Big Data and Artificial Intelligence in Medicine and Drug Discovery

Hiroshi Tanaka
Biomedical Data Science
Tokyo Medical and Dental University
and
Tohoku Medical Megabank Organization,
Tohoku University

Coming ! of the era of

Big Data Medicine

In Next Decade

Framework (paradigm) of Medicine

Will be Totally Changed!!

Big Data?

Difficult to treat by conventional information processing method because it is too large, too many kinds and too frequently changing

So what is

Medical Big Data?

Big Data in Medicine

Rapid and Huge Accumulation of Big Data

(1) **Precision Medicine** : Comprehensive **Genome-Omics data** brought by advance of biotechnology (e.g. NGS, Molecular Images)

(2) **Genomic Biobank**: **Genomic and Environmental (exposomic) data** of Genomic Cohort participants

(3) **mHealth**: **Continuous physiological and behavioral data** by **mobile Health** (wearable sensor monitoring)

Enormously **Cost Reduced**, nevertheless
High Quality Massive Data

Whole Genome seq : 13 yr, 3,500 M$ (2003) →
1day, 1000$ (2016)

How we should cope with this Medical Big Data

Tremendous Improvement of **Preciseness** of Medical Care

Groundbreaking Change of Medicine

New type of Big Data emerges

Medical **Big Data** Revolution

- Clinical Conventional “Large scaled Data”
  - Clinical Lab Tests, Prescriptions, Images
  - Ex. claim DB. Jp. Sentinel Project
- Socio-Medical epidemiological “Large scaled Data”
  - Ordinary epidemiological data
  - life style, health exams, questionnaire

Due to recent spread of “Digitalization”

Conventional
Medical
“Large data”

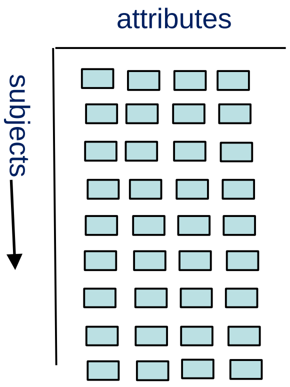
- **Big data of “Genome-Omics Medicine”**
  - Genome Omics Medicine
  - Due to Rapid Advance of **Clinical Sequencing**
  - **Molecular biomedical images**
- **Big Data of “Continuously monitoring biosignal”**
  - Life-course-oriented healthcare
  - Lifestyle, behavioral information, **mHealth**
  - Due to Rapid Advance of **Wearable Sensor**

New type of
(Genuine)
Medical **Big Data**

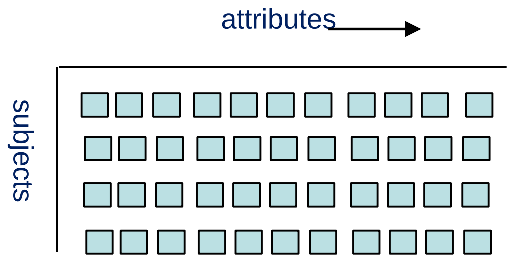
TMDU

New type of Medical Big Data

Data Structure



- Conventional Medical “Big Data”
  - **“$\mathcal{N}$– Big Data”**
    - For one subject (patient)
      Num. of attributes is “Small” (**n>>p**)
    - Num.(*n*) of subjects (patients) is “Big”
    - Conventional statistical method works well



- **New type of Big Data** (omics, mHealth)
  - **“$\mathcal{P}$- Big Data”**
    - Num. of attributes(*p*) for one subject is “Big”
    - “**New NP problem**” (**p>>n**)
    - But Num. of subject (patients) is comparatively “Small”
    - Conventional statistical method does not work well

Necessity of
New Data Science of Medicine

New type of Medical Big Data

Purpose to Collect Big Data

- Conventional Medical “Big Data”
  - **Population Medicine**
  - To reveal the **“collective law**” (“laws in group-level”) by collecting large number of samples
  - which can not be found by seeing each individual subject
- **New type pf Big Data** (**genome, omics, mHealth**)
  - **Personalized (Stratified) Medicine**
  - To comprehensively **enumerate all the individualized (stratified) patterns** existing under the same name of disease; **How many individualized patterns** exists?
  - For exhaustive and complete search, **Big amount of samples** are necessary.

Intention to Collect **Big Data** is Quite **Opposite**
Toward collective vs individualized pattern

Paradigm Changes

Medical Big Data Revolution Causes

- “**Population medicine**” **paradigm disrupts**
  - **“One size fit for all” medicine** is no more valid
  - **Towards “Individualized Medicine”**
    - How many “Personalized (Stratified) Patterns” (**intrinsic subtypes**) of **disease** exit under **the same name of disease**
    - How fine granularity of stratification should be?
    - **Big Data** is needed for **enumeration of these intrinsic subtypes**
- “**RCT and Evidence-based Medicine**” **paradigm disrupts**
  - Liberation from the “gold standard” of RCT and EBM
  - **RCT**: Random (Artificial) Controlled Trials with Small-ish populations **outside the Real Medical Practice**
  - These concepts are **before the discovery of “individualized medicine”** and are **no more valid**
  - **Randomization can not eliminate** the **difference of intrinsic subtypes** of disease unlike conventional confounding factors
  - **Towards Learning from “Real World Data”** (Disease registry, EHR big data) for clinical evaluation of drugs, devises, etc.

Big Data
in Genome-Omics Medicine

TMDU

Two Streams of Genome-Omics Medicine

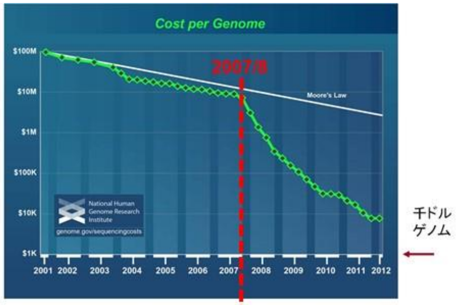
Genome Medicine in United States: **Precision Medicine**

- Surging Wave of **Rapid Clinical Implementation of Genomic medicine** (2010) shortly **after** “**Sequence Revolution** (2007)”
- Aiming at **dramatic improvement in therapeutic medicine** for **individual patient** by genome information
  - **POC** (Point of care) ID of **causative gene for rare disease**
  - **POC** (point of care) ID of **driver gene mutation for cancer**
  - **Preemptive PGx**: polymorphism of **drug metabolizing enzyme**

Genome Medicine in Europe: **Genomic Biobank**

- Recognition of the Value of “**Collective Genome Information**” (island) to the **Spread of Genomic Biobank** today
- Aiming at **dramatic improvement in preventive medicine** for the **general public** (a nation) by genome information: based on the concept of ”welfare state”
  - **Prospective Population-based Large Genomic Cohort**
  - Prediction of **Occurrence of “Multifactorial Disease**”
  - Estimate the **interaction of genomic predisposition and environmental factors**

Genome Medicine of United States



DNA Sequencing Cost: the National Human Genome Research Institute

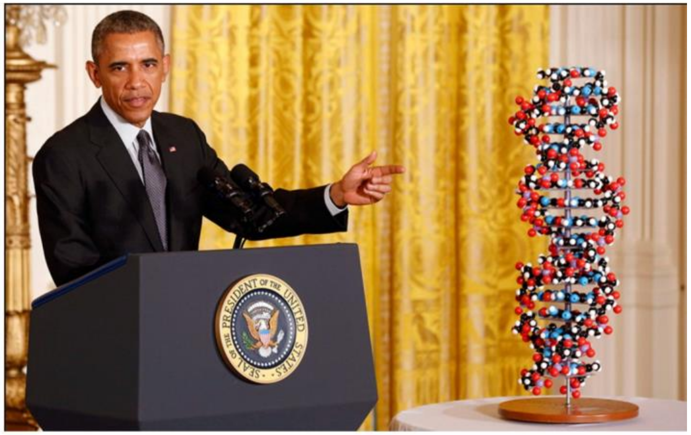
Sequence Revolution 2007/8

2005～ NGS 454 (LS,Roche)
2007/8～454, Solexa (Ilumina),
SOLiD (LT,TF)
Sequence Revolution
Faster than Moor's law

Ilumina 2500

Ion Torrent

President Obama Precision Medicine Initiative



2015.1 State of the Union Address

1st term
Early
adopters

2005～ NGS
(454,Solexa,SOLID)
2007/8～
Sequence Revolution

Identification of
Causative Gene for
Undiagnosed Disease
Exome Sq.
MCW Children H

Polymorphism genome
Hapmap project 2002～
GWAS 2003～

Polymorphism of Drug
Metabolizer, Preemptive PGx
Integrated with EHR
Vanderbilt UH

TCGA (2006), ICCG
(2008) ,Results
were Published
from 2011～

Identification of Cancer
Driver Gene, Molecular
Targeted Drug
Dana Faber CC

Genome Omics Medicine
Clinical Implementation

2nd term
National
project

National project
BD2K, many consortium, WG

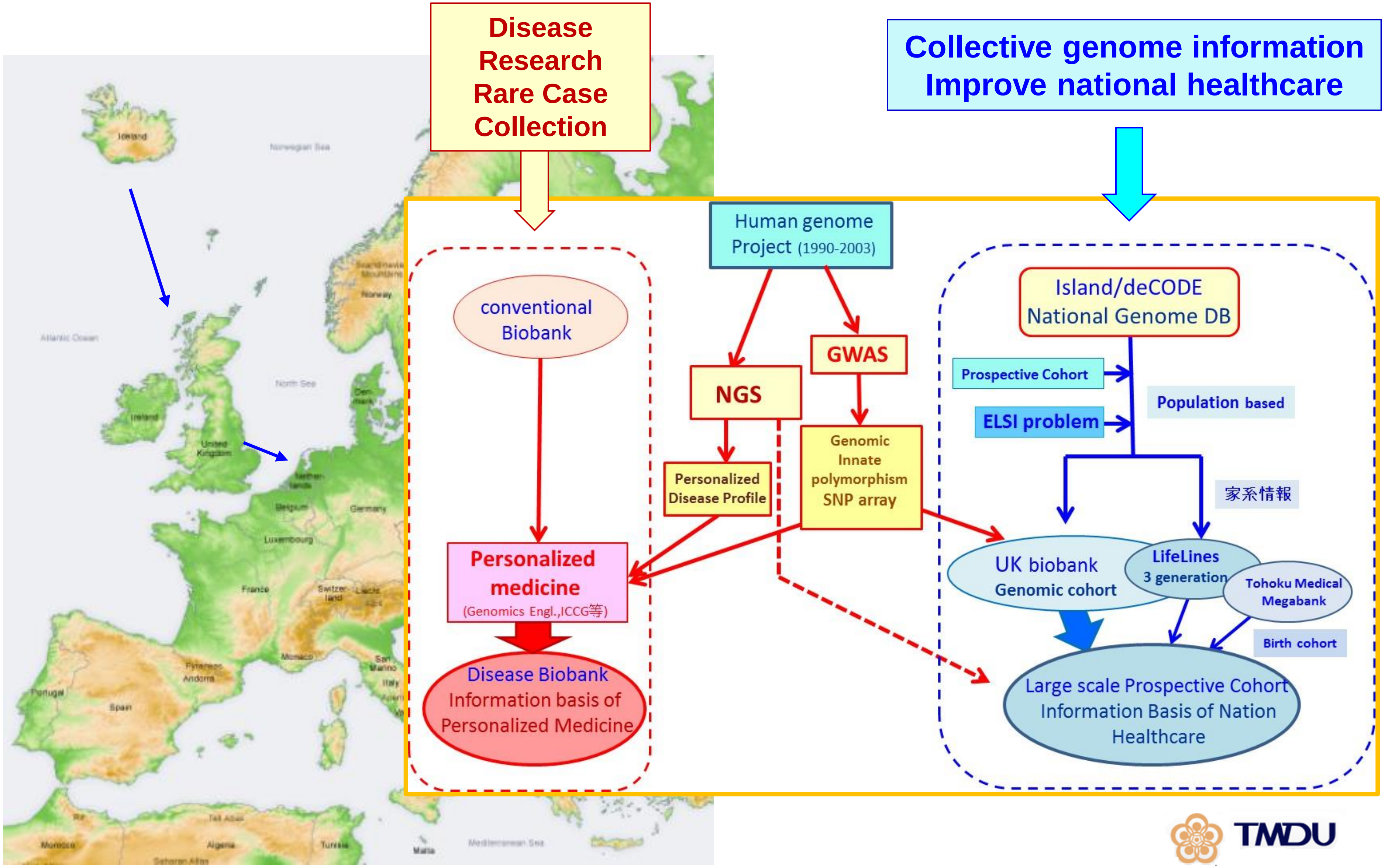
Precision Medicine Initiative
2016
President Obama, State of Union Address

3rd term
Spread of
precision
medicine

Prevail of Precision Medicine
1M cohort "All of Us"
National Cancer Moonshot Project

2007
2009
2010
2011
2012
2013
2014
2015
2016
2017

Genome Medicine of Europe



Disease Research Rare Case Collection

Collective genome information Improve national healthcare

Human genome Project (1990-2003)

NGS

GWAS

Personalized Disease Profile

Genomic Innate polymorphism SNP array

conventional Biobank

Personalized medicine (Genomics Engl.,ICCG等)

Disease Biobank Information basis of Personalized Medicine

Island/deCODE National Genome DB

Prospective Cohort

ELSI problem

Population based

家系情報

UK biobank Genomic cohort

LifeLines 3 generation

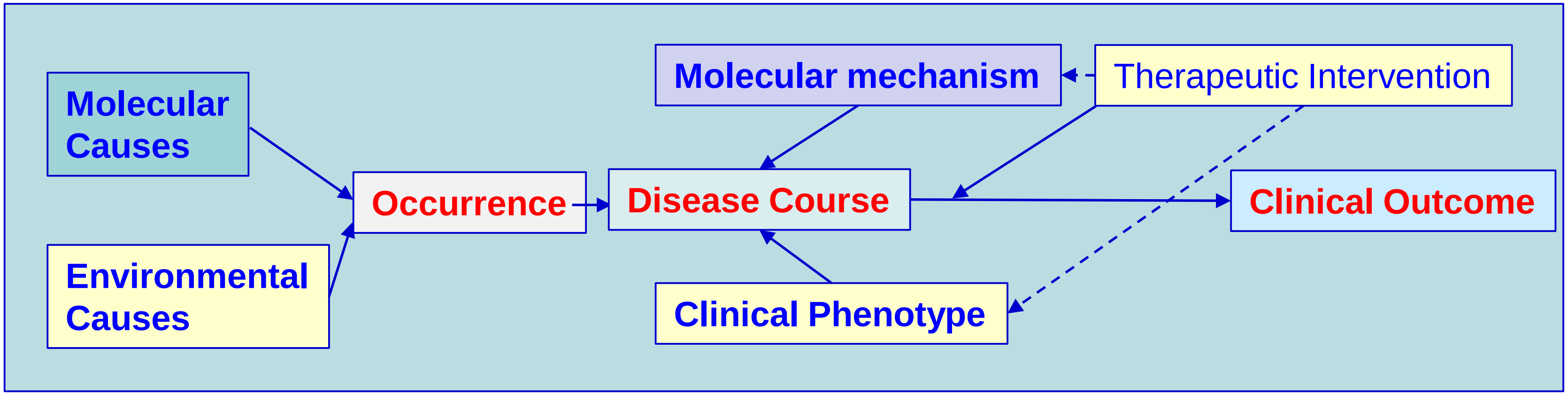
Tohoku Medical Megabank

Birth cohort

Large scale Prospective Cohort Information Basis of Nation Healthcare

Challenge of Big Data Medicine I

Disambiguation of corresponding “non-genomic” information



Ontology of disease course

Molecular Mechanism of Disease Occurrence and Progression

Disambiguation of corresponding **exposomic and phenotypic** information

Environmental factors **Clinical Phenotype** Therapeutic Intervention

eMERGE and PheKB

phase I (2007-2011)

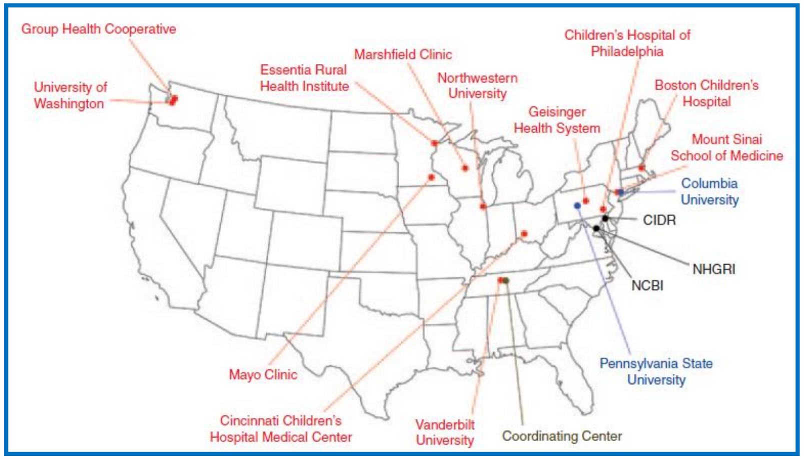
- **Phenotyping from EMR**
  - **genomic discovery and genomic medicine implementation** research.
- **EMR-based GWAS**
  - Each with its own biorepository (DNA etc) linked to phenotypic data contained within EMRs
- **eMERGE-I**: 5 Institutes, **PheKB**

phase II (2011-2015)

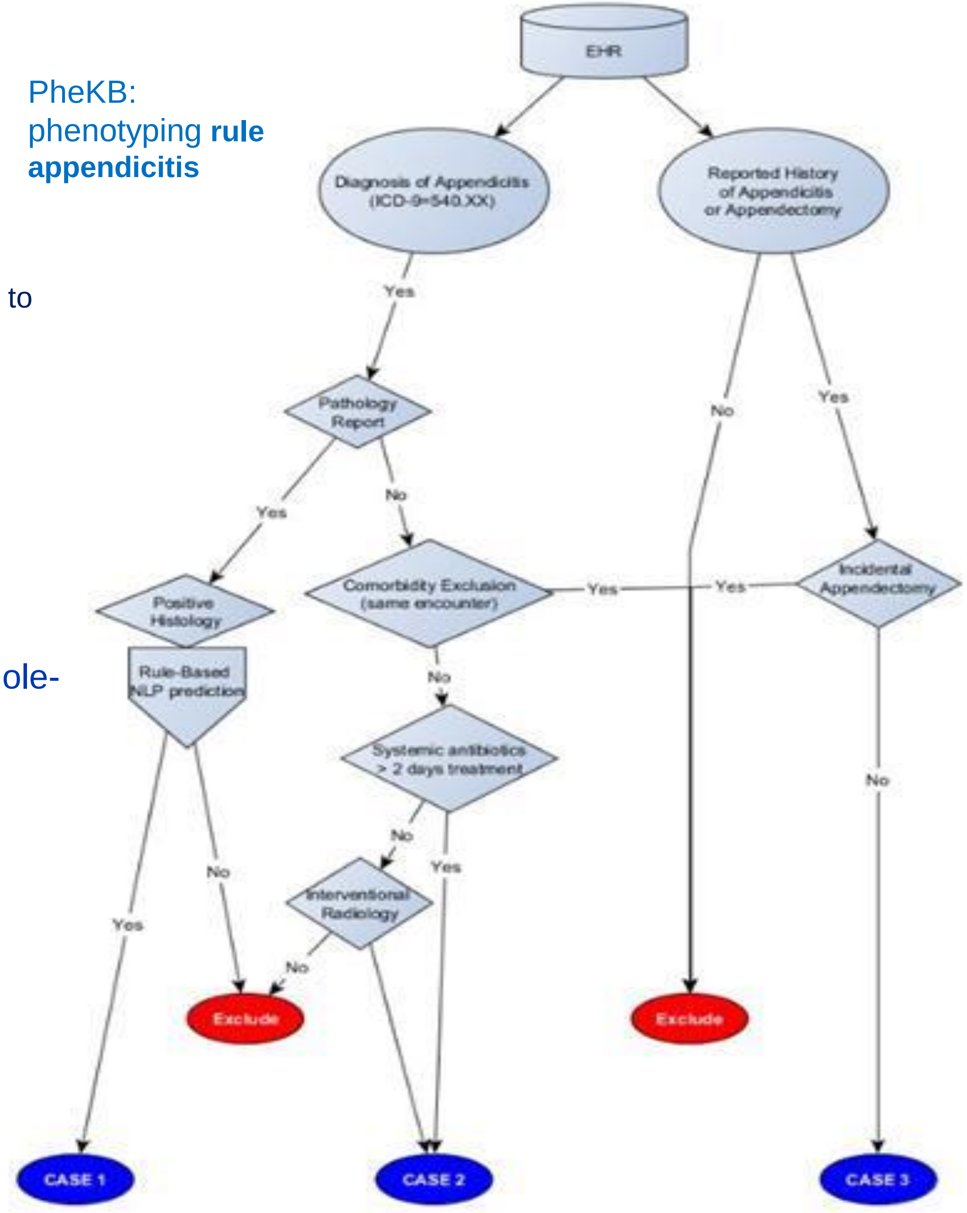
- **Integration of Genomic Information into EMR** (Clinical Implementation)
- PGx implementation in EMR
- Return of (Genomic) Result (RoR)

Phase III (2015～2019)

- explore the potential of whole-genome and whole-exome


Group Health Cooperative
University of Washington
Marshfield Clinic
Essentia Rural Health Institute
Northwestern University
Children's Hospital of Philadelphia
Boston Children's Hospital
Geisinger Health System
Mount Sinai School of Medicine
Columbia University
CIDR
NHGRI
NCBI
Pennsylvania State University
Mayo Clinic
Cincinnati Children's Hospital Medical Center
Vanderbilt University
Coordinating Center

PheKB: phenotyping **rule appendicitis**


EHR
Diagnosis of Appendicitis (ICD-9=540.XX)
Reported History of Appendicitis or Appendectomy
Yes
No
Yes
Pathology Report
Yes
No
Positive Histology
Comorbidity Exclusion (same encounter)
Yes
Yes
Incidental Appendectomy
Rule-Based NLP prediction
No
Systemic antibiotics > 2 days treatment
No
No
Yes
Yes
Interventional Radiology
No
No
Exclude
Exclude
CASE 1
CASE 2
CASE 3

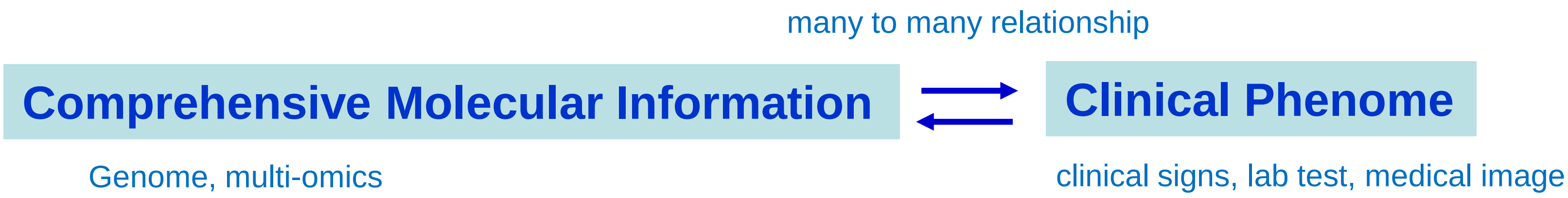
Challenge of Big Data Medicine II

Contraction Methodology to extract the Intrinsic Information Structure

Medical Big Data

⇨ **Hyper Multi-dimensional Correlation Network of Data**

Clinical Genome Medicine

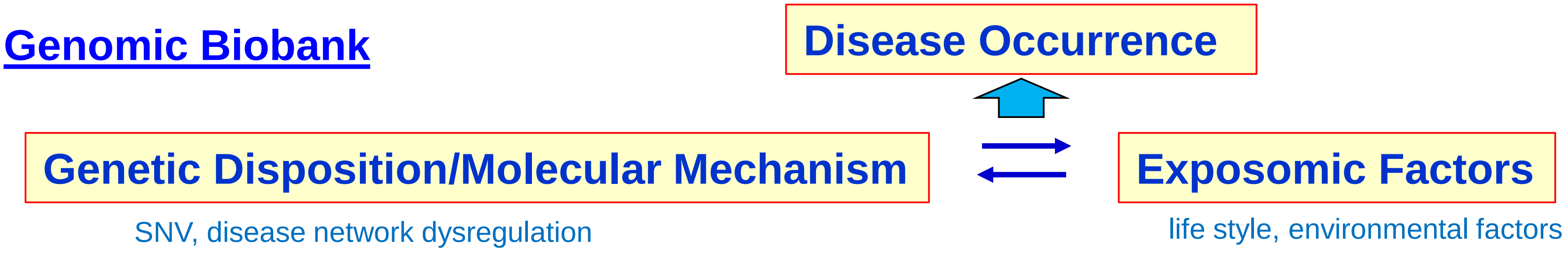

many to many relationship

Comprehensive Molecular Information ⇄ Clinical Phenome

Genome, multi-omics

clinical signs, lab test, medical image

Genomic Biobank


Disease Occurrence

↑

Genetic Disposition/Molecular Mechanism ⇄ Exposomic Factors

SNV, disease network dysregulation

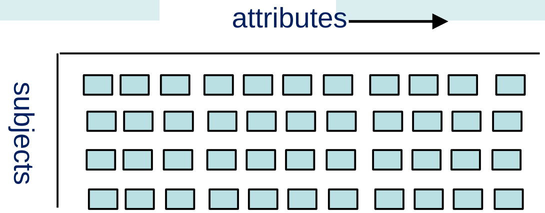
life style, environmental factors

Data Principle of Big Data

Challenge: num. attributes(p)≫num. subjects (n)

p: may be billion

n: at most, tens thousand



If this huge number of attributes are independent, we can not do anything

Big Data・Sparse Assumption

Big data is intrinsically determined by the latent variables, number of which is less than number of subjects

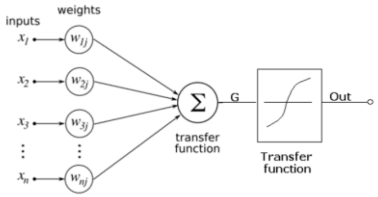
principle of compositionality

Big data is hierarchically composed of **nested structure**

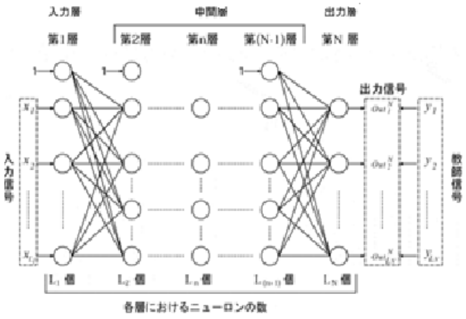
Multi-dimensional medical big data should be contrasted to intrinsic structure

Distinguishably Effective Method
AI, Deep Learning

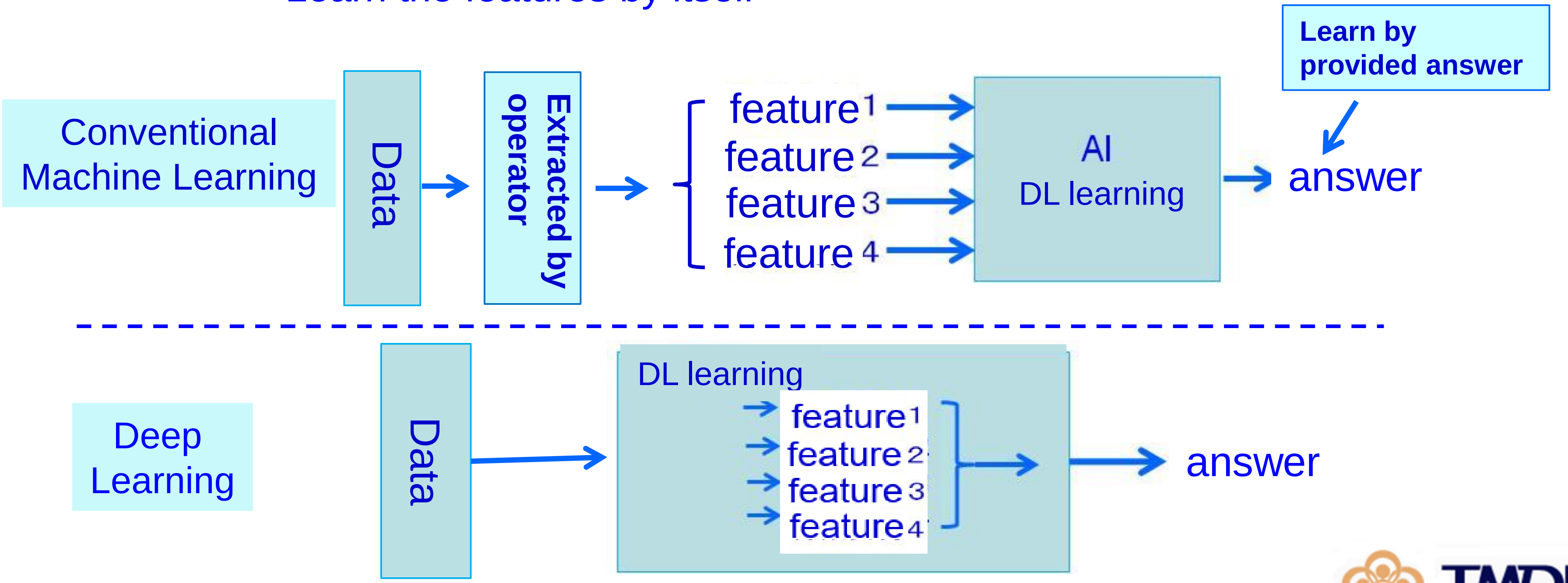
- Limitation of Machine Learning
  - "Supervised learning"
    - Construct AI by providing the feature and answer
- Deep Learning revolution
  - "Unsupervised learning"
    - Learn the features by itself



Neural information element



Multiple Layer neuro-network



Learn by provided answer

Conventional Machine Learning

Data

Extracted by operator

feature 1

feature 2

feature 3

feature 4

AI

DL learning

answer

Deep Learning

Data

DL learning

feature 1

feature 2

feature 3

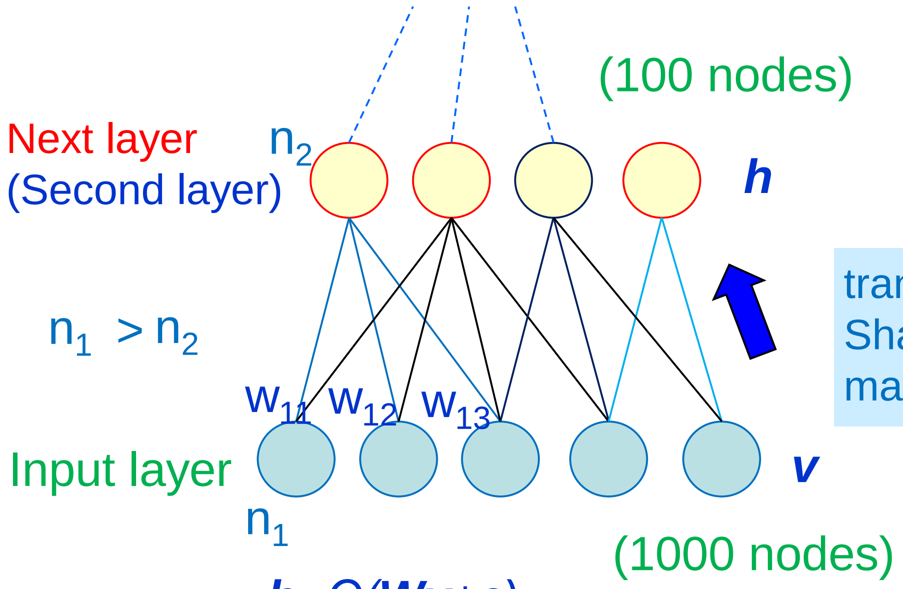
feature 4

answer

Revolutionary point of DL Autoencoder

- Principle of **autoencoder**: Learn specific intrinsic features of the big data
- Restore the node values of input layer from the node values of next layer where the number of nodes is decreasing compared with input layer. → Intrinsic features should be explored so that the input layer to be recovered as same as possible

→ discover intrinsic features



$h=\Theta(Wv+c)$

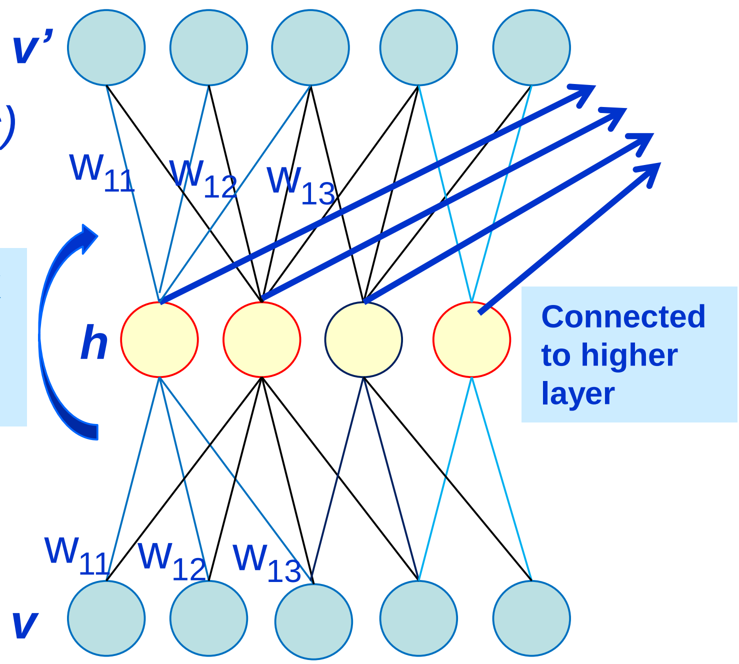
Same as input layer $v \fallingdotseq v'$

$v'=\Theta(W^t h+c)$

$W'=W^t$

transposed matrix
Share the weight matrix

$h=\Theta(Wv+c)$

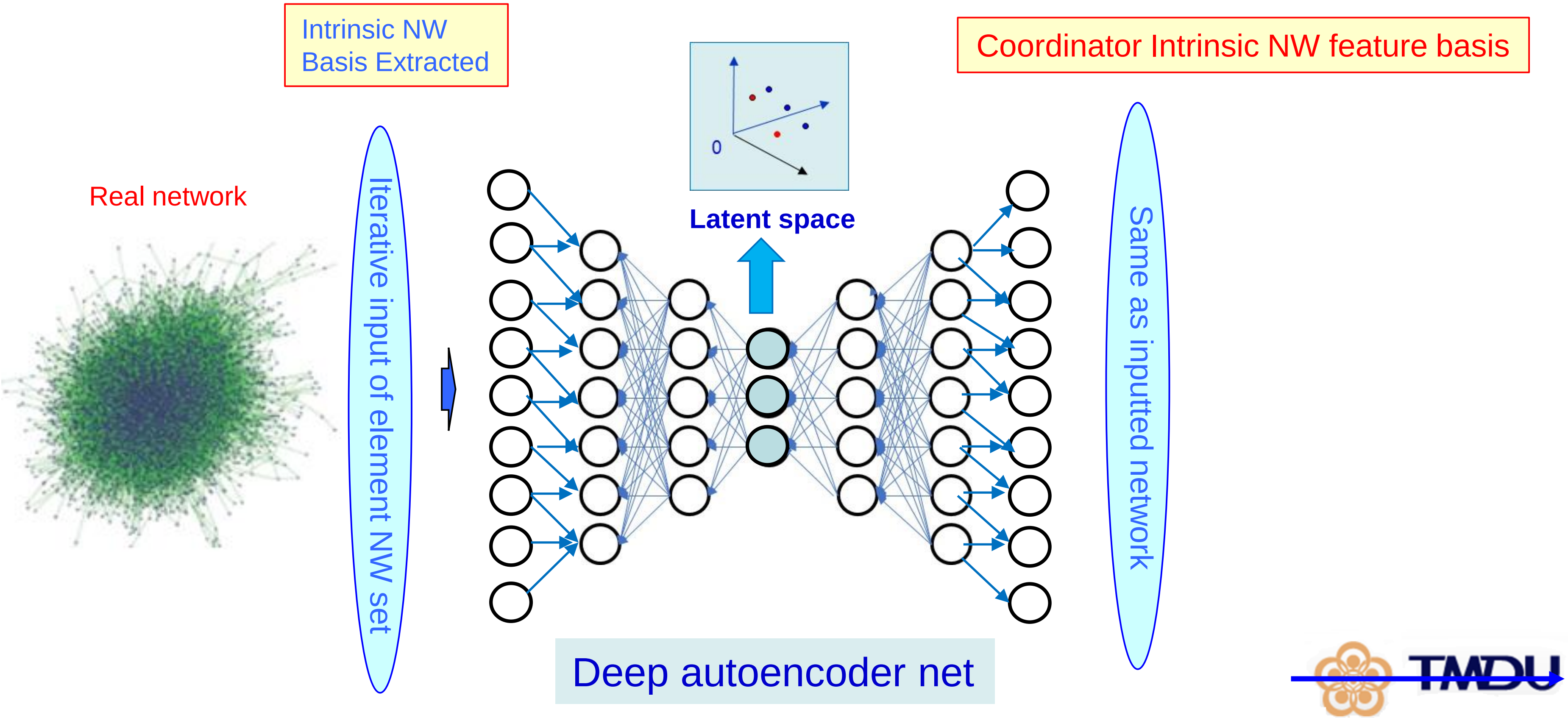

v'
w_{11} w_{12} w_{13}
h
Connected to higher layer
w_{11} w_{12} w_{13}
v

Deep Autoencoder Network

- Deep Learning-based CorrelationNetwork Contraction

Multi-dimensional correlation network information structure
⇒ Contract to be composed of a few network variables

- Projection of data to be composed of intrinsic bases by nonlinear contraction. Contraction to “latent space”



Intrinsic NW
Basis Extracted

Coordinator Intrinsic NW feature basis

Real network

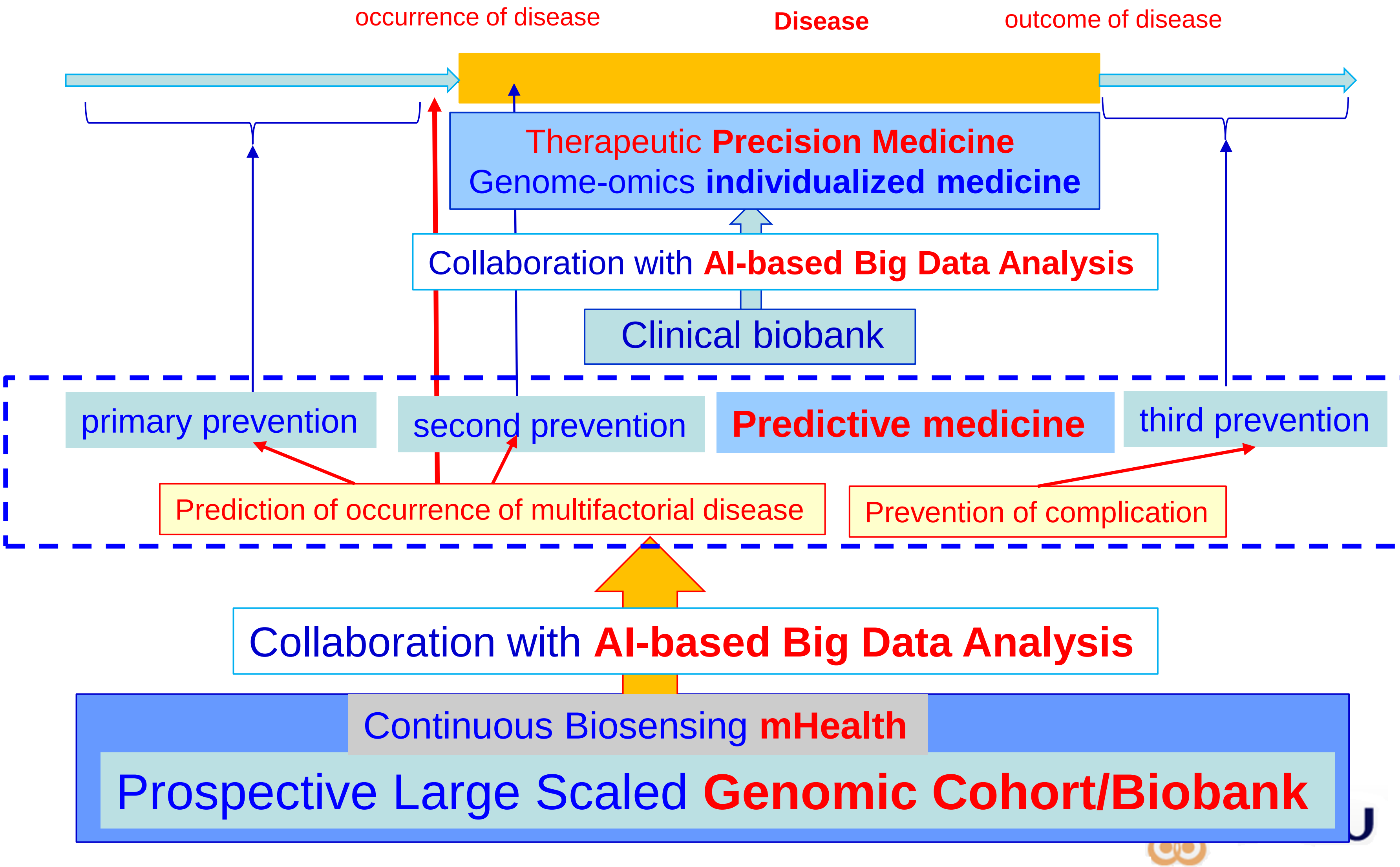
Iterative input of element NW set

Latent space

Same as inputted network

Deep autoencoder net

Integration of Big Data Medicine into life-course oriented healthcare



occurrence of disease

Disease

outcome of disease

Therapeutic **Precision Medicine**

Genome-omics **individualized medicine**

Collaboration with **AI-based Big Data Analysis**

Clinical biobank

primary prevention

second prevention

Predictive medicine

third prevention

Prediction of occurrence of multifactorial disease

Prevention of complication

Collaboration with **AI-based Big Data Analysis**

Continuous Biosensing **mHealth**

Prospective Large Scaled **Genomic Cohort/Biobank**

Future Big Data Medicine

- Genome Medicine, Genomic Biobank and mHealth are integrated in
- **Life-cource oriented healthcare**
  - Understand Individual in **his Totality** with respect to **Overall Susceptibility of Contacting Diseases** through Person's Whole Life
  - 1. throughout **total life span of his life**
    - "from uterus to grave"; DOHaD theory, life course healthcare
  - 2. throughout **total ecosystem he lives in**
    - Gut Microbiome as mediator between environment factor and biosystem, basis of various diseases

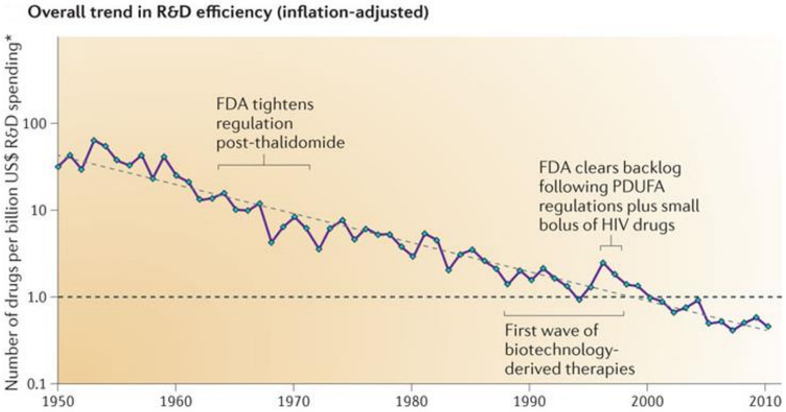
AI-based Drug Discovery

Current Situation of Drug Discovery

- Rapid increase of R&D expenditure
  - More than 1B $ for one marketed drug
- Decrease of success rate

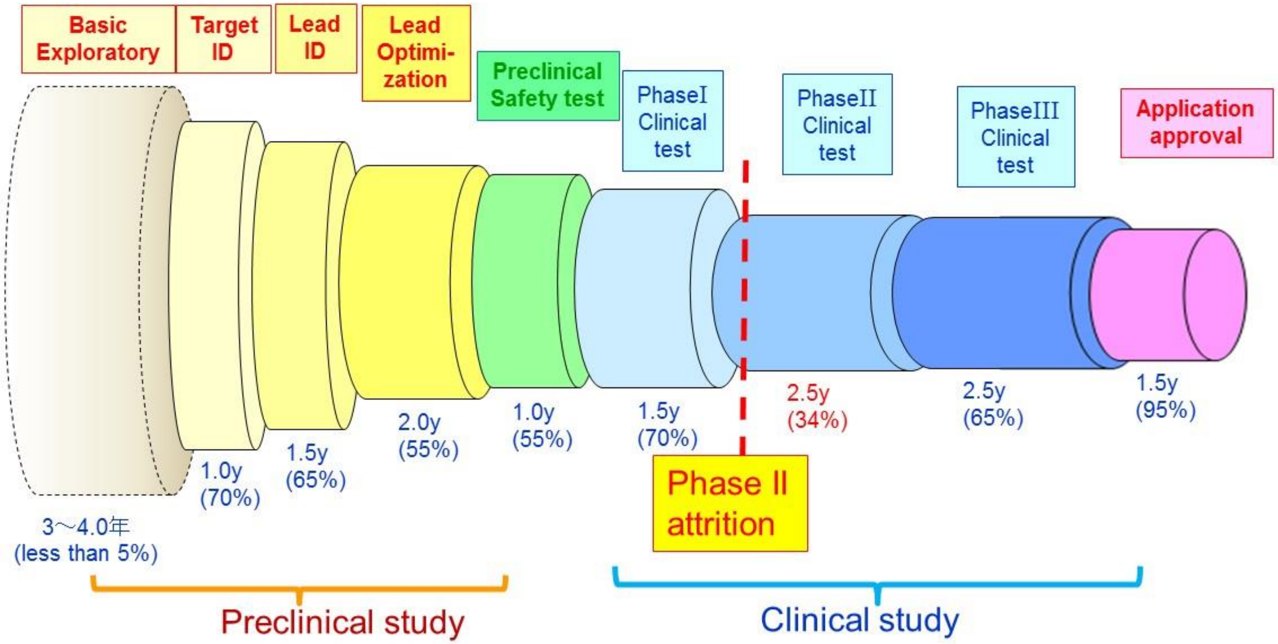
  now about 1/20,000～1/30,000
  - Remarkable Drop Between non-clinical and clinical test **(phase II attrition)**
- **Clinical Predictability**
  - At as early as possible stage,

    Estimation of clinical efficacy and toxicity
- **Efficient measures**
  - **Use Disease-specific iPS cell**
  - **Use of Human Bio Big Data in early stage**



Overall trend in R&D efficiency (inflation-adjusted)

Nature Reviews Drug Discovery (2012）



Basic Exploratory

Target ID

Lead ID

Lead Optimi-zation

Preclinical Safety test

PhaseI Clinical test

PhaseII Clinical test

PhaseIII Clinical test

Application approval

3～4.0年 (less than 5%)

1.0y (70%)

1.5y (65%)

2.0y (55%)

1.0y (55%)

1.5y (70%)

2.5y (34%)

2.5y (65%)

1.5y (95%)

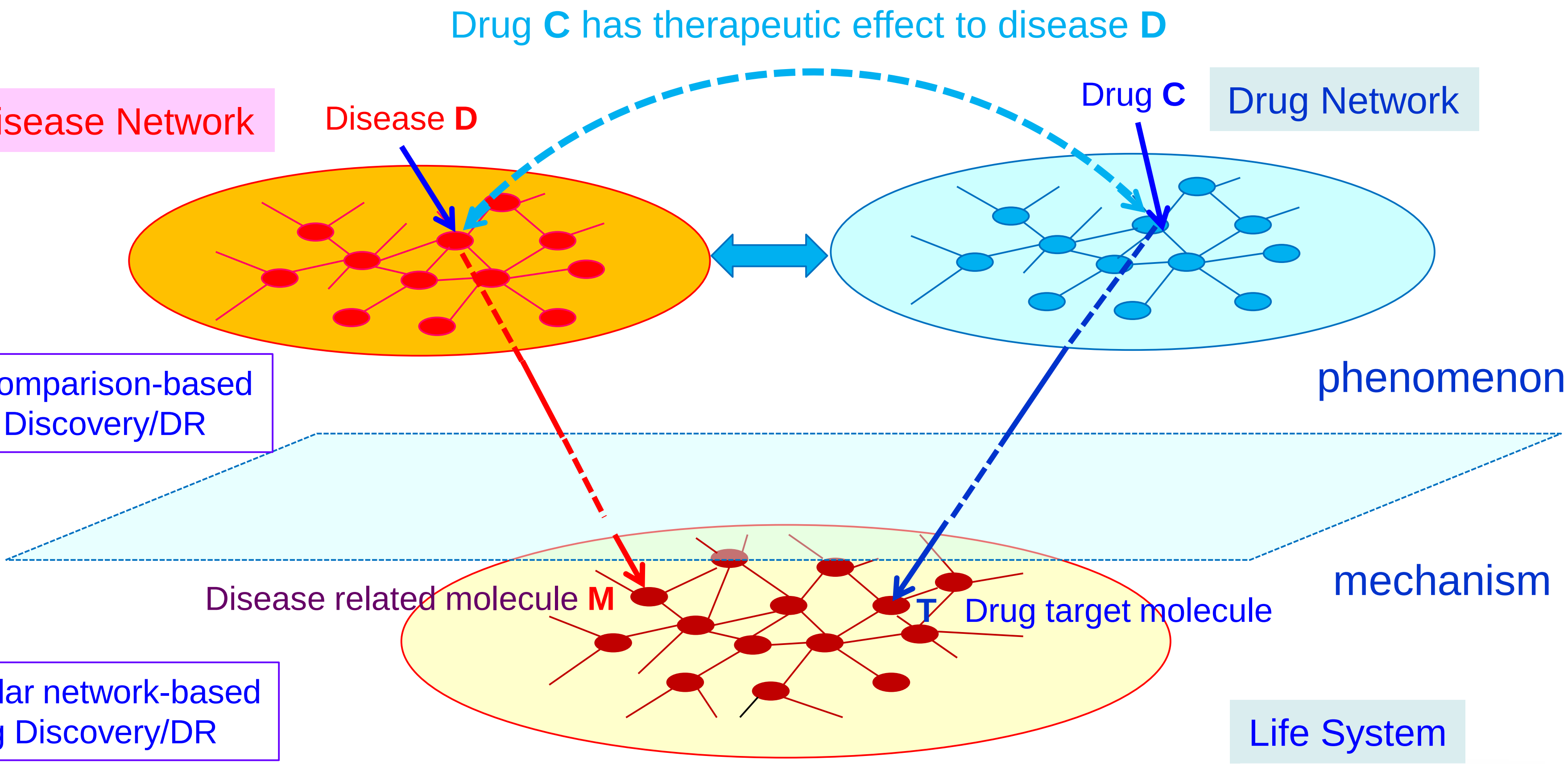
Phase II attrition

Preclinical study

Clinical study

Basic structure of profile-based computational drug discovery

Framework of Triple-layer disease and drug network



Drug **C** has therapeutic effect to disease **D**

Disease Network

Disease **D**

Drug **C**

Drug Network

Profile comparison-based
Drug Discovery/DR

phenomenon

mechanism

Disease related molecule **M**

T Drug target molecule

Molecular network-based
Drug Discovery/DR

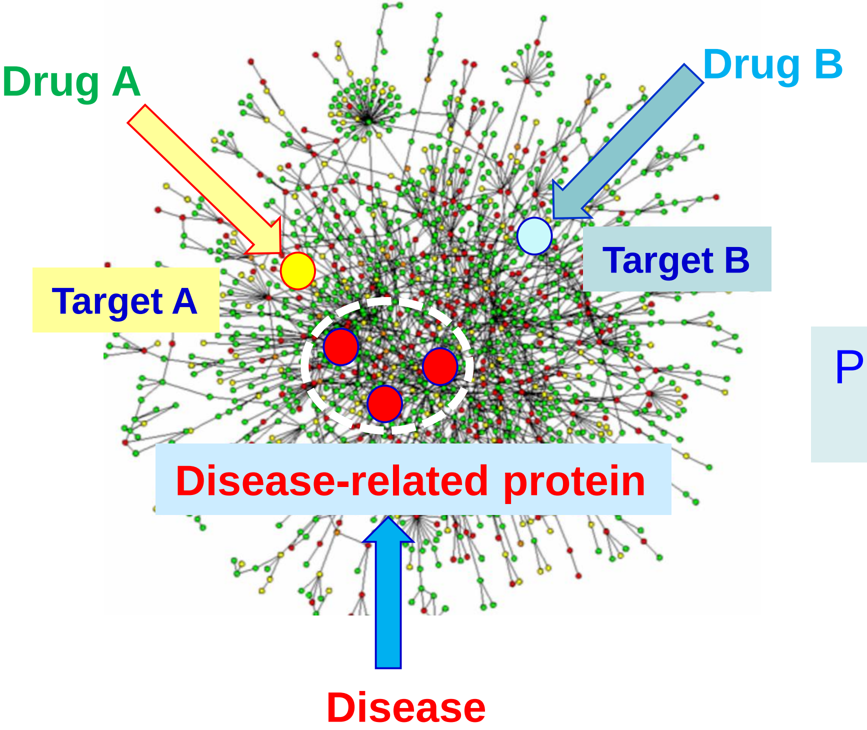
Life System

DR: Drug Repositioning: is the application of known drugs (compounds) to treat new indications (i.e., new diseases)

Common Platform of DrugDiscovery/DR
Protein-Protein interaction network（PPIN）

- Common Platform bionetwork: mediating disease and drug action
- **Protein-protein interaction network（PPIN）** as common platform
- **Disease:** Scaffolding in PPIN: **Disease-related protein** (gene)
- **Drug** : Scaffolding in PPIN**: Drug Target protein**
- Based on **the distance (proximity)** between **Disease-related protein** and **target protein**,

  the impact of the drug is measured



Protein-protein Interaction Network（PPIN）

Proximity between Drug and Disease at PPIN

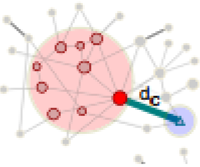
Relaltive Proximity Index d_c :

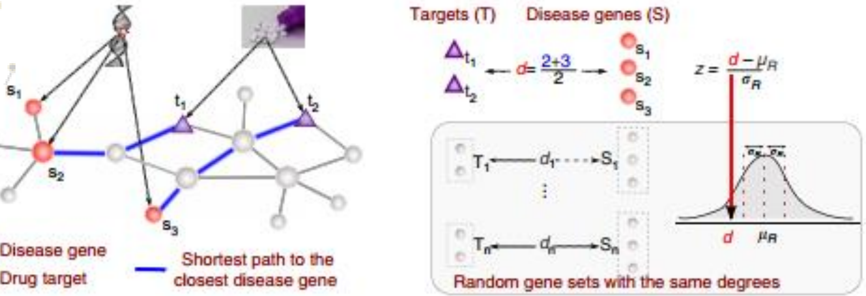
①Distance between Target and the nearest protein among disease-related protein module

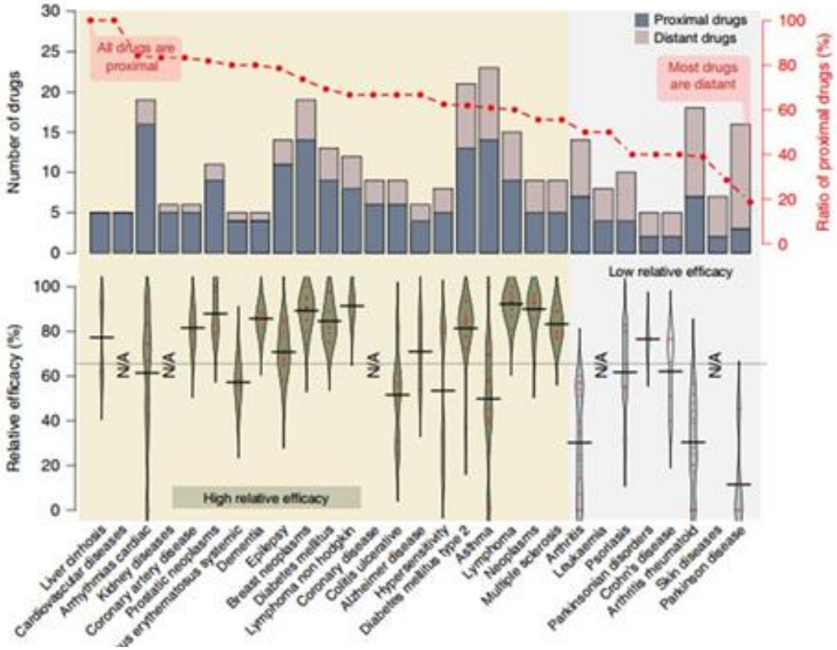
②The distance is normalized among the distance of the molecules in same context

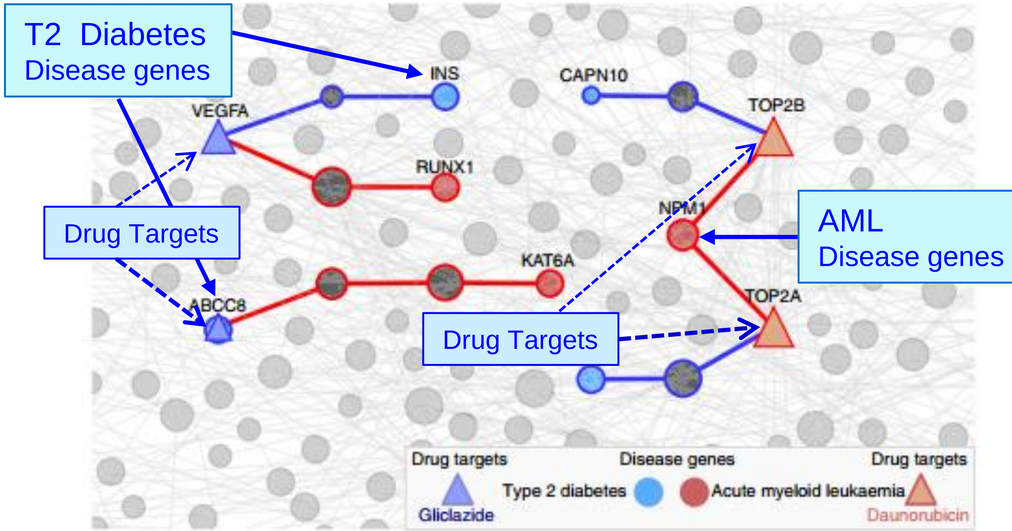
$z < -0.15$ ⇒proximate

③closest measure d_c : best index to measure efficacy







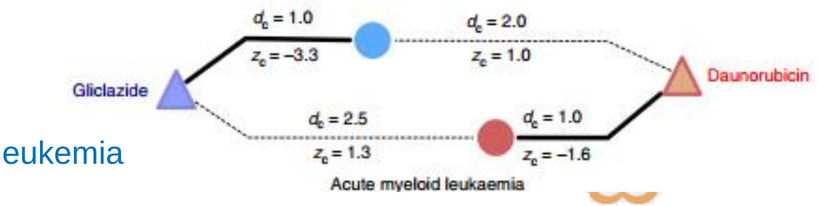

T2 Diabetes
Disease genes

Drug Targets

AML
Disease genes

Drug Targets

Average distance: about 2 rinks



(Guney, Barabasi,2016, Nat. Com)

AML: acute myeloid leukemia

Need for Learning

- We are **still missing in understanding** of the necessary conditions for molecule to be effective to disease
- We should find these conditions by **learning from the succeeded <disease-drug-target molecule> combinations**
- **Artificial Intelligence** (AI), specially **Deep Learning** is now the most powerful method

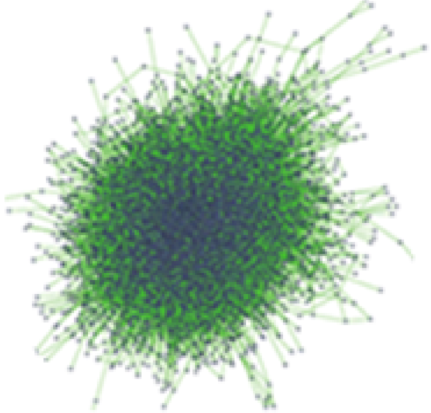
Our Approach

- By using deep learning and genome-wide protein interaction network,
- We build a computational framework to predict potential Drug Target genes and
- Repositionable drugs for Alzheimer’s disease.

Our computational workflow

Step1: Input data

Genome-wide PIN

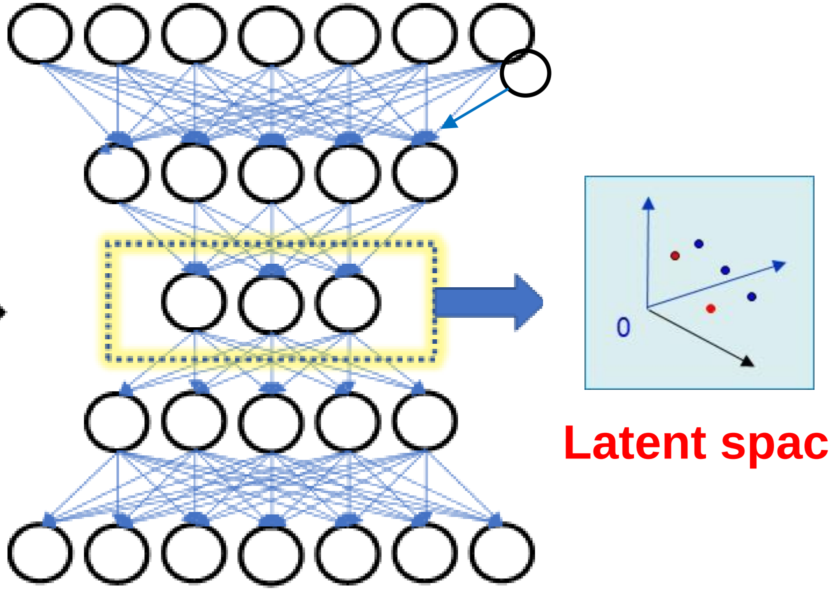


Drugs and their targets information

Step 2: Feature Engineering

Feature engineering by "deep autoencoder" and a state-of-the-art feature selection algorithm

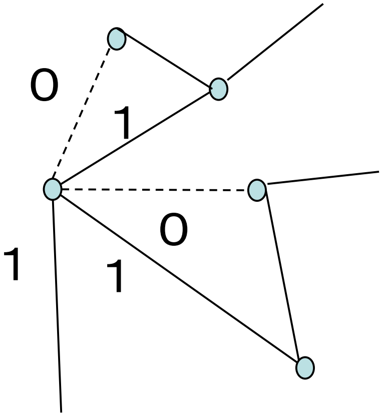
Dimensional reduction by "deep autoencoder"

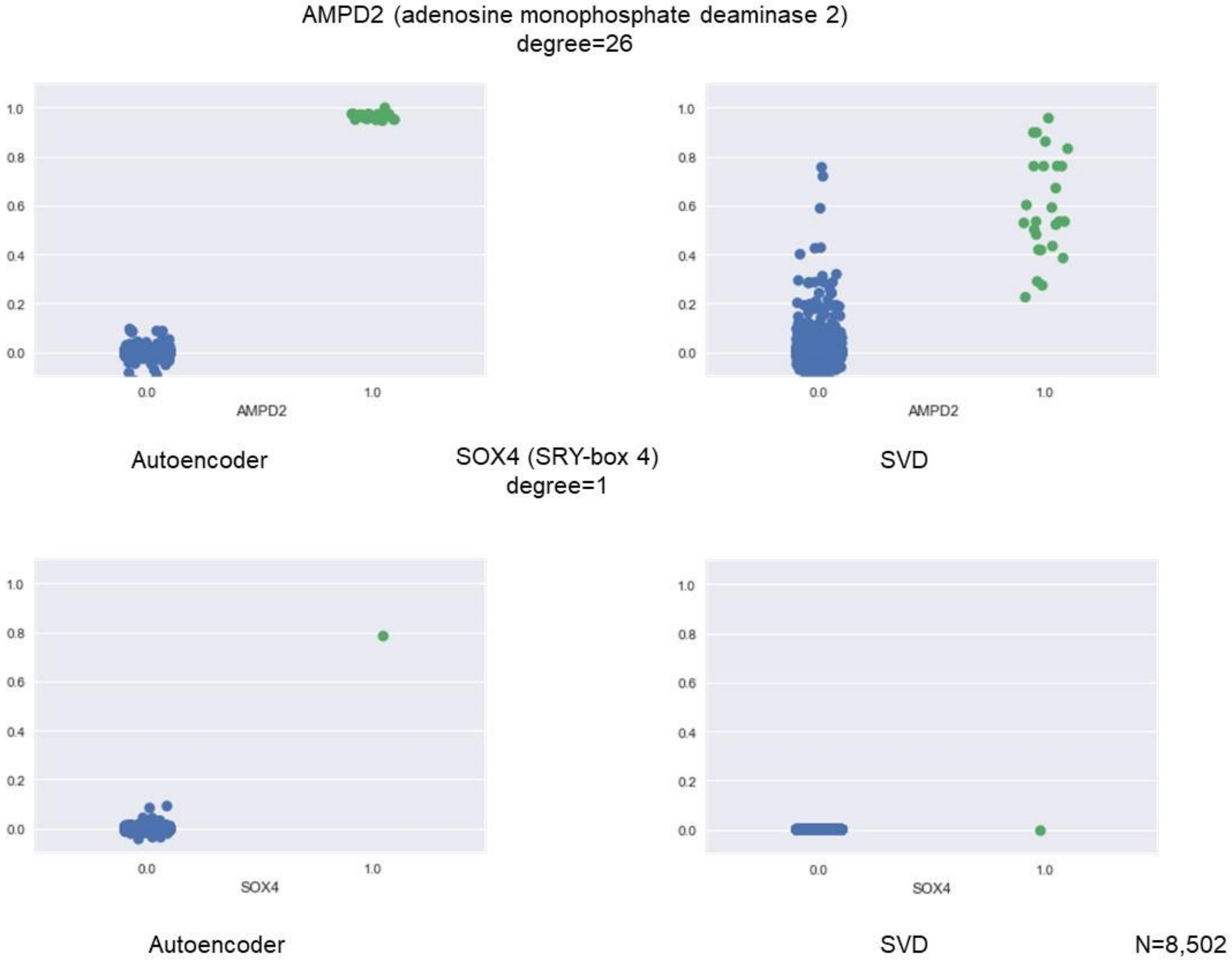


Latent space

Restoration Accuracy between **Deep Learning** and **SVD** (singular value decomposition)

For a certain protein, the connections are described by adjacency vector;
(0,0,0,1,0,1,0,……..),
where
0 $_{(i)}$: not connected
to i th node
1$_{(i)}$: connected to i th node




AMPD2 (adenosine monophosphate deaminase 2)
degree=26

Autoencoder

SVD

SOX4 (SRY-box 4)
degree=1

Autoencoder

SVD

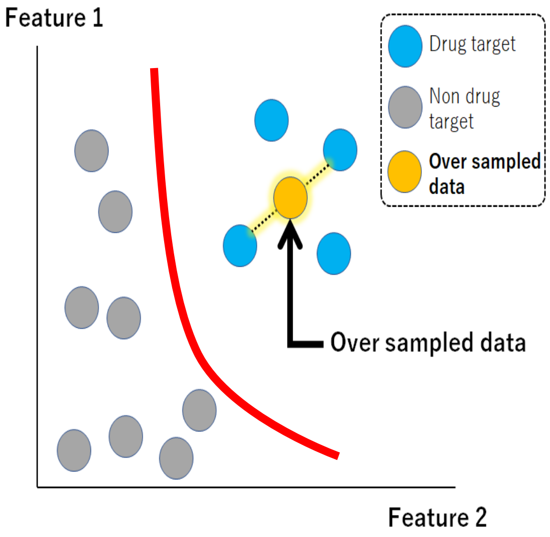
N=8,502

Step 3: Classifier model

Step 4: Target prioritization

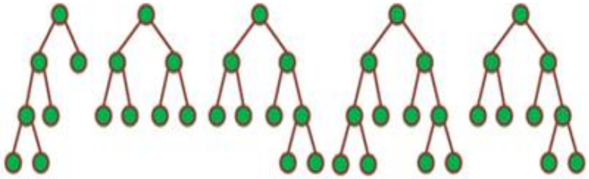
A binary classifier model to target prioritization by state-of-the-art machine learning algorithms

SMOTE algorithm to build a training data

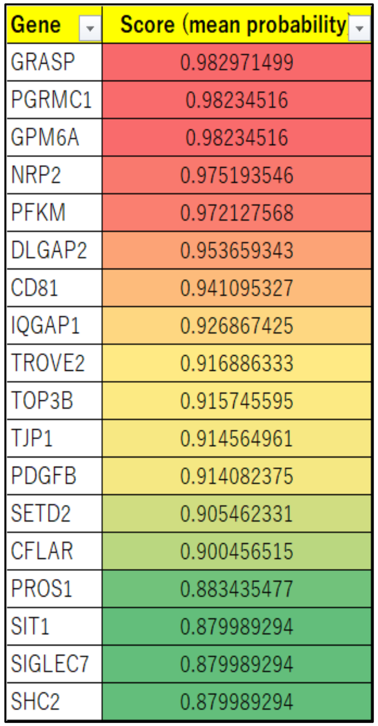


Latent space

Xgboost algorithm to build a binary classfier



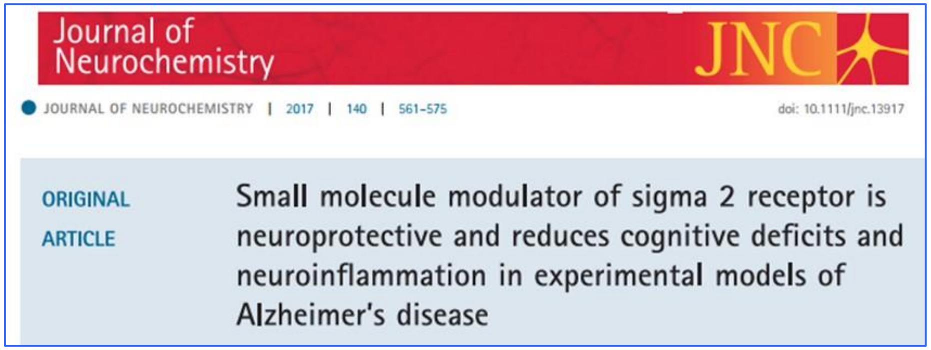
Scores for potential targets



Gene	Score (mean probability
GRASP	0.982971499
PGRMC1	0.98234516
GPM6A	0.98234516
NRP2	0.975193546
PFKM	0.972127568
DLGAP2	0.953659343
CD81	0.941095327
IQGAP1	0.926867425
TROVE2	0.916886333
TOP3B	0.915745595
TJP1	0.914564961
PDGFB	0.914082375
SETD2	0.905462331
CFLAR	0.900456515
PROS1	0.883435477
SIT1	0.879989294
SIGLEC7	0.879989294
SHC2	0.879989294

Correspondent with Wet reseach

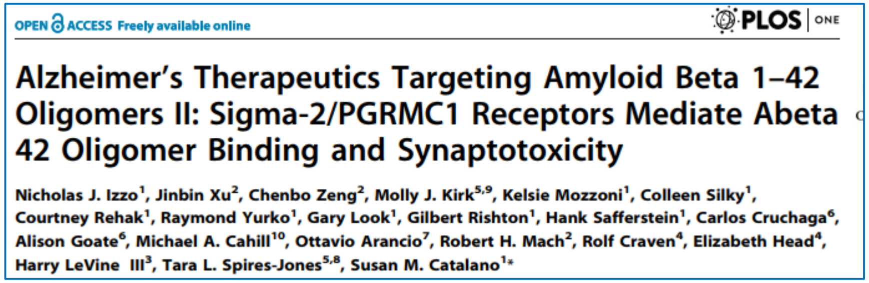
PGCM1：progesterone receptor membrane 1


Journal of Neurochemistry JNC

JOURNAL OF NEUROCHEMISTRY | 2017 | 140 | 561–575 doi: 10.1111/jnc.13917

ORIGINAL ARTICLE

Small molecule modulator of sigma 2 receptor is neuroprotective and reduces cognitive deficits and neuroinflammation in experimental models of Alzheimer's disease

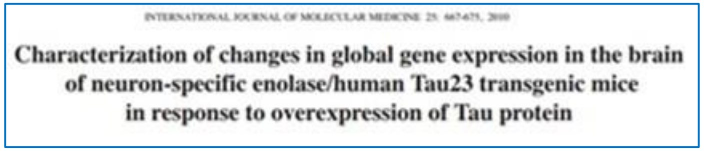

OPEN ACCESS Freely available online PLOS ONE

Alzheimer's Therapeutics Targeting Amyloid Beta 1–42 Oligomers II: Sigma-2/PGRMC1 Receptors Mediate Abeta 42 Oligomer Binding and Synaptotoxicity

Nicholas J. Izzo[1], Jinbin Xu[2], Chenbo Zeng[2], Molly J. Kirk[5,9], Kelsie Mozzoni[1], Colleen Silky[1], Courtney Rehak[1], Raymond Yurko[1], Gary Look[1], Gilbert Rishton[1], Hank Safferstein[1], Carlos Cruchaga[6], Alison Goate[6], Michael A. Cahill[10], Ottavio Arancio[7], Robert H. Mach[2], Rolf Craven[4], Elizabeth Head[4], Harry LeVine III[3], Tara L. Spires-Jones[5,8], Susan M. Catalano[1]*

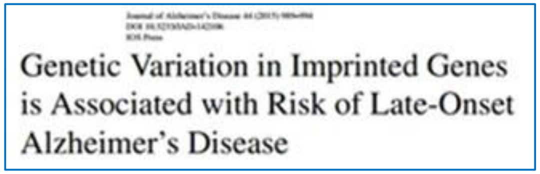
神経保護的効果（neuroprotective)認知不全・炎症に治療効果

GPM6A：Glycoprotein M6A


INTERNATIONAL JOURNAL OF MOLECULAR MEDICINE 25: 667-675, 2010

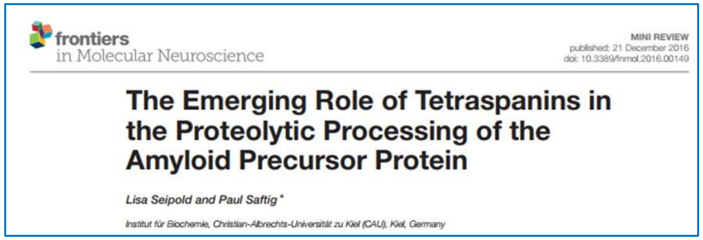
Characterization of changes in global gene expression in the brain of neuron-specific enolase/human Tau23 transgenic mice in response to overexpression of Tau protein

DLGAP2：DLG-Associated Protein 2


Journal of Alzheimer's Disease 44 (2015) 989–994
DOI 10.3233/JAD-142106
IOS Press

Genetic Variation in Imprinted Genes is Associated with Risk of Late-Onset Alzheimer's Disease

CD81:Tetraspanins family


frontiers in Molecular Neuroscience

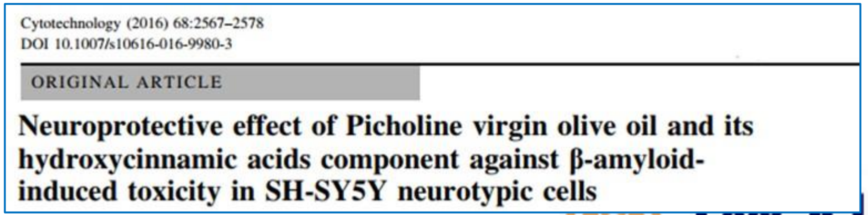
MINI REVIEW
published: 21 December 2016
doi: 10.3389/fnmol.2016.00149

The Emerging Role of Tetraspanins in the Proteolytic Processing of the Amyloid Precursor Protein

Lisa Seipold and Paul Saftig*

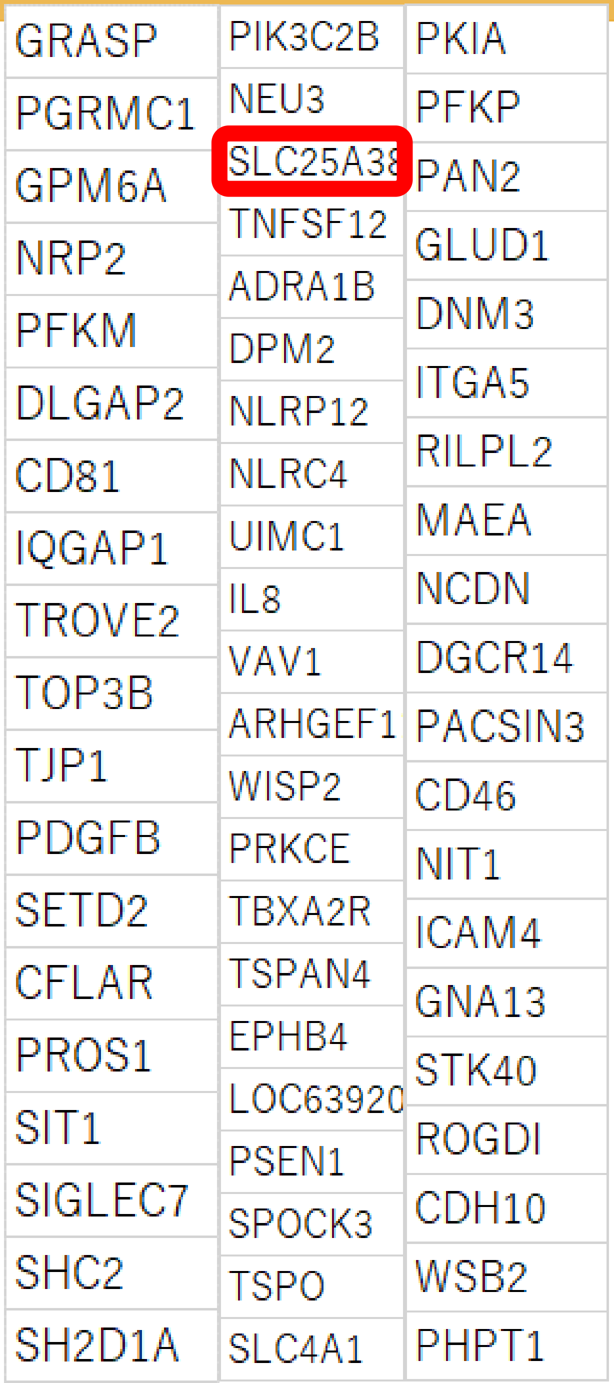
Institut für Biochemie, Christian-Albrechts-Universität zu Kiel (CAU), Kiel, Germany

PFKM: Phospofructokinase


Cytotechnology (2016) 68:2567–2578
DOI 10.1007/s10616-016-9980-3

ORIGINAL ARTICLE

Neuroprotective effect of Picholine virgin olive oil and its hydroxycinnamic acids component against β-amyloid-induced toxicity in SH-SY5Y neurotypic cells



GRASP
PGRMC1
GPM6A
NRP2
PFKM
DLGAP2
CD81
IQGAP1
TROVE2
TOP3B
TJP1
PDGFB
SETD2
CFLAR
PROS1
SIT1
SIGLEC7
SHC2
SH2D1A

PIK3C2B
NEU3
SLC25A38
TNFSF12
ADRA1B
DPM2
NLRP12
NLRC4
UIMC1
IL8
VAV1
ARHGEF1
WISP2
PRKCE
TBXA2R
TSPAN4
EPHB4
LOC63920
PSEN1
SPOCK3
TSPO
SLC4A1

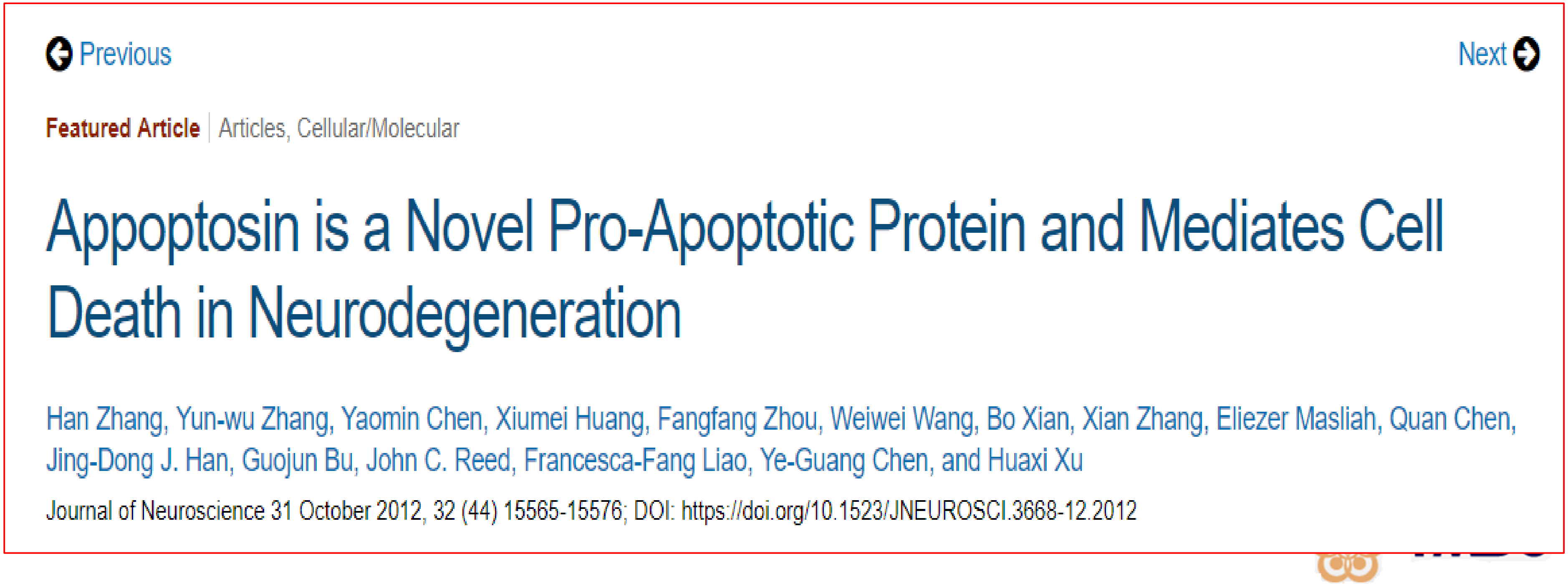
PKIA
PFKP
PAN2
GLUD1
DNM3
ITGA5
RILPL2
MAEA
NCDN
DGCR14
PACSIN3
CD46
NIT1
ICAM4
GNA13
STK40
ROGDI
CDH10
WSB2
PHPT1

By using the AI-based method, we successfully predict potential drug targets (more than 100 genes) for Alzheimer’s disease.

Example,

SLC25A38 (APPOPTOSIN)

SLC25A3 increases in the brain from Alzheimer’s disease patients as well as from infarct patients. Further, SLC25A38 downregulation is likely to inhibit apoptosis induced by Bax/BH3I and neuronal death induced by Aβ/glutamate.



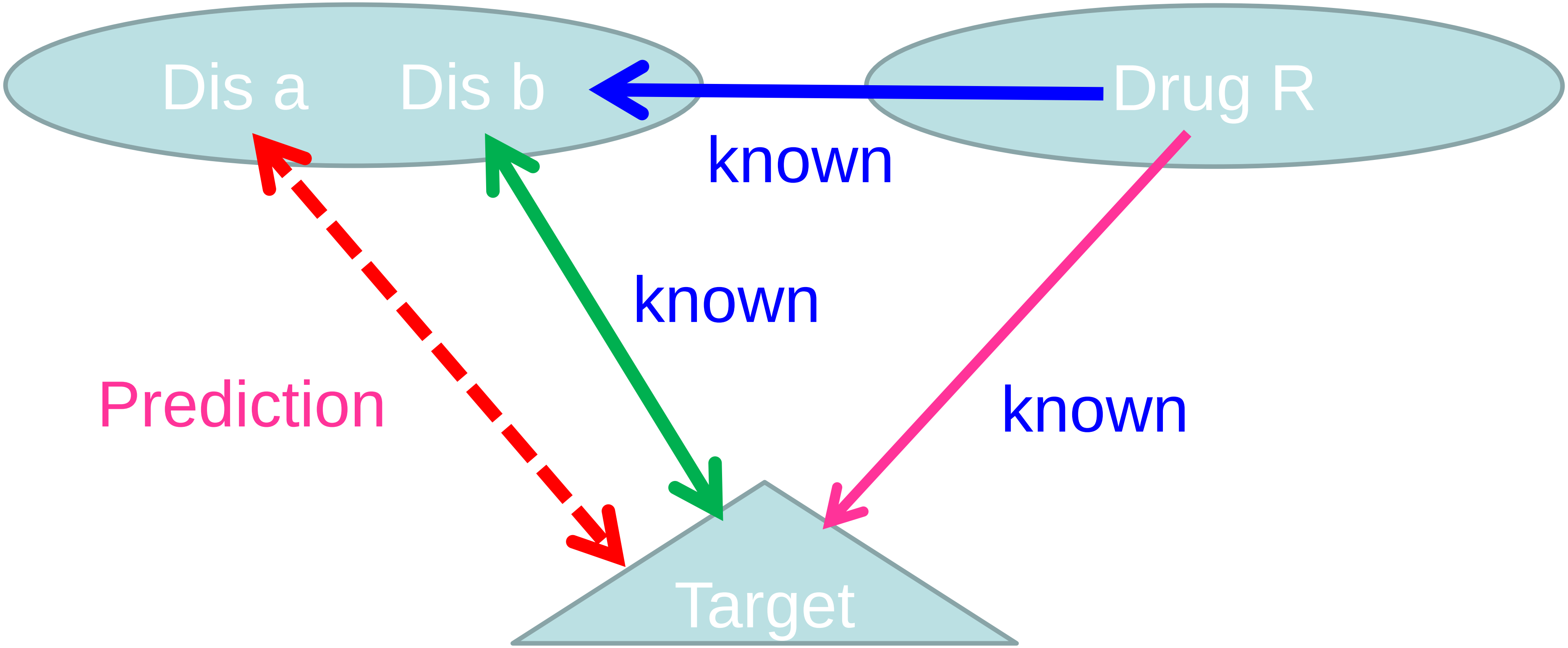
Featured Article | Articles, Cellular/Molecular

Appoptosin is a Novel Pro-Apoptotic Protein and Mediates Cell Death in Neurodegeneration

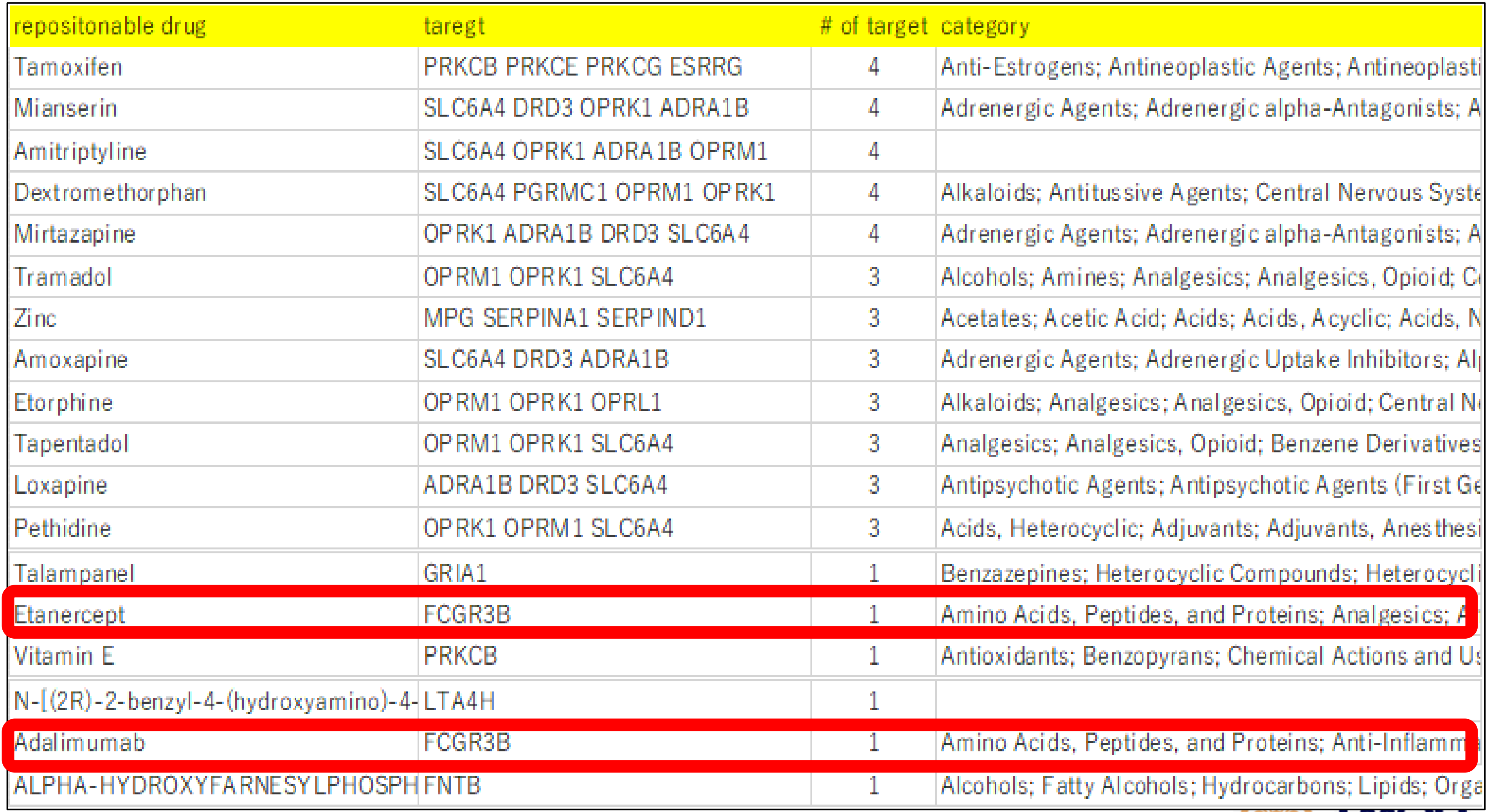
Han Zhang, Yun-wu Zhang, Yaomin Chen, Xiumei Huang, Fangfang Zhou, Weiwei Wang, Bo Xian, Xian Zhang, Eliezer Masliah, Quan Chen, Jing-Dong J. Han, Guojun Bu, John C. Reed, Francesca-Fang Liao, Ye-Guang Chen, and Huaxi Xu

Journal of Neuroscience 31 October 2012, 32 (44) 15565-15576; DOI: https://doi.org/10.1523/JNEUROSCI.3668-12.2012

If predicted target for disease A is known drug-target of drug R for disease B, the drug R may be repositionable drug for disease A.



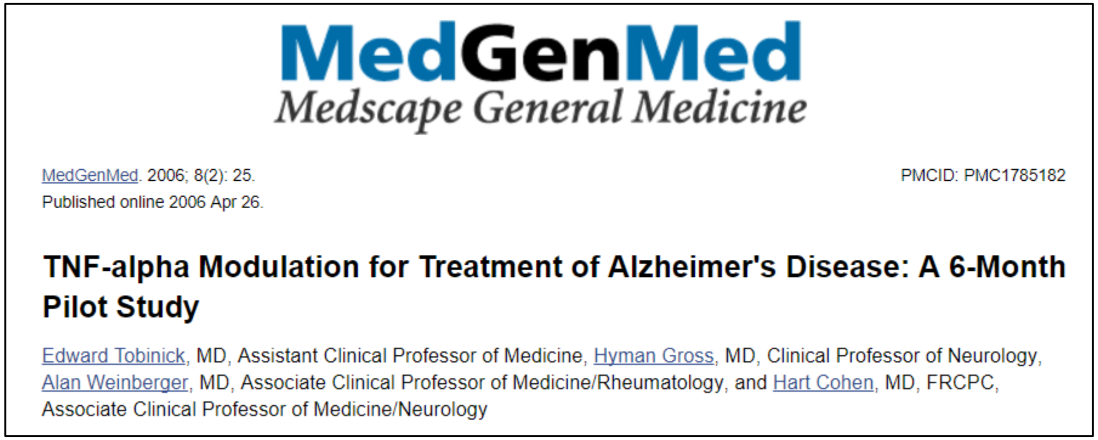
Potential (predicted) repositionable drugs for Alzheimer’s disease



repositonable drug	taregt	# of target	category
Tamoxifen	PRKCB PRKCE PRKCG ESRRG	4	Anti-Estrogens; Antineoplastic Agents; Antineoplasti
Mianserin	SLC6A4 DRD3 OPRK1 ADRA1B	4	Adrenergic Agents; Adrenergic alpha-Antagonists; A
Amitriptyline	SLC6A4 OPRK1 ADRA1B OPRM1	4	
Dextromethorphan	SLC6A4 PGRMC1 OPRM1 OPRK1	4	Alkaloids; Antitussive Agents; Central Nervous Syste
Mirtazapine	OPRK1 ADRA1B DRD3 SLC6A4	4	Adrenergic Agents; Adrenergic alpha-Antagonists; A
Tramadol	OPRM1 OPRK1 SLC6A4	3	Alcohols; Amines; Analgesics; Analgesics, Opioid; C
Zinc	MPG SERPINA1 SERPIND1	3	Acetates; Acetic Acid; Acids; Acids, Acyclic; Acids, N
Amoxapine	SLC6A4 DRD3 ADRA1B	3	Adrenergic Agents; Adrenergic Uptake Inhibitors; Al
Etorphine	OPRM1 OPRK1 OPRL1	3	Alkaloids; Analgesics; Analgesics, Opioid; Central N
Tapentadol	OPRM1 OPRK1 SLC6A4	3	Analgesics; Analgesics, Opioid; Benzene Derivatives
Loxapine	ADRA1B DRD3 SLC6A4	3	Antipsychotic Agents; Antipsychotic Agents (First Ge
Pethidine	OPRK1 OPRM1 SLC6A4	3	Acids, Heterocyclic; Adjuvants; Adjuvants, Anesthesi
Talampanel	GRIA1	1	Benzazepines; Heterocyclic Compounds; Heterocycli
Etanercept	FCGR3B	1	Amino Acids, Peptides, and Proteins; Analgesics; A
Vitamin E	PRKCB	1	Antioxidants; Benzopyrans; Chemical Actions and Us
N-[(2R)-2-benzyl-4-(hydroxyamino)-4-	LTA4H	1	
Adalimumab	FCGR3B	1	Amino Acids, Peptides, and Proteins; Anti-Inflamm
ALPHA-HYDROXYFARNESYLPHOSPH	FNTB	1	Alcohols; Fatty Alcohols; Hydrocarbons; Lipids; Orga

Example,

The two FDA-approved drugs, **adalimumab and etanercept**, may be most promising candidates, because they are inhibitors of TNF-alpha (a key cytokine to regulate immune response) and overexpression of TNF-alpha cause inflammation in various organs, especially in central nerve system.



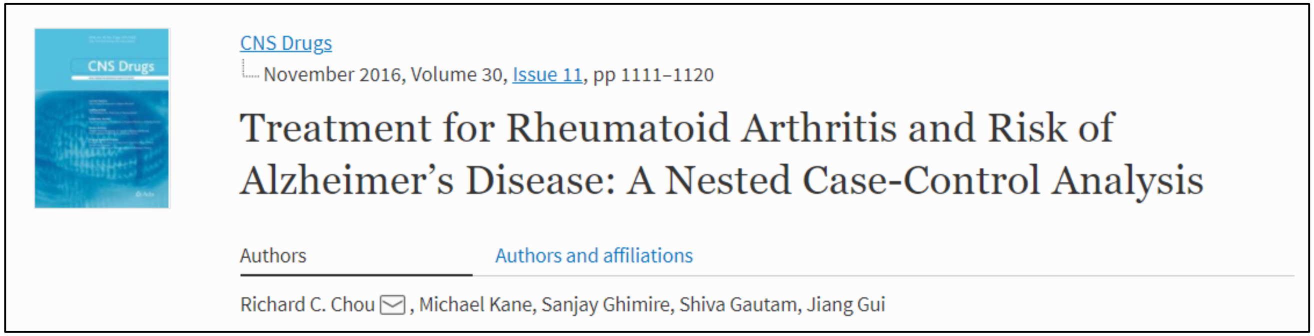
MedGenMed
Medscape General Medicine

MedGenMed. 2006; 8(2): 25.
Published online 2006 Apr 26.

PMCID: PMC1785182

TNF-alpha Modulation for Treatment of Alzheimer's Disease: A 6-Month Pilot Study

Edward Tobinick, MD, Assistant Clinical Professor of Medicine, Hyman Gross, MD, Clinical Professor of Neurology, Alan Weinberger, MD, Associate Clinical Professor of Medicine/Rheumatology, and Hart Cohen, MD, FRCPC, Associate Clinical Professor of Medicine/Neurology



CNS Drugs
November 2016, Volume 30, Issue 11, pp 1111–1120

Treatment for Rheumatoid Arthritis and Risk of Alzheimer's Disease: A Nested Case-Control Analysis

Authors | Authors and affiliations

Richard C. Chou, Michael Kane, Sanjay Ghimire, Shiva Gautam, Jiang Gui

Future strategies and trends

- Big Data era of **genomic medicine and drug discovery** has come
- **Contracting multidimensional network by Deep Learning**
  - Apply to big data in medicine
  - Correlative network structure of <comprehensive molecular information – clinical phenotype> in genome medicine
  - Disease onset and <interaction between genetic – environmental factors> in biobank
- AI drug discovery has now ready to be realized
- We started to organize “Big data medicine/AI drug discovery consortium of Japan” to promote the project, coordinated by pharmaceutical company, IT company and medical institution

Thank you for kind attention