

実験プロトコルの電子化



山田 涼太
fuku株式会社
 @roy29fuku

Google Scholar



☐ すべての言語 ☒ 英語 と 日本語のページを検索

COVID-19に関する記事

CDC

NEJM

JAMA

Lancet

Cell

BMJ

Nature

Science

Elsevier

Oxford

Wiley

medRxiv

巨人の肩の上に立つ

Google Scholar



☐ すべての言語 ☒ 英語 と 日本語のページを検索

COVID-19に関する記事

CDC

NEJM

JAMA

Lancet

Cell

BMJ

Nature

Science

Elsevier

Oxford

Wiley

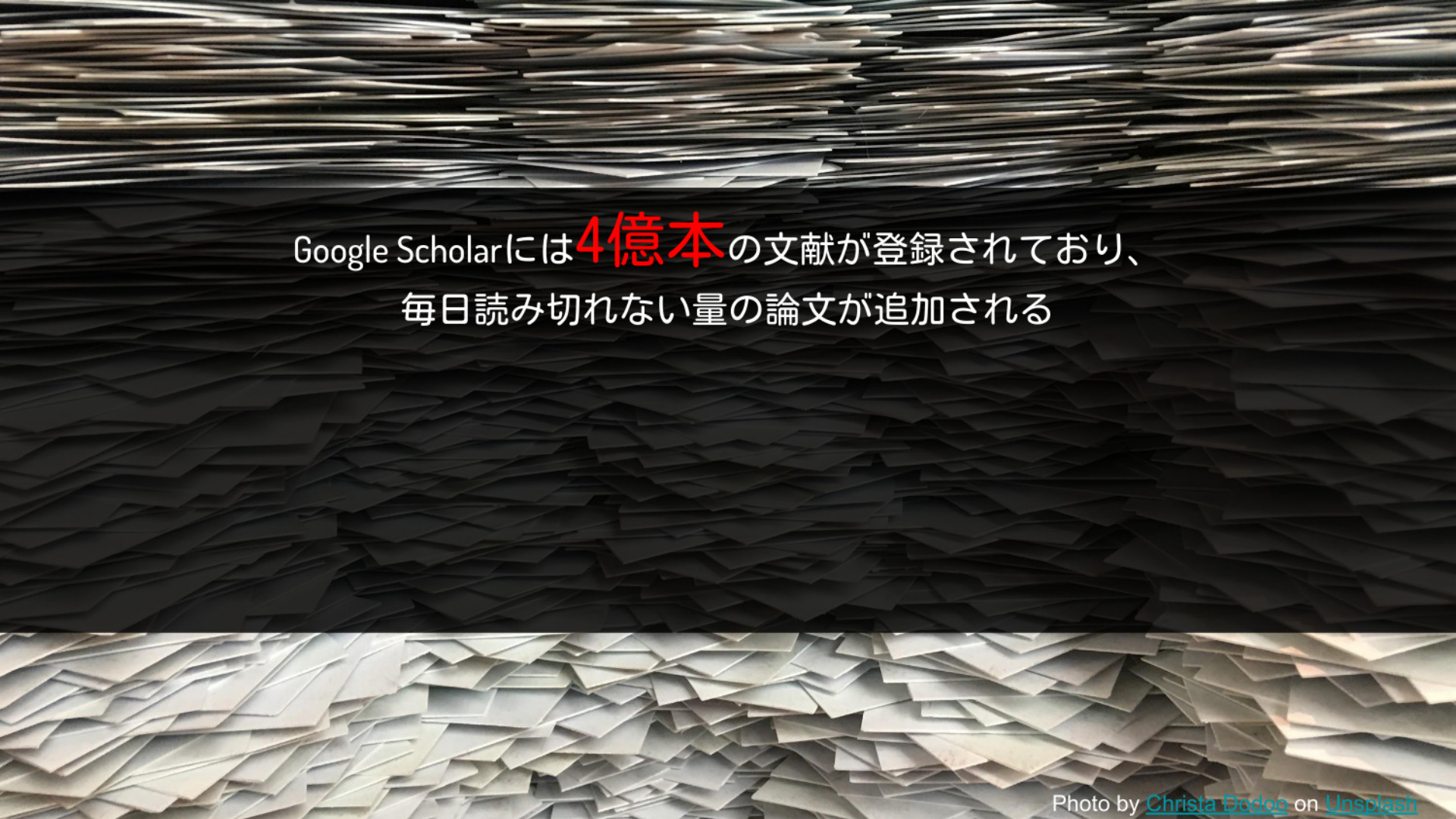
medRxiv

巨人の肩の上に立つ

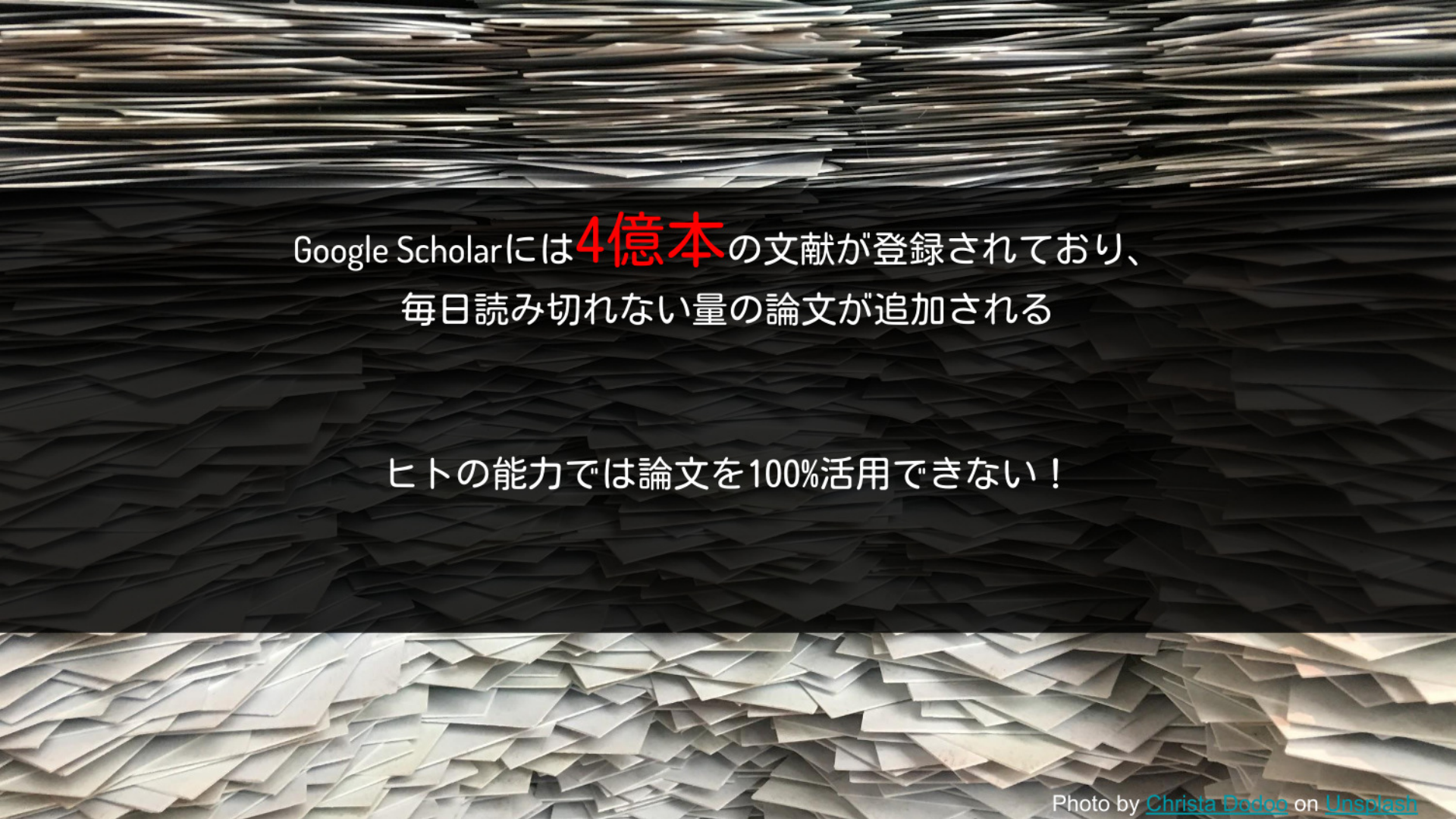


私たちは巨人の肩の上に
うまく立てているだろうか？ 🤔

巨人の肩の上に立つ

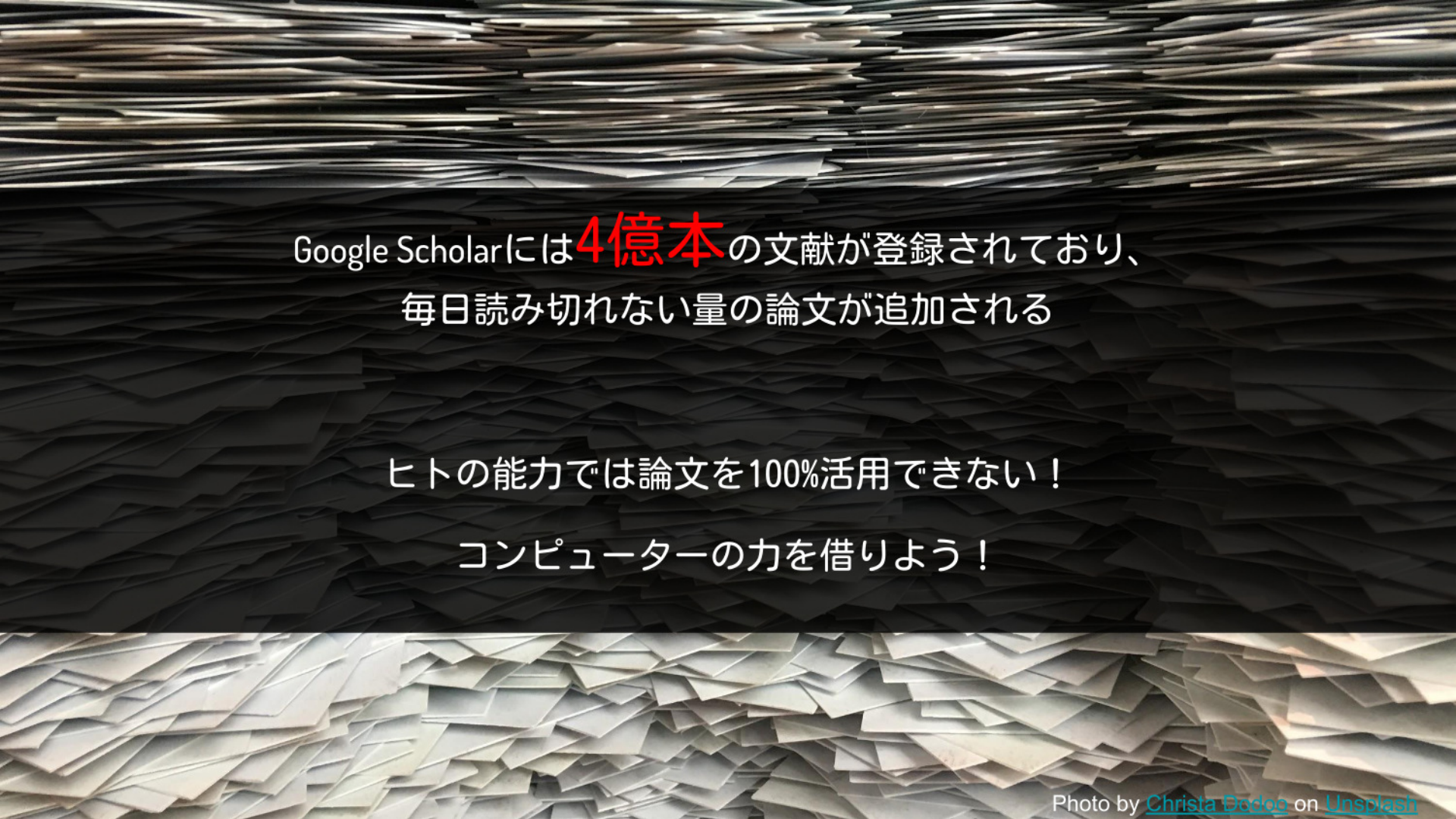


Google Scholarには**4億本**の文献が登録されており、
毎日読み切れない量の論文が追加される



Google Scholarには**4億本**の文献が登録されており、
毎日読み切れない量の論文が追加される

ヒトの能力では論文を100%活用できない！



Google Scholarには**4億本**の文献が登録されており、
毎日読み切れない量の論文が追加される

ヒトの能力では論文を100%活用できない！
コンピューターの力を借りよう！

Ornithine and breast cancer: a matched case–control study

Jiayi Zhang, Baihui Tao, Yiran Chong, Shuang Ma, Gang Wu, Hailong Zhu, Yi Zhao, Shitao Zhao, Mengmeng Niu, Shutian Zhang, Tianyi Wang, Shuman Yang, Wenjing Qiao, Ann M. Vuong, Jincheng Li, Demiao Zhu & Wei Tao 

Scientific Reports **10**, Article number: 15502 (2020) | [Cite this article](#)

109 Accesses | **2** Altmetric | [Metrics](#)

Abstract

In vivo and vitro evidence indicates that ornithine and its related metabolic products play a role in tumor development. Whether ornithine is associated with breast cancer in humans is still unclear. We examined the association between circulating ornithine levels and breast cancer in females. This 1:1 age-matched case–control study identified 735 female breast cancer cases and 735 female controls without breast cancer. All cases had a pathological test to ascertain a breast cancer diagnosis. The controls were ascertained using pathologic testing, clinical examinations, and/or other tests. Fasting blood samples were used to measure ornithine levels. The average age for cases and controls were 49.6 years (standard deviation [SD] 8.7 years) and 48.9 years (SD 8.7 years), respectively. Each SD increase in

[Zhang et al., 2020](#)

【HTML】

scientific reports

Explore our content ▾ Journal information ▾

nature > scientific reports > articles > article

Article | [Open Access](#) | Published: 23 September 2020

Ornithine and breast cancer: a matched case–control study

Jiayi Zhang, Baihui Tao, Yiran Chong, Shuang Ma, Gang Wu, Hailong Zhu, Yi Zhao, Shitao Zhao, Mengmeng Niu, Shutian Zhang, Tianyi Wang, Shuman Yang, Wenjing Qiao, Ann M. Vuong, Jincheng Li, Demiao Zhu & Wei Tao 

Scientific Reports **10**, Article number: 15502 (2020) | [Cite this article](#)

109 Accesses | **2** Altmetric | [Metrics](#)

Abstract

In vivo and vitro evidence indicates that ornithine and its related metabolic products play a role in tumor development. Whether ornithine is associated with breast cancer in humans is still unclear. We examined the association between circulating ornithine levels and breast cancer in females. This 1:1 age-matched case–control study identified 735 female breast cancer cases and 735 female controls without breast cancer. All cases had a pathological test to ascertain a breast cancer diagnosis. The controls were ascertained using pathologic testing, clinical examinations, and/or other tests. Fasting blood samples were used to measure ornithine levels. The average age for cases and controls were 49.6 years (standard deviation [SD] 8.7 years) and 48.9 years (SD 8.7 years), respectively. Each SD increase in

[Zhang et al., 2020](#)

【HTML】

- テキスト操作可能
 - 検索
 - コピー

scientific reports

Explore our content ▾ Journal information ▾

nature > scientific reports > articles > article

Article | [Open Access](#) | Published: 23 September 2020

Ornithine and breast cancer: a matched case–control study

Jiayi Zhang, Baihui Tao, Yiran Chong, Shuang Ma, Gang Wu, Hailong Zhu, Yi Zhao, Shitao Zhao, Mengmeng Niu, Shutian Zhang, Tianyi Wang, Shuman Yang, Wenjing Qiao, Ann M. Vuong, Jincheng Li, Demiao Zhu & Wei Tao [✉](#)

Scientific Reports **10**, Article number: 15502 (2020) | [Cite this article](#)

109 Accesses | **2** Altmetric | [Metrics](#)

Abstract

In vivo and vitro evidence indicates that ornithine and its related metabolic products play a role in tumor development. Whether ornithine is associated with breast cancer in humans is still unclear. We examined the association between circulating ornithine levels and breast cancer in females. This 1:1 age-matched case–control study identified 735 female breast cancer cases and 735 female controls without breast cancer. All cases had a pathological test to ascertain a breast cancer diagnosis. The controls were ascertained using pathologic testing, clinical examinations, and/or other tests. Fasting blood samples were used to measure ornithine levels. The average age for cases and controls were 49.6 years (standard deviation [SD] 8.7 years) and 48.9 years (SD 8.7 years), respectively. Each SD increase in

[Zhang et al., 2020](#)

【HTML】

○ テキスト操作可能

- 検索
- コピー

x 構造化

- 出版社ごとにフォーマットがバラバラ
- 本文は非構造化データ

scientific reports

Explore our content ▾ Journal information ▾

nature > scientific reports > articles > article

Article | [Open Access](#) | Published: 23 September 2020

Ornithine and breast cancer: a matched case–control study

Jiayi Zhang, Baihui Tao, Yiran Chong, Shuang Ma, Gang Wu, Hailong Zhu, Yi Zhao, Shitao Zhao, Mengmeng Niu, Shutian Zhang, Tianyi Wang, Shuman Yang, Wenjing Qiao, Ann M. Vuong, Jincheng Li, Demiao Zhu & Wei Tao [✉](#)

Scientific Reports **10**, Article number: 15502 (2020) | [Cite this article](#)

109 Accesses | **2** Altmetric | [Metrics](#)

Abstract

In vivo and vitro evidence indicates that ornithine and its related metabolic products play a role in tumor development. Whether ornithine is associated with breast cancer in humans is still unclear. We examined the association between circulating ornithine levels and breast cancer in females. This 1:1 age-matched case–control study identified 735 female breast cancer cases and 735 female controls without breast cancer. All cases had a pathological test to ascertain a breast cancer diagnosis. The controls were ascertained using pathologic testing, clinical examinations, and/or other tests. Fasting blood samples were used to measure ornithine levels. The average age for cases and controls were 49.6 years (standard deviation [SD] 8.7 years) and 48.9 years (SD 8.7 years), respectively. Each SD increase in

[Zhang et al., 2020](#)

【HTML】

- テキスト操作可能
 - 検索
 - コピー
- x 構造化
 - 出版社ごとにフォーマットがバラバラ
 - 本文は非構造化データ
- x 正規化
 - 表記揺れ
 - 同音異義語

scientific reports

Explore our content ▾ Journal information ▾

nature > scientific reports > articles > article

Article | [Open Access](#) | Published: 23 September 2020

Ornithine and breast cancer: a matched case–control study

Jiayi Zhang, Baihui Tao, Yiran Chong, Shuang Ma, Gang Wu, Hailong Zhu, Yi Zhao, Shitao Zhao, Mengmeng Niu, Shutian Zhang, Tianyi Wang, Shuman Yang, Wenjing Qiao, Ann M. Vuong, Jincheng Li, Demiao Zhu & Wei Tao [✉](#)

Scientific Reports **10**, Article number: 15502 (2020) | [Cite this article](#)

109 Accesses | **2** Altmetric | [Metrics](#)

Abstract

In vivo and vitro evidence indicates that ornithine and its related metabolic products play a role in tumor development. Whether ornithine is associated with breast cancer in humans is still unclear. We examined the association between circulating ornithine levels and breast cancer in females. This 1:1 age-matched case–control study identified 735 female breast cancer cases and 735 female controls without breast cancer. All cases had a pathological test to ascertain a breast cancer diagnosis. The controls were ascertained using pathologic testing, clinical examinations, and/or other tests. Fasting blood samples were used to measure ornithine levels. The average age for cases and controls were 49.6 years (standard deviation [SD] 8.7 years) and 48.9 years (SD 8.7 years), respectively. Each SD increase in

[Zhang et al., 2020](#)

【HTML】

- テキスト操作可能
 - 検索
 - コピー
- x 構造化
 - 出版社ごとにフォーマットがバラバラ
 - 本文は非構造化データ
- x 正規化
 - 表記揺れ
 - 同音異義語

ヒトには解釈しやすいが、
コンピューターには解釈し辛い...

scientific reports

Explore our content ▾ Journal information ▾

nature > scientific reports > articles > article

Article | Open Access | Published: 22 September 2020

【HTML】

- テキスト操作可能

! 電子化すりゃいいってわけじゃない! !

in vivo and *in vitro* evidence indicates that ornithine and its related metabolic products play a role in tumor development. Whether ornithine is associated with breast cancer in humans is still unclear. We examined the association between circulating ornithine levels and breast cancer in females. This 1:1 age-matched case-control study identified 735 female breast cancer cases and 735 female controls without breast cancer. All cases had a pathological test to ascertain a breast cancer diagnosis. The controls were ascertained using pathologic testing, clinical examinations, and/or other tests. Fasting blood samples were used to measure ornithine levels. The average age for cases and controls were 49.6 years (standard deviation [SD] 8.7 years) and 48.9 years (SD 8.7 years), respectively. Each SD increase in

[Zhang et al., 2020](#)

- 表記揺れ
- 同音異義語

ヒトには解釈しやすいが、
コンピューターには解釈し辛い...

紙の記憶に囚われている！！
ゼロベースでコンピューターにとって
優しい形を考えよう



スペースに制限がある



（紙に比べたら）制限はないに等しい
生データ全部掲載することも可能

紙の記憶に囚われている！！
ゼロベースでコンピューターにとって
優しい形を考えよう



人間に読みやすいことが大事



一次データを構造的にして、
加工して人間可読にできる

人間可読 → 機械可読

人間が読む分には問題ないけど、機械が扱うには問題がある

... なんか聞いたことある🤔

人間可読 → 機械可読

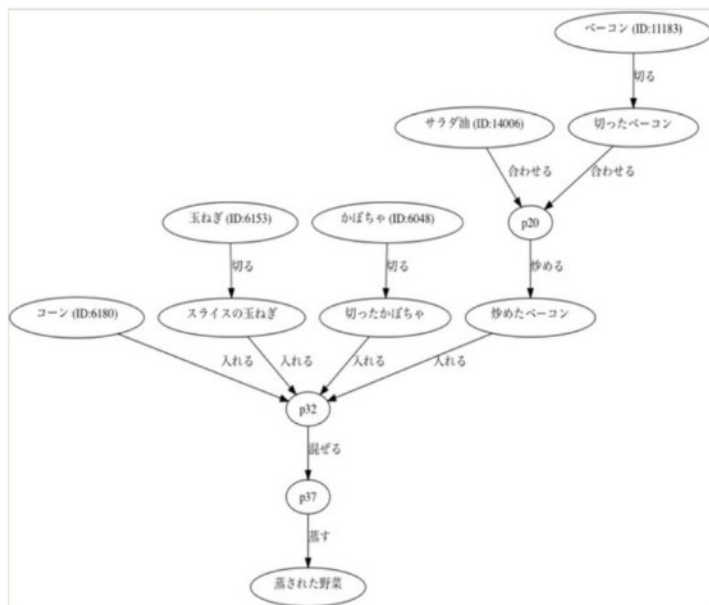
人間が読む分には問題ないけど、機械が扱うには問題がある

... なんか聞いたことある 🤔

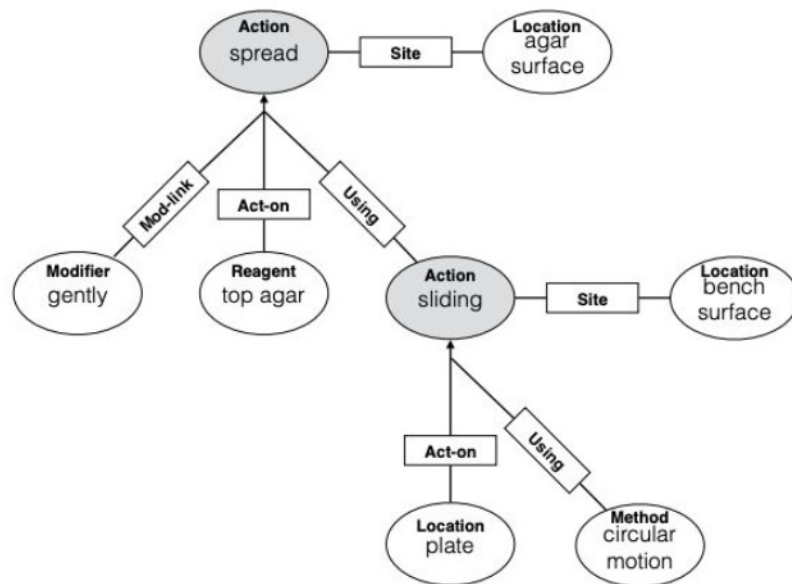
Semantic Webと同じ！

ドキュメントのWeb → データのWeb

コンピューターが解釈できる知識形態 ～非定型されど構造的～



[MRR \(Machine Readable Recipe\) の開発 / Development of MRR @ CookpadTechConf2019](#)



[Kulkarni et al., 2018](#)

- ①文脈を考慮した検索ができる
- ②新しいプロトコルの生成ができる

文献調査を 広く、早く。

研究支援サービス Sophiscope

ライフサイエンス研究者の文献調査の課題を解決し、
研究を加速させるプラットフォーム。

β版トライアル申し込み >

Cell Line-Derived Xenograft

Filter

Sort

4 hidden fields

19941Hits

Bibliography

Title

Date

1

Dual targeting of BCL2 and MCL...

2020-05-01

2

EGFR-targeted CAR-T cells are p...

2020-05-01

3

EGFR-targeted CAR-T cells are p...

2020-05-01

4

EGFR-targeted CAR-T cells are p...

2020-05-01

5

EGFR-targeted CAR-T cells are p...

2020-05-01

6

EGFR-targeted CAR-T cells are p...

2020-05-01

7

EGFR-targeted CAR-T cells are p...

2020-05-01

8

IMET726 suppresses human ski...

2020-05-01

9

The Histone Deacetylase Inhibito...

2020-05-01

10

The Histone Deacetylase Inhibito...

2020-05-01

11

The Histone Deacetylase Inhibito...

2020-05-01

12

Combined inhibition of PIM and C...

2020-04-28

13

Hydrogen gas represses the prog...

2020-04-28

14

MIF600HG suppresses metastasi...

2020-04-28

15

SLC8A14, a Na+/Cl-coupled am...

2020-04-23

16

SLC8A14, a Na+/Cl-coupled am...

2020-04-23

17

Tripartite motif containing 35 cont...

2020-04-23

Animal

Species

Strain

Sex

Age

mouse

NSG

female

6 weeks old

mouse

SCID

female

6-12 weeks old

mouse

SCID

female

6-12 weeks old

mouse

SCID

female

6-12 weeks old

mouse

SCID

female

6-12 weeks old

mouse

SCID

female

6-12 weeks old

mouse

SCID

female

6-12 weeks old

mouse

SCID

female

6-7 weeks old

mouse

CD-1

—

7 weeks old

mouse

Nude

—

7 weeks old

mouse

CD-1

female

6-7 weeks old

mouse

Fox Chase SCID

female

—

mouse

BALB/c Nude

male

—

mouse

Athymic Nude

female

6 weeks old

mouse

Nude

male

—

mouse

C57BL/6

male

—

mouse

BALB/c

female

4 weeks old

Transplantation

Cancer type

Cell line

Plasma cell myeloma

U266

Breast adenocarcinoma

MDA-MB-231

Breast adenocarcinoma

MDA-MB-231

Breast adenocarcinoma

MDA-MB-231

Breast adenocarcinoma

MDA-MB-231

Invasive ductal carcinoma, not otherwise...

MCF-7

Invasive ductal carcinoma, not otherwise...

MCF-7

Vulvar squamous cell carcinoma

A-431

Bladder carcinoma

RT-112

Bladder carcinoma

RT-112

Bladder carcinoma

RT-112

Lung adenocarcinoma

A-549

Lung adenocarcinoma

A-549

Colon adenocarcinoma

Caco2

Colon adenocarcinoma

LS173T

Colon adenocarcinoma

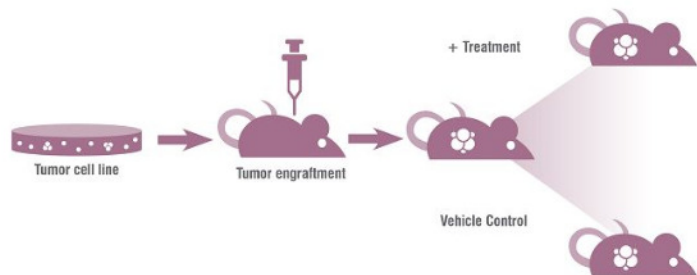
LS173T

Large cell lung carcinoma

NCI-H460

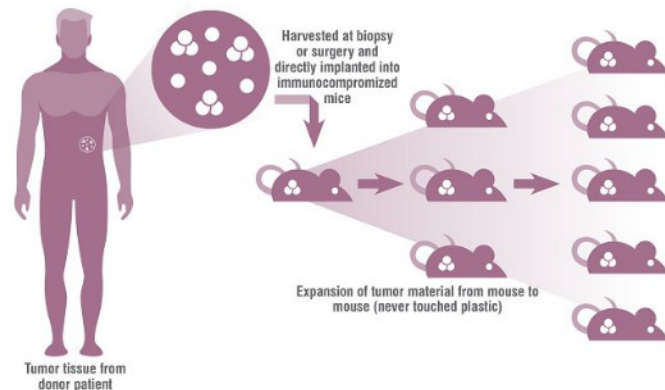
Sophi scopeの対応状況

Cell Line Derived Xenograft (CDX) Model



[Cell Line-Derived Xenograft – CDX Model Studies in Mice](#)

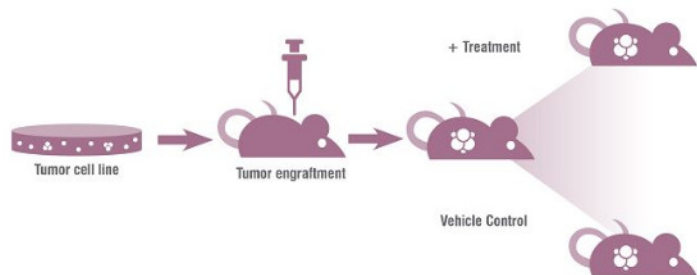
Patient Derived Xenografts (PDX)



[Patient-Derived Xenografts - PDX Models](#)

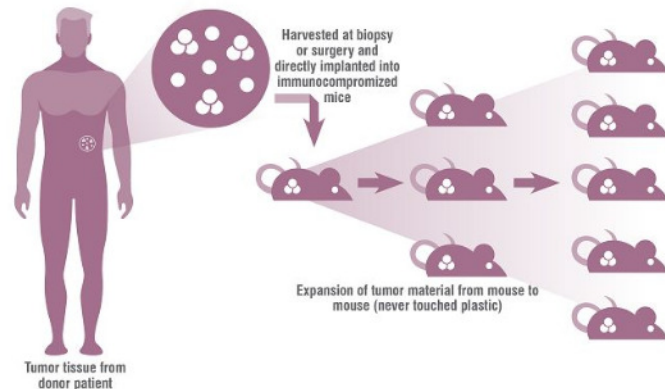
Sophi scopeの対応状況

Cell Line Derived Xenograft (CDX) Model



[Cell Line-Derived Xenograft – CDX Model Studies in Mice](#)

Patient Derived Xenografts (PDX)



[Patient-Derived Xenografts - PDX Models](#)

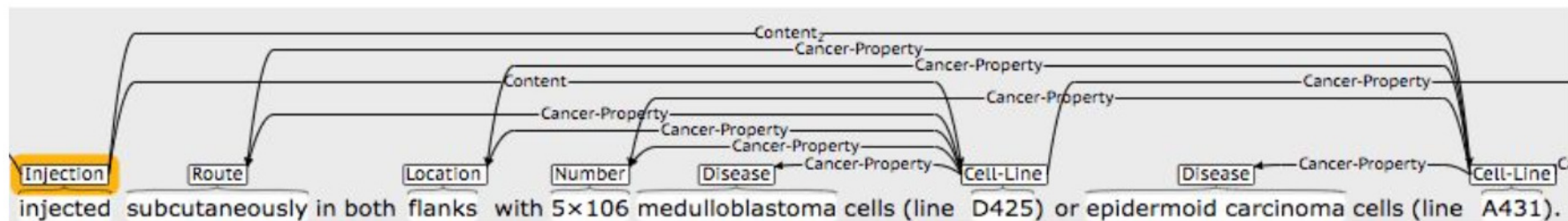
糖尿病・認知症・AIDSなどに領域に展開予定
共同研究、PoCお待ちしております！！

失敗と対策

失敗①

あらゆる固有表現・関係性をアノテーションしようとした

→コストが甚大 & 実用レベルに至っていない



brat

対策①

優先度の高い項目を選別、マニュアル化、自動化

→ (ある程度) 誰でもできるようにした

	A Title	A PMCID	A URL	Is CDX	A Strain	Strain norm	A Sex	Sex no.	A Age	Age norm	A Cell line	Cell line norm
1	Claudin-1 regulates cellular trans...	PMC1150288	https://www.ncbi.nlm...	✓	athymic nude	Athymic Nude	female	female	6-week-old	6 weeks old		不明
2	Claudin-1 regulates cellular trans...	PMC1150288	https://www.ncbi.nlm...	✓	athymic/nude	Athymic Nude	female	female	6-week-old	6 weeks old	SW480	SW480
3	Claudin-1 regulates cellular trans...	PMC1150288	https://www.ncbi.nlm...	✓	athymic/nude	Athymic Nude	female	female	6-week-old	6 weeks old	SW620	SW620
4	Newly identified biologically acti...	PMC2140032	https://www.ncbi.nlm...	✓	'nude', 'Swiss...	Swiss Nude			6-wk-old	6 weeks old	HEK293	HEK293
5	Newly identified biologically acti...	PMC2140032	https://www.ncbi.nlm...	✓	Nude	Nude			6-wk-old	6 weeks old	MCF-7	MCF-7
6	Targeting EGFR with metabolical...	PMC2203207	https://www.ncbi.nlm...	✓	CB17 SCID	Fox Chase S...	female	female	6-8 weeks o...	6-8 weeks old	SKOV3ip1	SKOV3.ip1
7	Distinct Angiogenic Mediators Ar...	PMC2366879	https://www.ncbi.nlm...	✓	SCID/Beige	Fox Chase S...	female	female	4-wk-old	4 weeks old	MDA-MB-231	MDA-MB-231
8	MATRIX METALLOPROTEINASE-...	PMC2394502	https://www.ncbi.nlm...	✓	nude	Nude						不明
9	MATRIX METALLOPROTEINASE-...	PMC2394502	https://www.ncbi.nlm...	✓	C57BL/6	C57BL/6	Female	female			B16-BL6	B16-BL6
10	Sorafenib inhibits STAT3 signalin...	PMC2592687	https://www.ncbi.nlm...	✓	nude	Nude			4- to 6-wee...	4-6 weeks old	amb_Daoy	Daoy
11	P21/Cip1 and P27/Kip1 are Essen...	PMC2633420	https://www.ncbi.nlm...	✓	athymic nu/nu	Athymic Nude	male	male	6-week-old	6 weeks old	DU145	DU145
12	Acquired resistance to 17-allylam...	PMC2652695	https://www.ncbi.nlm...	✓	NCr athymic	NCr Nude	female	female				不明

Airtable

失敗②

SOTAを追い続ける日々

→「技術の検証はできました！あとはデータさえ集めれば完成します！」

Table 6. Test results in biomedical named entity recognition

Type	Datasets	Metrics	BERT		BioBERT v1.0			BioBERT v1.1	
			SOTA	(Wiki + Books)	(+ PubMed)	(+ PMC)	(+ PubMed + PMC)	(+ PubMed)	
Disease	NCBI disease	P	<u>88.30</u>	84.12	86.76	86.16	<u>89.04</u>	88.22	
		R	89.00	87.19	88.02	89.48	<u>89.69</u>	91.25	
		F	88.60	85.63	87.38	87.79	<u>89.36</u>	89.71	
	2010 i2b2/VA	P	<u>87.44</u>	84.04	85.37	85.55	87.50	86.93	
		R	<u>86.25</u>	84.08	85.64	85.72	85.44	86.53	
		F	<u>86.84</u>	84.06	85.51	85.64	86.46	<u>86.73</u>	
	BC5CDR	P	<u>89.61</u>	81.97	85.80	84.67	85.86	<u>86.47</u>	
		R	83.09	82.48	86.60	85.87	<u>87.27</u>	87.84	
		F	<u>86.23</u>	82.41	86.20	85.27	86.56	87.15	
Drug/chem.	BC5CDR	P	<u>94.26</u>	90.94	92.52	92.46	93.27	<u>93.68</u>	
		R	92.38	91.38	92.76	92.63	<u>93.61</u>	<u>93.26</u>	
		F	93.31	91.16	92.64	92.54	<u>93.44</u>	93.47	
	BC4CHEMD	P	<u>92.29</u>	91.19	91.77	91.65	92.23	92.80	
		R	90.01	88.92	<u>90.77</u>	90.30	90.61	91.92	
		F	91.14	90.04	91.26	90.97	<u>91.41</u>	92.36	
	Gene/protein	BC2GM	P	81.81	81.17	81.72	82.86	<u>85.16</u>	<u>84.32</u>
			R	81.57	82.42	83.38	<u>84.21</u>	83.65	85.12
			F	81.69	81.79	82.54	83.53	<u>84.40</u>	<u>84.72</u>
JNLPBA		P	74.43	69.57	71.11	71.17	<u>72.68</u>	72.24	
		R	<u>83.22</u>	81.20	83.11	82.76	83.21	83.56	
		F	78.58	74.94	76.65	76.53	<u>77.59</u>	77.49	
Species		LINNAEUS	P	<u>92.80</u>	91.17	91.83	91.62	93.84	90.77
			R	<u>94.29</u>	84.30	84.72	85.48	<u>86.11</u>	85.83
			F	93.54	87.60	88.13	88.45	<u>89.81</u>	88.24
	Species-800	P	<u>74.34</u>	69.35	70.60	71.54	<u>72.84</u>	72.80	
		R	<u>75.96</u>	74.05	75.75	74.71	<u>77.97</u>	75.36	
		F	<u>74.98</u>	71.63	73.08	73.09	75.31	74.06	

対策②

データの数が少ない段階から機能する手法に切り替え

→～1,000データでも検証に足る性能

急がば回るな！

効率よくやろうとすると永遠に完成しない...



知識の記述形態を改革する仲間を募集中！

エンジニア・営業を採用中！

<https://fuku-inc.com>