | X                        |
| Carbonic anhydrase 5A, mitochondrial                             | CA5A              | CVD II        | P35218       | X                        |
| C-C motif chemokine 3  | CCL3              | CVD II        | P10147       |                          |
| T-cell surface glycoprotein CD4                                  | CD4               | CVD II        | P01730       |                          |
| CD40 ligand  | CD40L             | CVD II        | P29965       |                          |
| SLAM family member 5   | CD84              | CVD II        | Q9UIB8       |                          |
| Carcinoembryonic antigen-related cell adhesion molecule 8        | CEACAM8           | CVD II        | P31997       |                          |
| Chymotrypsin C   | CTRC              | CVD II        | Q99895       |                          |
| Cathepsin L1   | CTSL1             | CVD II        | P07711       |                          |
| C-X-C motif chemokine 1 (CVD2)                                   | CXCL1             | CVD II        | P09341       |                          |
| Decorin  | DCN               | CVD II        | P07585       |                          |
| 2,4-dienoyl-CoA reductase, mitochondrial                         | DECR1             | CVD II        | Q16698       |                          |
| Dickkopf-related protein 1                                       | DKK1              | CVD II        | O94907       |                          |
| Fatty acid-binding protein, intestinal                           | FABP2             | CVD II        | P12104       |                          |
| Fibroblast growth factor 21 (CVD2)                               | FGF21             | CVD II        | Q9NSA1       |                          |
| Fibroblast growth factor 23 (CVD2)                               | FGF23             | CVD II        | Q9GZV9       |                          |
| Follistatin  | FS                | CVD II        | P19883       |                          |
| Galectin-9   | GAL9              | CVD II        | O00182       |                          |
| Growth/differentiation factor 2                                  | GDF2              | CVD II        | Q9UK05       |                          |
| Growth hormone   | GH                | CVD II        | P01241       |                          |
| Gastric intrinsic factor   | GIF               | CVD II        | P27352       |                          |
| Lactoylglutathione lyase   | GLO1              | CVD II        | Q04760       |                          |
| Gastrotropin   | GT                | CVD II        | P51161       | X                        |
| Hydroxyacid oxidase 1  | HAOX1             | CVD II        | Q9UJM8       |                          |
| Proheparin-binding EGF-like growth factor                        | HBEGF             | CVD II        | Q99075       |                          |
| Heme oxygenase 1   | HO1               | CVD II        | P09601       |                          |
| Osteoclast-associated immunoglobulin-like receptor               | HOSCAR            | CVD II        | Q8IYS5       |                          |
| Heat shock 27 kDa protein  | HSP27             | CVD II        | P04792       |                          |
| Alpha-L-iduronidase  | IDUA              | CVD II        | P35475       |                          |
| Low affinity immunoglobulin gamma Fc region receptor II-b        | IGGFCRECEPTOR IIB | CVD II        | P31994       |                          |
| Pro-interleukin-16   | IL16              | CVD II        | Q14005       |                          |
| Interleukin-17D  | IL17D             | CVD II        | Q8TAD2       | X                        |
| Interleukin-18 (CVD2)  | IL18              | CVD II        | Q14116       |                          |
| Interleukin-1 receptor antagonist protein                        | IL1RA             | CVD II        | P18510       | X                        |
| Interleukin-1 receptor-like 2                                    | IL1RL2            | CVD II        | Q9HB29       |                          |
| Interleukin-27   | IL27              | CVD II        | Q8NEV9       |                          |
| Interleukin-4 receptor subunit alpha                             | IL4RA             | CVD II        | P24394       |                          |
| Interleukin-6 (CVD2)   | IL6               | CVD II        | P05231       |                          |
| Melusin  | ITGB1BP2          | CVD II        | Q9UKP3       | X                        |
| Kidney injury molecule 1   | KIM1              | CVD II        | Q96D42       |                          |

|   |              |        |        |   |
|---|--------------|--------|--------|---|
| Leptin  | LEP          | CVD II | P41159 |   |
| Lectin-like oxidized LDL receptor 1                   | LOX1         | CVD II | P78380 |   |
| Lipoprotein lipase                                    | LPL          | CVD II | P06858 |   |
| Macrophage receptor MARCO                             | MARCO        | CVD II | Q9UEW3 |   |
| Tyrosine-protein kinase Mer                           | MERTK        | CVD II | Q12866 |   |
| Matrix metalloproteinase-12                           | MMP12        | CVD II | P39900 |   |
| Matrix metalloproteinase-7                            | MMP7         | CVD II | P09237 |   |
| NF-kappa-B essential modulator                        | NEMO         | CVD II | Q9Y6K9 |   |
| Pappalysin-1  | PAPPA        | CVD II | Q13219 | X |
| Proteinase-activated receptor 1                       | PAR1         | CVD II | P25116 |   |
| Poly [ADP-ribose] polymerase 1                        | PARP1        | CVD II | P09874 | X |
| Platelet-derived growth factor subunit B              | PDGFSUBUNITB | CVD II | P01127 |   |
| Programmed cell death 1 ligand 2                      | PDL2         | CVD II | Q9BQ51 |   |
| Polymeric immunoglobulin receptor                     | PIGR         | CVD II | P01833 |   |
| Placenta growth factor                                | PLGF         | CVD II | P49763 |   |
| Prolargin   | PRELP        | CVD II | P51888 |   |
| Brother of CDO  | PROTEINBOC   | CVD II | Q9BWW1 |   |
| Serine protease 27                                    | PRSS27       | CVD II | Q9BQR3 |   |
| Prostasin   | PRSS8        | CVD II | Q16651 |   |
| P-selectin glycoprotein ligand 1                      | PSGL1        | CVD II | Q14242 |   |
| Pentraxin-related protein PTX3                        | PTX3         | CVD II | P26022 |   |
| Receptor for advanced glycosylation end products      | RAGE         | CVD II | Q15109 |   |
| Renin   | REN          | CVD II | P00797 |   |
| Stem cell factor (CVD2)                               | SCF          | CVD II | P21583 |   |
| Serpin A12  | SERPINA12    | CVD II | Q8IW75 | X |
| SLAM family member 7                                  | SLAMF7       | CVD II | Q9NQ25 | X |
| Superoxide dismutase [Mn], mitochondrial              | SOD2         | CVD II | P04179 |   |
| Sortilin  | SORT1        | CVD II | Q99523 |   |
| Spondin-2   | SPON2        | CVD II | Q9BUD6 |   |
| Proto-oncogene tyrosine-protein kinase Src            | SRC          | CVD II | P12931 |   |
| Serine/threonine-protein kinase 4                     | STK4         | CVD II | Q13043 |   |
| Tissue factor   | TF           | CVD II | P13726 |   |
| Protein-glutamine gamma-glutamyltransferase 2         | TGM2         | CVD II | P21980 |   |
| Thrombospondin-2                                      | THBS2        | CVD II | P35442 |   |
| Thrombopoietin  | THPO         | CVD II | P40225 |   |
| Angiopoietin-1 receptor                               | TIE2         | CVD II | Q02763 |   |
| Thrombomodulin  | TM           | CVD II | P07204 |   |
| Tumor necrosis factor receptor superfamily member 10A | TNFRSF10A    | CVD II | O00220 | X |
| Tumor necrosis factor receptor superfamily member 11A | TNFRSF11A    | CVD II | Q9Y6Q6 |   |
| Tumor necrosis factor receptor superfamily member 13B | TNFRSF13B    | CVD II | O14836 |   |
| TNF-related apoptosis-inducing ligand receptor 2      | TRAILR2      | CVD II | O14763 |   |
| Vascular endothelial growth factor D                  | VEGFD        | CVD II | O43915 |   |
| V-set and immunoglobulin domain-containing protein 2  | VSIG2        | CVD II | Q96IQ7 | X |
| Lymphotoxin   | XCL1         | CVD II | P47992 |   |

Supplemental Table 2. Biomarkers selected by the LASSO procedure performed in hypertensive subjects vs. matched controls

|                         | Lasso procedure |      |
|-------------------------|-----------------|------|
|                         | n               | %    |
| Gdf15 > T2 (836.98)     | 941             | 94.1 |
| NGAL > T2 (427.61)      | 163             | 16.3 |
| PIIINP > T2 (4.70)      | 81              | 8.1  |
| Sdf1a > T1 (1455.95)    | 230             | 23   |
| BMP 6 > T1 (3.82)       | 81              | 8.1  |
| ADM > median (5.99)     | 178             | 17.8 |
| PGF > T2 (6.96)         | 247             | 24.7 |
| BOC                     | 999             | 99.9 |
| IL6 > median (2.23)     | 178             | 17.8 |
| TNFRSF11A > T2 (4.14)   | 167             | 16.7 |
| PAR 1                   | 150             | 15   |
| TRAIL R2 > T1 (4.00)    | 167             | 16.7 |
| TF > median (4.62)      | 168             | 16.8 |
| IL1RL2 > median (2.78)  | 135             | 13.5 |
| IL 27 > median (2.16)   | 190             | 19   |
| SCF                     | 999             | 99.9 |
| FGF 21                  | 999             | 99.9 |
| FGF 23 > median (2.28)  | 174             | 17.4 |
| GH                      | 998             | 99.8 |
| GLO1 > median (3.32)    | 70              | 7    |
| CD84                    | 10              | 1    |
| REN > T1 (5.96)         | 206             | 20.6 |
| DECR1 > T1 (2.90)       | 168             | 16.8 |
| KIM1 > T1 (6.41)        | 128             | 12.8 |
| THBS2                   | 27              | 2.7  |
| SORT1                   | 999             | 99.9 |
| CEACAM8 > median (2.49) | 115             | 11.5 |
| FABP2 > T1 (7.34)       | 820             | 82   |
| CTSL1 > T1 (3.79)       | 283             | 28.3 |
| hOSCAR > T1 (9.10)      | 344             | 34.4 |
| TGM2                    | 969             | 96.9 |
| LEP                     | 1000            | 100  |
| HSP 27                  | 230             | 23   |
| NEMO                    | 860             | 86   |
| VEGFD > T2 (6.69)       | 206             | 20.6 |
| HAOX1                   | 230             | 23   |

In green are the biomarkers selected >80% of the times in the 1000x bootstrap model

Supplemental Table 3. Biomarkers selected by the LASSO procedure performed in obese subjects vs. matched controls

|                         | Lasso procedure |      |
|-------------------------|-----------------|------|
|                         | n               | %    |
| Sdf1a > T1 (1303.36)    | 639             | 63.9 |
| TNFRSF11A               | 9               | 0.9  |
| CXCL1 > median (7.57)   | 639             | 63.9 |
| LOX 1 > median (4.94)   | 1000            | 100  |
| FGF 21                  | 1000            | 100  |
| TM                      | 980             | 98   |
| PRELP                   | 988             | 98.8 |
| XCL1 > T2 (3.64)        | 791             | 79.1 |
| CEACAM8 > median (2.42) | 999             | 99.9 |
| PRSS8                   | 1000            | 100  |
| GDF 2 > T2 (3.50)       | 986             | 98.6 |
| ACE2                    | 1000            | 100  |
| LEP                     | 1000            | 100  |
| HSP 27                  | 1000            | 100  |
| VEGFD                   | 1000            | 100  |
| HAOX1 > T2 (3.99)       | 996             | 99.6 |

In green are the biomarkers selected >80% of the times in the 1000x bootstrap model

Supplemental Table 4. Biomarkers selected by the LASSO procedure performed in diabetic subjects vs. matched controls

|                                      | Lasso procedure |      |
|--------------------------------------|-----------------|------|
|                                      | n               | %    |
| Gdf15                                | 1000            | 100  |
| Mmp1 > T1 (285.90)                   | 158             | 15.8 |
| NGAL                                 | 999             | 99.9 |
| PIINP > median (4.50)                | 37              | 3.7  |
| Sdf1a > median (1631.55)             | 275             | 27.5 |
| BOC > T1 (3.18)                      | 252             | 25.2 |
| PAR 1 > T1 (6.67)                    | 28              | 2.8  |
| CXCL1                                | 1000            | 100  |
| SCF > T2 (8.20)                      | 998             | 99.8 |
| CTRC > T1 (8.74)                     | 136             | 13.6 |
| GH > median (8.13)                   | 708             | 70.8 |
| REN > median (6.50)                  | 999             | 99.9 |
| KIM1 > T1 (6.23)                     | 116             | 11.6 |
| XCL1 > T1 (3.38)                     | 37              | 3.7  |
| SORT1 > median (6.82)                | 83              | 8.3  |
| CCL17 > T2 (6.95)                    | 60              | 6    |
| IgG Fc receptor II b > median (1.81) | 275             | 27.5 |
| LPL                                  | 1000            | 100  |
| AGRP > T1 (2.30)                     | 275             | 27.5 |
| FABP2 > median (7.73)                | 890             | 89   |
| THPO > T2 (1.75)                     | 549             | 54.9 |
| MMP12                                | 373             | 37.3 |
| TNFRSF13B                            | 978             | 97.8 |
| HSP 27 > T2 (9.94)                   | 35              | 3.5  |
| CD4                                  | 632             | 63.2 |
| NEMO                                 | 69              | 6.9  |

In green are the biomarkers selected >80% of the times in the 1000x bootstrap model

Supplemental Table 5. Biomarkers identified in both methods and biomarker overlap (green).

| Biomarker     | Bioprofile of phenotypic groups |         |          |
|---------------|---------------------------------|---------|----------|
|               | HTN                             | Obesity | Diabetes |
| CD4           |                                 | X       |          |
| CEACAM-8      |                                 | +       |          |
| CXCL-1        |                                 | X       | +        |
| FABP-2        | +                               |         |          |
| GDF-15        | •                               |         | +        |
| IL-1 $\alpha$ | X                               | X       |          |
| LEP           | •                               | •       |          |
| LPL           | X                               | X       |          |
| PRELP         | X                               | •       |          |
| PSGL-1        | X                               | X       |          |
| REN           | X                               | X       | +        |
| SCF           | •                               | X       | +        |
| SORT-1        | •                               |         |          |

Legend: HTN, hypertension;

X biomarkers retained in the bioinformatical model.

+ biomarkers retained in the biostatistical model.

• Green, biomarkers retained both in the biostatistical and bioinformatical models.

Note: For diabetes the bioinformatical results were inaccurate (>50% of misclassification), due to the low number of cases (n=36), and are not presented.

Supplemental Table 6. Descriptive statistics for the biomarkers selected for each condition (presented in the main Table 2)

| <b>Hypertension</b> |                |             |                |
|---------------------|----------------|-------------|----------------|
| <b>Biomarker</b>    | <b>Control</b> | <b>Case</b> | <b>p-value</b> |
| GDF-15, log pg/mL   | 6.5 ± 0.3      | 6.7 ± 0.4   | 0.015          |
| LEP, NPX            | 4.1 ± 1.1      | 4.6 ± 1.1   | 0.003          |
| SORT-1, NPX         | 6.8 ± 0.4      | 6.7 ± 0.5   | 0.024          |
| FABP-2, NPX         | 7.7 ± 0.7      | 7.5 ± 0.8   | 0.039          |
| <b>Obesity</b>      |                |             |                |
| LEP, NPX            | 4.3 ± 1.0      | 5.5 ± 0.9   | <0.001         |
| CEACAM-8, NPX       | 2.5 ± 0.7      | 2.7 ± 0.5   | 0.049          |
| PRELP, NPX          | 5.3 ± 0.3      | 5.4 ± 0.3   | 0.16           |
| <b>Diabetes</b>     |                |             |                |
| GDF-15, log pg/mL   | 6.7 ± 0.5      | 7.1 ± 0.7   | 0.002          |
| REN, NPX            | 6.3 ± 0.7      | 6.8 ± 0.7   | 0.005          |
| CXCL-1, NPX         | 7.4 ± 1.0      | 6.9 ± 0.8   | 0.021          |
| SCF, NPX            | 8.1 ± 0.4      | 7.7 ± 0.7   | 0.002          |

See the Supplemental Addenda 1 for the description of the NPX values.



## Supplemental addenda 1. Material and Methods for the Olink® technology

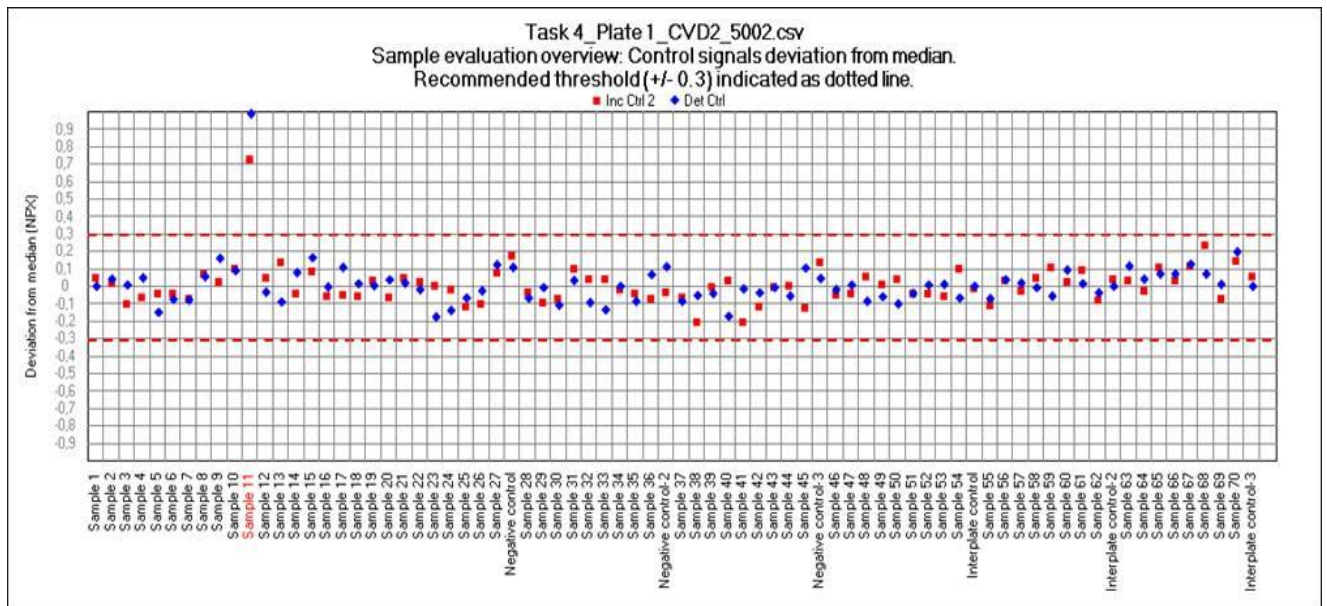
Proteins were measured using the Olink® CVDII, CVDIII, and Inflammation panels\* (Olink Proteomics AB, Uppsala, Sweden) according to the manufacturer's instructions. The Proximity Extension Assay (PEA) technology used for the Olink protocol has been well described (Assarsson et al, 2014; accessible at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0095192>), and enables 92 analytes to be analyzed simultaneously, using 1 µL of each sample. In brief, pairs of oligonucleotide-labeled antibody probes bind to their targeted protein, and if the two probes are brought in close proximity the oligonucleotides will hybridize in a pair-wise manner. The addition of a DNA polymerase leads to a proximity-dependent DNA polymerization event, generating a unique PCR target sequence. The resulting DNA sequence is subsequently detected and quantified using a microfluidic real-time PCR instrument (Biomark HD, Fluidigm). The resulting Ct-data is then quality controlled and normalized using a set of internal and external controls. The final assay read-out is presented in Normalized Protein eXpression (NPX) values, which is an arbitrary unit on a log<sub>2</sub>-scale where a high value corresponds to a higher protein expression. The four internal controls are designed to mimic and monitor the different steps of the PEA. They consist of two non-human proteins with matching antibody-probes that functions as incubation/immuno controls, an IgG antibody with two matching probes attached to it that functions as an extension control and a complete double-stranded amplicon that function as detection control. The internal controls are introduced to all samples as well as to the external controls. The external controls consist of a triplicate of negative controls used to calculate the limit of detection (LOD), as well as a triplicate of interplate controls (IPCs) containing 92 sets of antibodies with both matching probes for each assay attached to them that are used for normalization. For each sample and assay NPX is calculated by the following equations: 1.  $Ct_{(analyte)} - Ct_{(extension\ control)} = dCt_{(analyte)}$  (to decrease technical variation) 2.  $dCt_{(analyte)} - Ct_{(median\ IPC)} = ddCt_{(analyte)}$  (to improve inter plate variation) 3.  $Correction\ factor_{(analyte)} - ddCt_{(analyte)} = NPX_{(analyte)}$  (for more intuitive data). The correction factor is a set variable unique for each assay and reagent lot.

Quality control of the data is performed in two steps: First, the standard deviation for each of the incubation/immuno controls and the detection control is calculated for each run. A run will only pass quality control if the standard deviation for each of the controls is below 0.2. Secondly, each sample is quality controlled using incubation control 2 and the detection control. The run median of each of the controls is calculated and all samples within the run is compared to that. Samples that fall more than +/- 0.3 NPX from the plate median with regards to these two controls will fail the quality control and receive a QC warning in the data output file.

All assay validation data (detection limits, intra- and inter-assay precision data, etc.) are available on manufacturer's website ([www.olink.com](http://www.olink.com)).

\* Previously branded as Olink® Multiplex panels.

An example of a QC-plot is provided below:



## Technical description Olink Proximity Extension Assay

### Overview

Olink Proteomics has developed a unique protein detection system called Proximity Extension Assay (PEA), which enables highly specific and sensitive multiplex immunoassays for the detection of 92 different proteins in 90 samples simultaneously, generating 9260 data points in a single run. These proteins can be measured in a large variety of human sample types using as little as one microliter volume of sample. To date, >25 different biological matrices have been tested, all of which have proven compatible with PEA. Olink Proteomics currently offers assays for more than 977 proteins. During PEA, protein targets are ID-tagged by unique DNA amplicons that are subsequently detected and quantified by real-time PCR using Fluidigm's BioMark HD system. In doing so, the platform harnesses the highly parallel nature of Fluidigm's Dynamic Array IFC (Integrated Fluidic Circuit) technology, bringing the assays to the high throughput, reproducibility and superior sensitivity realm of the BioMark HD.

Measurements performed with the Proseek<sup>®</sup> Multiplex assays are semi-quantitative, where changes in protein levels in one group or population are quantified relative to another. The assays generate quantitation cycle (Cq) values that are then normalized and converted to Olink's arbitrary Normalized Protein eXpression (NPX) unit, which can be used for further statistical analyses. NPX gives relative quantification. NPX is on a log<sub>2</sub> scale. NPX should be compared between samples for each assay separately.

The analytical performance is thoroughly validated during product development to ensure the highest standards of sensitivity, precision, specificity and dynamic range. Since traceable calibrators are not yet available for all proteins in Olink's portfolio, performance is evaluated by using full length recombinant proteins as well as well characterized normal and diseased plasma/serum samples.

Repeatability of the measurements is reflected by an average intra-assay variation of 5-10% CV, and inter-assay variation 10-20% CV.

Multiple head-to-head comparisons have been performed against golden-standard single-plex ELISA assays in clinical settings (e.g. NT-pro-BNP, OPN, ST2). So far, these tests have demonstrated strong positive correlation, and have indicated similar, and sometimes superior, levels of sensitivity for Olink's PEA-based assays.

The real strength of Olink's technology is that it offers a high-throughput capacity coupled to a growing list of analytes, therefore placing the platform in a strong position for biomarker selection. Furthermore, the development of new multiplex assays is very rapid thanks to a conjugation procedure that works well on all antibody subtypes tested so far, universal incubation conditions and minimal probe concentrations used in our assays. Olink Proteomics currently offers 12 pre-designed 92-plex panels.

### **Detailed technical description**

#### Measurements

##### 1. Accuracy and reproducibility of the measurements:

Olink's multiplex assays are validated following the standard procedure for immunoassays.

Proseek® Multiplex measurements are semi-quantitative, where changes in protein levels are determined and compared for each protein separately. Due to the nature of relative quantification, no calibration standards are needed leaving room for more biological samples on the plate and, thus, higher throughput.

The output from the Proseek Multiplex protocol is in quantitation cycles (Cq) produced by the BioMark's Real-Time PCR Software. To minimize variation within and between runs, the data is normalized using both an internal control (Extension Control) and an Inter-Plate Control (IPC), and data is transformed using a pre-determined correction factor. The pre-processed data will be in the arbitrary unit NPX (Normalized Protein eXpression), which can be used for further statistical analyses.

##### 2. Volume:

As little as 1  $\mu$ l of sample is required to analyse 92 proteins in a single measurement. The amount of data generated increases proportionally with the volume of analysed sample, so that 2  $\mu$ l produces data for 184 proteins and so forth.

##### 3. Number of proteins:

92 proteins are quantified for each sample in a single run.

##### 4. Sample types:

Proseek® Multiplex assays are designed for detection of proteins in complex human biofluids.

Proseek® Multiplex assays are validated using both EDTA plasma and serum samples. Furthermore, a wide range of additional sample types are compatible with the technology, such as citrate and heparin plasma, CSF, urine, saliva, tissue and cell lysates, cell culture media, dried blood spots and

many more. See Table 1 at the end of this document for information about some of the sample types that have been analyzed so far. Olink is open to discussion and exploration of untested sample types.

#### 5. Correlation and agreement:

Correlation studies have been carried out for a subset of proteins to compare Olink's Proseek® Multiplex assays to commercial singleplex ELISA/conventional analyses in clinical settings. Olink assays have so far displayed high correlations with these established tests, both in Olink's own Service Laboratory and external, customer-led evaluations, demonstrating the accuracy and parallelism of the PEA technology. PEA is described by a peer-reviewed publication in PLOS one: Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability. Assarsson and Lundberg *et al.*, 2014.

#### 6. Validation:

During the validation procedure, Olink analyzes matched plasma (three different anti-coagulants) and serum samples, and reports their respective protein levels relative to the measurements in EDTA plasma. This varies for each biomarker and is presented in the Validation Data package for each panel. Intra- and inter-assays precision values are also determined for each protein and presented in the Validation Data Packages which can be found on <http://www.olink.com/proseek-multiplex/>.

In a recent pilot wellness study, EDTA plasma was taken from 88 well-characterized individuals at three different time points (every 3 months) and each sample was analyzed using multiple Proseek® Multiplex panels. A large proportion of biomarkers correlated significantly well across the different time points for each individual, while some proteins revealed distinct expression patterns that varied with time. It is also worth mentioning that the inflammation panel was able to identify the signs of infection and single out an individual who had been suffering from illness during one of the three time points (data not shown).

These data further strengthen the case for the robustness of the technology and its potential for future longitudinal studies.

#### 7. Information:

The platform was recently used to investigate the effects of freezer storage, chronological age at sampling, season and month of the year on the abundance levels of 108 proteins in 380 plasma samples collected from 106 Swedish women: "Effects of Long-Term Storage Time and Original Sampling Month on Biobank Plasma Protein Concentrations" Enroth *et al.*, 2016.

The study found that storage time can explain 4.8-34.9% of the observed variance, and that the chronological age at sample collection after adjustment for storage-time explains 1.1-33.5% of the observed variance.

Another recent study performed at the Karolinska Institute, using samples from the BBMRI BioBank in Stockholm, found that proteins are stable for consistent measurements by PEA, even after 8 short, successive freeze-thaw cycles (manuscript in preparation).

The first Olink panel was launched on the 1<sup>st</sup> of March 2013 and since then the technology has been used to analyse over 235 000 samples. This has so far led to 57 publications, in the form of technology reviews and customer applications, demonstrating that Olink's semi-quantitative approach for measuring biomarker levels is approved by referees. More publications are in the pipeline as Olink continues to analyse protein biomarkers for life science researchers around the world. For a complete list of publications see; <http://www.olink.com/data-you-can-trust/publications/>

Olink will work closely with customers to produce smaller and focused custom panels, to support the future development of diagnostic tools based on clinically relevant protein signatures.

Finally, with the help of bio-informatics tools, Olink proteomics is currently developing ways of extending the biological information provided about the proteins in each panel. In doing so, Olink intends to go beyond the simple provision of technology and reagents by simultaneously catering knowledge for the field of protein biomarker discovery. olink Proteomics provides first-class support to our customers. With our unparalleled experience and skills working with Proseek® Multiplex panels, the support team quickly delivers the best assistance, free of charge. Olink Proteomics also offers professional statistical analysis services. Our in-house team of biostatisticians can help you with customized statistical analyses tailored to your requirements.

All assay validation data (detection limits, intra- and inter-assay precision data, etc) are available on manufacturer's website ([www.olink.com](http://www.olink.com)).

## **Instrumentation**

### 1. Automation:

The Fluidigm BioMark HD combined with a controller or Juno is able to load all reagents into the microfluidic circuits and can be left to operate protocols unattended during the run. The Dynamic Arrays have an integrated network of channels, chambers and valves that automatically combine the reactions, conserving time, reagents and sample volumes. Initial pipetting is needed to pipette samples and reagents into the micro-well section of the dynamic arrays. Automatic sample and assay mixing of Fluidigm integrated fluidic circuit (IFCs) reduces pipetting by 95% compared to conventional PCR microplate-based systems. The IFC is subsequently placed in the controller (either HX or Juno). The samples do not come in contact with the interface plate which is within the controllers. However, it is possible to remove the interface plate and clean with 70% Ethanol as an additional measure between runs. Furthermore, it is possible to automate this using liquid handling instruments as the dynamic arrays have a Society for Biomolecular Sciences (SBS)-compatible format enabling manual and/or 384-well plate format handling and standard dispensing equipment.

By using Fluidigm 96.96 IFCs it is possible to analyse 90 samples + 6 controls in a single run on the BioMark HD. At a rate of 3 to 4 IFCs per day, the platform is largely capable of supporting the proposed scale of proteomic studies. Dedicated staff can achieve up to 360 samples per day (4 IFCs), which translates into  $4 \times 9216 = 36864$  data points in one day per BioMark HD system.

## 2. Accuracy and Reproducibility:

The repeatability and reproducibility of Olink multiplex assays is determined by using spiked and unspiked serum and plasma samples. These reference samples are generated to cover a wide concentration range for each protein target in order to mimic large biological variation, and are analysed by several (external) operators/service providers. Olink assays display an average intra-assay variation of 5-10% and a 10-20% inter-assay variation. The variation between runs is naturally slightly higher, with the highest variation observed for assays that measure samples close to the Limit of Detection (LOD).

### Sample types:

Table 1. Assay detectability in different sample types: Number of proteins detected in healthy blood donors

| Sample Type           | Panels                          |                                 |                                     |                                 |                                     |
|-----------------------|---------------------------------|---------------------------------|-------------------------------------|---------------------------------|-------------------------------------|
|                       | CVD II                          | CVD III                         | ONC II                              | INF I                           | NEU I                               |
| <b>Whole blood</b>    | NT                              | NT                              | NT                                  | NT                              | NT                                  |
| <b>Plasma EDTA</b>    | AS<br>99% in >75%<br>of samples | AS<br>98% in >75%<br>of samples | AS<br>100% in<br>>75% of<br>samples | AS<br>80% in >75%<br>of samples | AS<br>100% in<br>>75% of<br>samples |
| <b>Plasma Heparin</b> | AS                              | AS                              | AS                                  | AS                              | AS                                  |
| <b>Plasma Citrate</b> | AS                              | AS                              | AS                                  | AS                              | AS                                  |
| <b>Serum</b>          | AS<br>97% in >75%<br>of samples | AS<br>99% in >75%<br>of samples | AS<br>100% in<br>>75% of<br>samples | AS<br>82% in >75%<br>of samples | AS<br>98% in<br>>75% of<br>samples  |
| <b>Urine</b>          | NT                              | NT                              | NT                                  | AS<br>37% in >75%<br>of samples | NT                                  |
| <b>PBMC</b>           | NT                              | NT                              | NT                                  | AS<br>73% in >75%<br>of samples | NT                                  |

Additional sample-types include, but are not limited to:

|                                  |  |  |   |   |                                    |
|----------------------------------|--|--|---|---|------------------------------------|
| <b>CSF</b>                       | AS<br>70% in >75%<br>samples                       | AS<br>61% in >75%<br>samples                       | AS<br>76% in >75%<br>samples                  | AS<br>62% in >75%<br>samples                  | AS<br>85% in<br>>75%<br>samples    |
| <b>Exosomes</b>                  | AS<br>48% in >75%<br>samples<br>(plasma<br>origin) | AS<br>38% in >75%<br>samples<br>(plasma<br>origin) | AS<br>37% in >75%<br>samples<br>(cell origin) | AS<br>27% in >75%<br>samples<br>(cell origin) | NT                                 |
| <b>Saliva</b>                    | NT   | NT   | NT  | AS<br>63% in >75%<br>of samples               | NT                                 |
| <b>Synovial fluid</b>            | AS<br>46% in >75%<br>of samples                    | NT   | NT  | AS<br>41% in >75%<br>of samples               | NT                                 |
| <b>Cell lysates</b>              | AS<br>33% in >75%<br>of samples                    | AS<br>34% in >75%<br>of samples                    | AS<br>41% in >75%<br>of samples               | AS<br>38% in >75%<br>of samples               | NT                                 |
| <b>Tissue lysates</b>            | NT   | NT   | NT  | AS<br>79% in<br>>75%<br>samples               | AS<br>86% in<br>>75%<br>samples    |
| <b>Cell culture supernatants</b> | NT   | NT   | AS<br>34% in >75%<br>of samples               | AS<br>47% in >75%<br>samples                  | NT                                 |
| <b>Dried blood spots</b>         | NT   | NT   | NT  | AS<br>73% in >75%<br>of samples               | AS<br>54% in<br>>75% of<br>samples |

AS - Run in Analysis Service

NT - Not Tested

## References

- Assarsson E, Lundberg M, Holmquist G, Björkesten J, Thorsen SB, Ekman D, Eriksson A, Rennel Dickens E, Ohlsson S, Edfeldt G, Andersson A-C, Lindstedt P, Stenvang J, Gullberg M, and Fredriksson S. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLOS one* (2014).
- Enroth S, Johansson Å, Bosdotter Enroth S and Ulf Gyllensten U. Strong effects of genetic and lifestyle factors on biomarker variation and use of personalized cutoffs. *Nature Communications*. (2014).
- Chen H, Zucknick M, Werner S, Knebel P, Brenner H. Head-to-Head Comparison and Evaluation of 92 Plasma Protein Biomarkers for Early Detection of Colorectal Cancer in a True Screening Setting. *Clin Cancer Res*. (2015).
- Andreasson U, Blennow K and Zetterberg H. Update on ultrasensitive technologies to facilitate research on blood biomarkers for central nervous system disorders. *Alzheimer's & Dementia* (2016).

## NPX, QC and CV

### Calculation of NPX

Normalization of the Ct values derived from the qPCR detection step improves precision and makes data more intuitive (higher value = higher protein concentration). NPX (Normalized Protein eXpression) represents the relative signal (on log<sub>2</sub> scale) and is used for downstream statistical analyses.

1. Extension Control:

$$Ct_{\text{analyte}} - Ct_{\text{Ext Ctrl}} = dCt_{\text{analyte}}$$

2. Interplate Control:

$$dCt_{\text{analyte}} - dCt_{\text{Interplate Ctrl}} = ddCt$$

3. Normalization against a correction factor:

$$\text{Correction factor} - ddCt = \text{NPX}$$

- Protein expression changes are measured for each protein separately by relative quantification
- Even if two different proteins have the same NPX values, their actual concentration may differ
- NPX should be compared between samples within a run for each assay separately
- NPX cannot be directly compared between runs

## Quality control

Olink has built-in quality controls in all multiplex panels. Each 92-plex panel contains 96 assays. Four of those are internal controls that allow for an in-process quality control designed to monitor different steps of the protocol: immuno reaction, extension and amplification/ detection.



- The two incubation controls consist of two different non-human antigens measured with PEA assays: Incubation Control 1 and 2. These controls monitor potential variation in all three steps of the reaction.
- The Extension Control is an antibody coupled to both DNA-tags (hence always in proximity). This control monitors the extension and amplification/detection step.
- The Detection Control is a complete double stranded DNA amplicon which does not require any proximity binding or extension step. This control monitors the amplification/detection step.

The internal controls are used for both sample and run QC. The quality control of data is performed separately for each sample plate.

### **Sample QC**

Each of the internal controls are spiked into the samples in the same concentration. The signal for these are therefore expected to be the same over the plate. Sample QC is performed using the Detection Control and Incubation Control 2. Within each run, the levels of these controls are monitored for each sample and compared against the median of all samples. If either of the controls deviate more than the acceptance criteria allow (see below), the sample is flagged.

The Extension Control is used in the normalization step and in generation of NPX, and hence not included in the quality control of data.

The IPC cannot flag on the Detection control. This function was removed already from the start, as the large number of extensions that is required for the IPC made the IPC to always be a bit behind the rest of the samples. I.e., the IPC would always flag on the Detection control. It is however always constant, and flagging on the Detection control was therefore excluded.

Acceptance criteria for passing a sample:

- Incubation Control 2 and Detection Control deviates  $< +/- 0.3$  NPX from plate median.

### **Run QC**

The internal controls are used also in the run QC. This QC assesses the variation over the plate for each of the Incubation Controls 1 and 2 and the Detection Control. If the variation for one of the controls is too large (see below) the entire run is considered unreliable.

Acceptance criteria for passing a run:

- Standard deviation of Incubation Control 1 & 2 and Detection Control:  $< 0.2$  NPX.
- Number of flagged samples:  $\leq 1/6$  of total number of samples (i.e.  $\leq 16$  in a full plate).

### **Correction factor**

The correction factor is calculated during the validation of the panels. It is a pre-determined value, a fixed background, used to shift signals to a more intuitive value. Several runs during the validation process is used for this. Data from all assays are used but only IPC and negative control are needed from the samples. A minimum of 12 runs is performed. All runs should have at least three IPCs and three negative controls. Median is used in the final steps of the calculations to make the analysis less sensitive for extreme values.

Calculation of correction factor

1. Normalize using extension control:

$$Ct_{\text{sample assay (x)}} - Ct_{\text{sample assay (ext ctrl)}} = dCt_{\text{sample assay (x)}}$$

2. Normalize using median IPC (per experiment):

$$dCt_{\text{sample assay (x)}} - \text{Median } dCt_{\text{IPC assay (x)}} = ddCt_{\text{sample assay (x)}}$$

3. Calculate average value for background (per experiment):

$$\text{Average } (ddCt_{\text{background assay(x)}}) = Bg \text{ IPC}_{(x)}$$

4. Calculate median value for the background of all assays:

$$\text{Median } (Bg \text{ IPC}_{(x)})$$

Since the background is shifting within an interval, this interval is usually estimated by assuming normal distribution and selecting which interval to use. Olink uses a confidence interval of 99% (+/- 3xSD from the background). As the customer is recommended to run only three background samples in every run, this will not be enough to exactly calculate the standard deviation for the background. Therefore, we have a fixed value for this, which is selected from the same data as the Correction Factor. In contrast to the correction factor, fixed SD has a function and impact on the data.

Normalization

Olink NPX Manager has three available methods for data normalization and minimization of systematic biases between sample plates. These methods are described below. An important concept when deciding on normalization procedure is randomization, which in this context applies to the sample placement across the plates. For randomized studies with more than one plate, intensity normalization is the default normalization. For other studies, IPC normalization is default.

### **IPC normalization**

Three inter-plate controls (IPC) are included on each plate and run as normal samples. The inter-plate control is a pool of 92 antibodies, each with one of the pairs of unique DNA tags on it positioned in fixed proximity (i.e. 92 Extension Controls).

The median of the three IPCs is used as normalizer for each assay, and this compensates for potential variation between runs/plates. This method is completely independent on the samples included on the plate and is therefore recommended for projects where complete randomization of samples cannot be guaranteed.

### **Intensity normalization**

The intensity normalization adjusts the data so that the median NPX for a protein on each plate is equal to the overall median. Each plate is adjusted so that the median of all assays is the same on all plates.

This method requires that the true median of each plate is the same. One way of ensuring this is to randomize the samples beforehand. If there is total randomization, this method outperforms other normalization methods. If there are specific types of samples that are only available on certain plates, this normalization method should not be used.

### **Control normalization**

The control normalization works in the same way as the IPC normalization, but with one addition. While the IPC adjusts each protein to the level of a reference, the control normalization also adjusts the total intensity of entire plates. This is useful if you have plates or projects that differ on average in a way that cannot be explained biologically. Control normalization can be used to reduce systematic variation in non-randomized studies. This normalization method requires the addition of at least two control samples on each plate.

NOTE This is an optional normalization method that should only be used when standard normalizations are not sufficient.

### **Coefficient of variation**

%CV is calculated using linear NPX values from replicate control samples, for assays detected above LOD.

Intra-CV represents the CV within a plate and inter-CV represents the CV between plates. The reported %CV is the mean %CV over all assays.

Coefficient of Variation = (Standard Deviation / Mean) \* 100

- Reference value for Inter %CV: < 25%
- Reference value for Intra %CV: < 15%