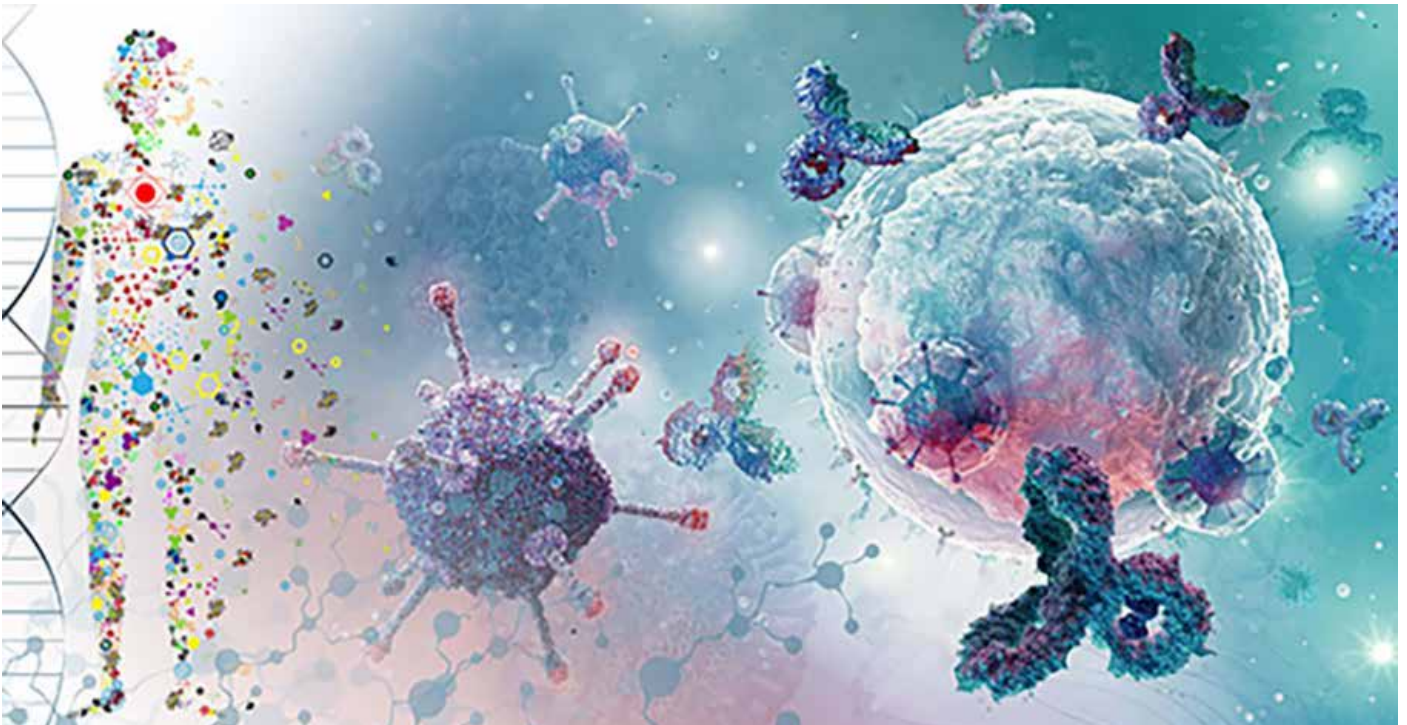


Yksilöity lääketiede syövän ja virusten torjunnassa

Solujen mallintaminen ja niiden toiminnan simuloiminen parantaa huomattavasti henkilökohtaisia hoitosuunnitelmia. PerMedCoE- hankkeessa yhdistetään kliinistä potilastietoa geenien ja proteiinien ja solujen toimintaan liittyvään tietoon. Tavoitteena on kehittää täsmälääketieteessä käytettäviä työkaluja. Solujen mallintaminen yksityiskohtaisesti on kuitenkin valtava urakka ja vaatii paljon supertietokoneiden laskentatehoa.



Yksilöity lääketiede avaa tulevaisuudessa suuria mahdollisuuksia. Tavoitteena on, että potilaan kliininen data voidaan yhdistää geneettiseen dataan ja näiden tietojen pohjalta voidaan laatia yksilöllisiä hoitosuunnitelmia. PerMedCoE-hankkeessa (HPC/Exascale Centre of Excellence in Personalised Medicine) pyritään parantamaan yksilöidyn lääketieteen mallinnusohjelmistojen yhteensopivuutta eksaskaalan supertietokonejärjestelmiin. Eksaskaalan supertietokoneet ovat seuraavan sukupolven järjestelmiä joiden teoreettinen laskentateho vastaa jopa 10^{18} laskutoimitusta sekunnissa. Hankkeeseen osallistuu tutkijoita useasta eurooppalaisesta yliopistosta ja sairaalasta. Projekti keskittyy neljään avoimen lähdekoodiin perustuvaan solutason mallinnusohjelmistoon. Ohjelmis-

tokehityksen lisäksi tutkimushankkeessa pyritään edistämään täsmälääketieteen työkalujen helppokäyttöisyyttä ja toimivuutta useissa eurooppalaisissa suurteholaskentakeskuksissa.

”Tavoitteena on että nämä neljä ohjelmistoa pystyisivät tulevaisuudessa toimimaan useassa supertietokoneessa”, sanoo CSC:n projektipäällikkö **Sampo Sillanpää**.

”Tällä hetkellä tämä on teknisesti hyvin haastavaa toteuttaa, koska jokainen suurteholaskentaympäristö on omanlaisensa johtuen järjestelmäarkkitehtuurista.”

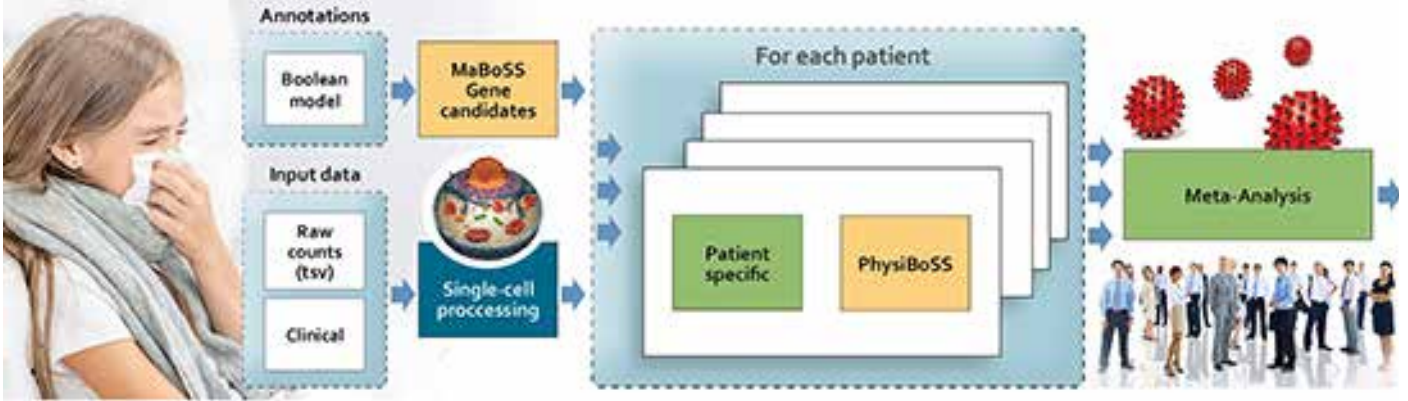
Ohjelmistojen ja datamassojen saumaton toiminta on tarkoitus saavuttaa yhteisesti sovituilla teknologioilla. PerMedCoE-hankkeessa tämä toteutetaan ns. konttitekniikan ja työvuoro-ohjelmistojen avulla. Työvuoro on

tutkimusprosessin automaatio, jonka aikana dokumentteja, tietoa ja tehtäviä siirretään suoritettaviksi tiettyjen sääntöjen mukaisesti. Konttitekniikan avulla voidaan määrittää vakioitu ympäristö, jossa tieteellisiä ohjelmistoja ajetaan jokaisessa hankkeeseen osallistuvassa suurteholaskentaympäristössä. Kun ohjelman koodi kirjastoinen ja asetuksineen asetetaan konttiin, sitä voidaan siirrellä koneelta ja konesalista toiseen.

”Ohjelmistot ja data on tavallaan paketoitu omaan laatikkoonsa, jotta niitä voidaan siirtää ympäristöstä toiseen. CSC:llä on useita konttitekniikan asiantuntijoita, joten työkaluja pystytään siirtämään alustalta toiselle”, Sillanpää sanoo.

”Kontteja hyödyntämällä asiantuntijat pystyvät rakentamaan käyttäjäystävällisiä

COVID19 use case (with scRNA-Seq)



COVID-19-tautia ja eri solupopulaatioiden käyttäytymistä voidaan tutkia moniskaalautuvien mallien ja yhden solun dataa käyttämällä. MaBoSS on ohjelmisto, joka mahdollistaa solupopulaatioiden simuloimisen ja mallintamisen stokastisesti (Boolean mallinnus) ja solunsisäisiä mekanismeja, joiden säätelyyn taudit vaikuttavat. PhysiBoSS yhdistää MaBoSS-ohjelmiston vapaan lähdekoodin PhysiCell-ohjelmistoon, joka simuloi isoja solujärjestelmiä. 3D-kudoksia voidaan tutkia standardeilla pöytäkoneilla. PhysiBoSS-ohjelmisto avulla tutkijat voivat analysoida yksittäisten solujen geneettisiä muutoksia väestötasolla.

työnkulkuja, eli työvoita. PerMedCoE-hankkeessa työvuot koostuvat useammasta rakennuspalikasta, joista jokainen toteuttaa tietyn täsmälääketieteen laskentatehtävän. Yhdessä rakennuspalikassa voidaan tehdä esikäsittelyä datalle, toisessa varsinainen analyysi ja viimeinen antaa tuloksen lopputulokseksi. Käyttäjän ei siis välttämättä tarvitse huolehtia miten useasta rakennuspalikasta rakennettu automatisointi toimii, vaan keskittyä tulosten tulkintaan.”

COVID-19 mallintaminen solutasolla

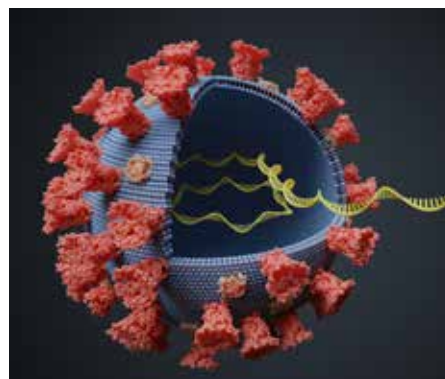
Hankkeessa rakennettujen teknologioiden hyödyllisyyttä arvioidaan erilaisten käyttötapauksien avulla. Työvoimien avulla analysoidaan, mitä häiriöitä taudit voivat aiheuttaa solutasolla tai miten lääkeaineet toimivat. Mallien avulla voidaan tutkia solujen aineenvaihduntaa tai signaalinvälitystä.

”PerMedCoE-käyttötapauksissa hyödynnetään julkisesti saatavilla olevia genomidata-aineistoja. Nyt voimme tutkia koronaviruspotilaista otettuja näytteitä ja etsiä genomidatista sellaisia markkereita, jotka ilmentävät, mitkä potilasryhmät ovat erityisen alttiita taudin vaaralliselle muodolle.”

Projektissa mallinnetaan ihon epiteelikudosta, joka reagoi koronavirusstartuntaan kutsumalla erilaisia immuuneja soluja vaikuttamaan virukseen. Näin voidaan

mahdollisesti paremmin tunnistaa sellaisia potilasryhmiä, jotka ovat alttiita koronan vakavalle tautimuodolle.

”Ajatuksena on, että pystytään rinnakkain ajamaan useita malleja yksittäisille potilaille. Näin voidaan tehokkaasti analysoida riittävän suuria datamääriä, jotta mallinnustuloksia voitaisiin käyttää yksilöidyn lääketieteen apuna”, sanoo vanhempi datatieteilijä **Jesse Harrison CSC:stä.**



COVID 19-käyttötapauksen mallinnuksessa käytetään solutason RNA-sekvenssi-dattaa. RNA-sekvensointi yhden solun tarkkuudella (scRNA-seq) voi paljastaa geenien välisiä säännöllisiä vuorovaikutusyhteyksiä, solujen syntyperälinjat, solujen eroavaisuuksia sekä solun viitekehysten ympäristössään.

Toinen tärkeä projektin käyttötapaus on syöpädiagnostiikka. Tavoitteena on luoda

mallinnustyökaluja syöpäkasvainten kasvun ennustamiseen ja potilaskohtaisten hoitojen kehittämiseen. Aineistona käytetään Wellcome-instituutin ja Massachusettsin syöpäkeskuksen keräämää aineistoa. Tietokantaan on kerätty yli tuhat erilaista kasvainkudoksen solulinjaa.

”Projektissa pyritään esimerkiksi tunnistamaan uusia lääkeyhdistelmiä, jotka voisivat olla syöpähoitossa hyödyllisiä” Jesse Harrison sanoo.

Tämä johtaisi toivottavasti potilaskohtaisten syöpähoitojen tarkempaan kohdistamiseen ja diagnostiikan nopeutumiseen.

”Jotta nämä tavoitteet täyttyisivät läheisempää yhteistyötä tarvitaan suurteholaskentakeskusten ja lääketieteellisen organisaatioiden kanssa. Tämä siksi, koska nyt puhutaan isoista datamassoista ja suurten potilaskohtaisen datan analysointi ei ole omalla pöytäkoneella mahdollista.”

PerMedCoE:n tulokset ja työkalut on tarkoitettu kaikille tutkijoille.

”Kun projekti päättyy kesällä 2023, meillä on päivitettyjä versioita avoimen lähdekoodin pohjalta kehitetyistä mallinnustyökaluista ja ne saatetaan tutkijayhteisön saataville. Hankkeessa luodaan myös uutta osaamista tukemaan täsmälääketieteen työkalujen käyttöä CSC:n laskentaympäristöissä.”

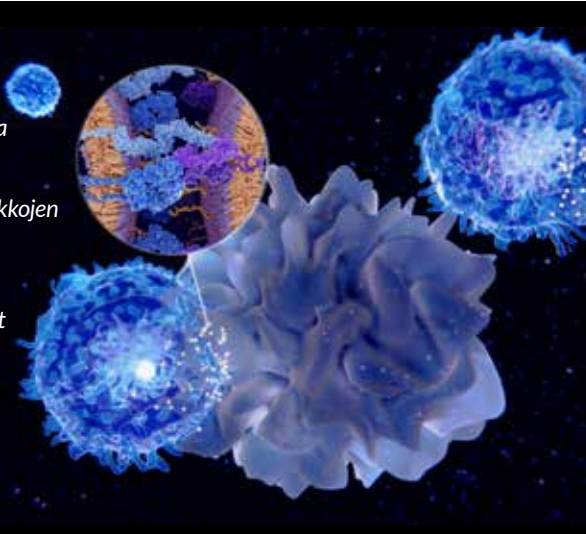
Ari Turunen

COBREXA.jl:
solun aineenvaihdunnan
simulointi genomien tasolla

CellNOpt:
solujen signaalinäilytysverkkojen
mallintaminen

MaBoSS:
Boolean mallien stokastiset
simulaatiot

PhysiCell:
Isojen solujärjestelmien
simulointiin



Miten solut toimivat eri tasoilla, yksittäisestä solusta laajoihin solupopulaatioihin? CSC – Tieteen tietotekniikan keskus ja Barcelonan superlaskentakeskus (BSC), yhdessä kymmenen muun akateemisen ja kaupallisen toimijan kanssa, aloittivat lokakuussa 2020 Euroopan komission huippuyksikköhankkeen HPC/Exascale Center of Excellence for Personalised Medicine (PerMedCoE). Hankkeessa kehitetään solutasoisen mallinnusohjelmistojen suurteholaskentaan soveltuviksi. Suurteholaskennan avulla biologinen data kuten genomiikka ja proteomiikka voidaan tuoda osaksi täsmälääketiedettä, koska datan analysointi nopeutuu huomattavasti. Esimerkiksi sairauden diagnooseja pitäisi käytännössä pystyä tekemään tuntien tai päivien sisällä. PerMedCoE on osa ELIXIR Suomen kehitysohjelman.

Syövän voittaminen

EU rahoittaa monia projekteja, jotka tulevaisuudessa mahdollistavat yksilölliset potilashoidot. Syöpä on yksi esimerkki taudista, joka on erittäin yksilöllinen, oli kyse sitten rintaa-, keuhko-, maksa-, tai eturauhassyövästä.

Esimerkiksi Horisontti Eurooppa -puiteohjelman *Conquering Cancer: Mission Possible* näyttää Suomen molekyyli lääketieteen instituutin (FIMM) tutkijan **Esa Pitkäsen** mukaan suuntaa tulevaisuuden syöpätutkimukselle ja -hoidoille. Kunnianhimoinen ohjelma tavoittelee syöpien syntyyn johtavien mekanismien ymmärtämistä, uusia menetelmiä syöpien aikaiseen havaitsemiseen, sekä henkilökohtaisen eli yksilöidyn syöpälääketieteen läpimurtoja.

”Kaikille näille tavoitteille yhteistä on monipuolisen ja laajan terveysdatan hyödyntäminen uusien laskennallisten menetelmien avulla. Koneoppimiseen perustuvien tekoäly-

algoritmien avulla onkin jo saavutettu rohkaisevia tuloksia esimerkiksi digitaalisen patologian alalla. Seuraavat harppaukset tehdään yhdistelemällä useita eri tietolähteitä yksilöllisten syöpäseulonta- ja hoitosuosituksen antamiseksi”, Pitkänen uskoo.

Ohjelmassa syöpäpotilaat halutaan mukaan syöpähoitojen kehitykseen esimerkiksi antamalla potilaille mahdollisuuksia lähettää tietoturvasuojalla omaa terveysdataansa tutkijoiden käyttöön. Samalla potilaat saavat myös uutta tutkimustietoa omasta sairaudestaan.

”On tärkeää, että hoitomuotojen kehittyessä pidetään huolta siitä, että ihmisille taataan tasa-arvoinen mahdollisuus hyötyä uusista hoidoista taustasta riippumatta. Olen iloinen siitä, että tämä on huomioitu ohjelman suosituksissa. Lisäksi lasten ja nuorten syöpiin kiinnitetään erityistä huomiota.”

LISÄTIETOJA:

HPC/Exascale Centre of Excellence in Personalised Medicine

<https://permedcoe.eu>

CSC – Tieteen tietotekniikan keskus Oy

on valtion omistama, opetus- ja kulttuuriministeriön hallinnoima, voittoa tavoittelematon osakeyhtiö. CSC ylläpitää ja kehittää valtion omistamaa keskitettyä tietotekniikkainfrastruktuuria.

<http://www.csc.fi>

<https://research.csc.fi/cloud-computing>

ELIXIR

rakentaa infrastruktuurin bioalan tutkimuksen tueksi. Se yhdistää 21 Euroopan maan ja Euroopan molekyylibiologian laboratorion EMBL:n johtavat organisaatiot yhteiseksi biologisen informaation infrastruktuuriksi. Sen Suomen keskus on CSC Tieteen tietotekniikan keskus Oy.

<http://www.elixir-finland.org>

<http://www.elixir-europe.org>

SUOMEN ELIXIR

Puh. +358 9 457 2821 e-mail: servicedesk@csc.fi
www.elixir-europe.org/about-us/who-we-are/nodes/finland

www.elixir-finland.org

ELIXIR PÄÄMAJA

EMBL-European Bioinformatics Institute
www.elixir-europe.org