

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Equipment software: Mass Spectrometry - Xcalibur 4.0

MS spectra conversion software: ProteoWizard 3.0.9935

Data analysis

General software: Microsoft Excel 16.0.4266.1001

Adobe Illustrator 23.0.1

LC-MS data analysis software: MAVEN 6.2

Computational software: MATLAB 2017b

Link to the code in an open source repository: <https://github.com/shovall/FlowInjectionMSOptimization>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The metabolomics data generated and analyzed in this study is available in Metabolomics Workbench with the identifier doi:10.21228/M8P41V

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Samples sizes were sufficient according to statistical methods, confidence intervals and p-values.
Data exclusions	No data were excluded from the analysis.
Replication	All replication attempts of the experiments were successful.
Randomization	Set of human serum samples was randomized during metabolomics analysis for gender prediction.
Blinding	Not relevant

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involvement
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involvement
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HCT116 - ATCC, CCL-247 HeLa - ATCC, CCL-2 Hek 293 - ATCC, CRL-1573 HepG2 - ATCC, HB-8065 MiaPaca2 - ATCC, CRL-1420 Panc-1 - ATCC, CRL-1469 A549 - ATCC, CCL-185 WM266-4 - ATCC, CRL-1676 Jurkat - ATCC, CRL-2899 CCRF-CEM - ATCC, CCL-119 Purchased from ATCC, USA
Authentication	Non of the cell lines used were authenticated.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified cell lines were used.

## Human research participants

---

Policy information about [studies involving human research participants](#)

Population characteristics	98 serum samples of healthy individuals (for gender information used for this study see Methods).
Recruitment	Samples were purchased from the Israeli MIDGAM Biobank.
Ethics oversight	Rambam Hospital, Haifa, Israel (IRB 0481-18-RMB).

Note that full information on the approval of the study protocol must also be provided in the manuscript.