Aus der Kinderklinik und Kinderpoliklinik
im Dr. von Haunerschen Kinderspital
Klinik der Universität München
Direktor: Prof. Dr. Dr. Christoph Klein

Die Detektion von Proteininteraktionen mittels iBRET als Grundlage systemmedizinischer Analysen von Pathomechanismen bei genetischen Erkrankungen

Dissertation
zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von
Mathias Woidy
aus
Landshut

2023

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter:	Prof. Dr. Ania C. Muntau
Mitberichterstatter:	PD Dr. Dr. Albrecht von Brunn
	Prof. Dr. Axel Imhof
	Prof. Dr. Frederick Klauschen
Mitbetreuung durch den promovierten Mitarbeiter:	Prof. Dr. Søren W. Gersting
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	05.10.2023

Affidavit

Eidesstattliche Versicherung

Woidy, Mathias

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Die Detektion von Proteininteraktionen mittels iBRET als Grundlage systemmedizinischer Analysen von Pathomechanismen bei genetischen Erkrankungen

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Hamburg, 29.01.2023 — Ort, Datum

Mathias Woidy — Unterschrift Doktorandin bzw. Doktorand

Inhaltsverzeichnis

Publikationsliste

Folgende veröffentlichte Originalarbeiten sind Bestandteil dieser kumulativen Promotionsarbeit.

Veröffentlichung I:

Inborn errors of metabolism and the human interactome: a systems medicine approach.

Woidy M, Muntau AC & Gersting SW

Journal of Inherited Metabolic Disease**,** 2018 41:285-296.

Veröffentlichung II:

iBRET screen of the ABCD1 peroxisomal network and mutation-induced network perturbations.

Lotz-Havla, AS*, **Woidy M*,** Guder P*, Friedel CC, Klingbeil JM, Bulau, A, Schultze A, Dahmen I, Noll-Puchta H, Kemp S, Erdmann R, Zimmer R, Muntau AC, Gersting SW

Journal of Proteome Research, 2021 20(9): 4366-4380

*geteilte Erstautorenschaft

Weitere Arbeiten, die im Rahmen meiner wissenschaftlichen Tätigkeit als Promovend mit Hilfe der hier entwickelten Methode entstanden sind, jedoch nicht Teil der kumulativen Promotionsleistung sind:

Edgetic perturbations contribute to phenotypic variability in PEX26 deficiency

Lotz-Havla AS, **Woidy M**, Guder P, Schmiesing J, Erdmann R, Waterham HR, Muntau AC, Gersting SW

Frontiers in Genetics, 2021 12:726174

Isoform-specific domain-organization determines conformation and function of the peroxisomal biogenesis factor PEX26

Guder P, Lotz-Havla AS, **Woidy M**, Reiß DD, Danecka MK, Schatz UA, Becker M, Ensenauer R, Pagel P, Büttner L, Muntau AC, Gersting SW

Biochimica et Biophysica Acta – Molecular Cell Research, 2018 1866(3): 518-531

Varicella zoster virus ORF25 gene product: an essential hub protein linking encapsidation proteins and the nuclear egress complex.

Vizoso Pinto MG, Pothineni VR, Haase R, **Woidy M**, Lotz-Havla AS, Gersting SW, Muntau AC, Haas J, Sommer M, Arvin AM, and Baiker A

Journal of Proteome Research, 2011 10(12): 5374–5382

Weitere wissenschaftliche Arbeiten mit meiner Beteiligung ohne direkten Bezug zur kumulativen Promotionsarbeit:

Long-Term Antibody Response to SARS-CoV-2 in Children

Dunay GA*, Barroso M*, **Woidy M*,** Danecka MK, Engels G, Hermann K, Neumann FS, Paul K, Beime J, Escherich G, Fehse K, Grinstein L, Haniel F, Haupt LJ, Hecher L, Kehl T, Kemen C, Kemper KJ, Kobbe R, Kohl A, Klokow T, Nörz D, Olfe J, Schlenker F, Schmiesing J, Schrum J, Sibbertsen F, Stock P, Tiede S, Vettorazzi E, Zazara DE, Zapf A, Lütgehetmann M, Oh J, Mir TS, Muntau AC, Gersting SW & the C19.CHILD Study Group

Journal of Clinical Immunology 2022 https://doi.org/10.1007/s10875-022-01355-w

Specific CD4+ T Cell Responses to Ancestral SARS-CoV-2 in Children Increase With Age and Show Cross-Reactivity to Beta Variant.

Paul K, Sibbertsen F, Weiskopf D, Lütgehetmann M, Barroso M, Danecka MK, Glau L, Hecher L, Hermann K, Kohl A, Oh J, Schulze zur Wiesch J, Sette A, Tolosa E, Vettorazzi E, **Woidy M**, Zapf A, Zazara DE, Mir TS, Muntau AC, Gersting SW and Dunay GA (2022)

Front. Immunol., 2022 13:867577

Dominant KPNA3 mutations cause infantile onset hereditary spastic paraplegia.

Schob C, Hempel M, Brozkova DS, Jiang H, Kim SY, Batzir NA, Orenstein N, Bierhals T, Johannsen J, Meszarosova AU, Chae JH, Seeman P, **Woidy M**, Fang F, Kubisch C, Kindler S, Denecke J

Annals of Neurology, 2021 90:738-750

Biallelic MADD variants cause a phenotypic spectrum ranging from developmental delay to a multisystem disorder

Schneeberger P, Kortüm F, Korenke GC, Alawi M, Santer R, **Woidy M**, Buhas D, Fox S, Juusola J, Alfadhel M, Webb BD, Coci EG, Jamra RA, Siekmeyer M, Biskup S, Heller C, Maier EM, Javaher-Haghigh P, Bedeschi MF, Ajmone PF, Iascone M, Peeters H, Ballon K, Jaeken J, Alonso RA, Palomares-Bralo M, Santos-Simarro F, Meuwissen MEC, Beysen D, Kooy RF, Houlden H, Murphy D, Doosti M, Karimiani EG, Mojarrad M, Maroofian R, Noskova L, Kmoch S, Honzik T, Cope H, Sanchez-Valle A, Undiagnosed Diseases Network, Gelb BD, Kurth I, Hempel M and Kutsche K

Brain, 2020 143(8): 2437-2453

Secondary BH4 deficiency links protein homeostasis to regulation of phenylalanine metabolism

Eichinger A, Danecka MK, Möglich T, Borsch J, **Woidy M,** Büttner L, Muntau AC and Gersting SW

Human Molecular Genetics, 2018 27(10): 1732-1742.

Mapping the functional landscape of frequent phenylalanine hydroxylase genotypes promotes personalised medicine in phenylketonuria.

Danecka MK, **Woidy M**, Zschocke J, Feillet F, Muntau AC, Gersting SW

Journal of Medical Genetics, 2015 52 (3): 175-185.

The interplay between genotype, metabolic state and cofactor treatment governs phenylalanine hydroxylase function and drug response.

Staudig M, Gersting SW, Danecka MK, Messing DD, **Woidy M**, Pinkas D, Kemter FK, Blau N, and Muntau AC

Human Molecular Genetics, 2011 20 (30): 2628-2641.

Beitrag zu den Veröffentlichungen

1.1 Beitrag zu Veröffentlichung I, Inborn errors of metabolism and the human interactome: a systems medicine approach.

Ziel dieser Studie war es herauszufinden, ob sich systemmedizinische Ansätze auf angeborene Stoffwechselerkrankungen (IEM) anwenden lassen. IEM bilden klinisch eine sehr heterogene Gruppe an Erkrankungen, deren Gemeinsamkeit Störungen in metabolischen Prozessen der Zellen sind. Diese Erkrankungsgruppe lässt sich nach verschiedenen Kriterien in Untergruppen einteilen. Eine Möglichkeit bietet die Society for the Study of Inborn Errors of Metabolism (SSIEM, https://www.ssiem.org/resources/resources/inborn-errors-classification). Sie listet mehr als 600 IEM und ihre assoziierten Proteine in 15 Untergruppen. Wir stellten uns zunächst die Frage, ob diese Proteine zufällig im humanen Interaktom, also der Gesamtheit aller Protein-Protein Interaktionen, verteilt sind, oder ob diese ein *disease module* bilden. Als *disease module* bezeichnet man ein graphentheoretisches Modell, das die Kolokalisation von krankheitsassoziierten Proteinen im humanem Interaktom beschreibt. Hierfür haben wir einen Datensatz eines humanen Interaktom mit insgesamt 13.460 Proteinen und 141.296 Interaktionen verwendet. Dieser Datensatz beruht auf experimentell nachgewiesenen Proteininteraktionen und wurde 2015 von Menche et al. veröffentlich (Menche *et al.*, 2015). Wir fanden darin für 427 der über 600 IEM assoziierten Proteine experimentell bestätigte Interaktionen. Wir konnten zeigen, dass IEM assoziierte Proteine (i) nicht zufällig im humanen Interaktom verteilt sind, sondern (ii) in gleicher Nachbarschaft liegen und (iii) ein großes *disease module* bilden. Außerdem bilden auch die einzelnen IEM-Untergruppen *disease module*, die teilweise stark, teilweise weniger stark miteinander verbunden sind. Als nächstes wurden die Verbindungen der IEM zu anderen Erkrankungsgruppen im humanen Interaktom untersucht, um Rückschlüsse über gemeinsame Pathomechanismen zu erlangen. Dies gelang über die Berechnung der Distanzen von IEM assoziierten Proteinen zu mit anderen Erkrankungsgruppen assoziierten Proteinen im humanen Interaktom. Hier ergaben sich mehrere Assoziationen zu anderen Erkrankungen, wie zum Beispiel zur Gruppe der Alzheimererkrankungen. Enge Beziehungen von Erkrankungsgruppen im Interaktom lassen auf gemeinsame Pathomechanismen schließen, so werden auch für Alzheimer und IEM in der wissenschaftlichen Literatur gemeinsame Pathomechanismen diskutiert und bestätigen unsere Ergebnisse (Barbero-Camps *et al.*, 2014; Mercer, Wang and Burke, 2017). Die Integration zusätzlicher Daten über Medikamente und deren Zielproteine stellt eine Möglichkeit des *drug-repurposing* dar. Wir haben daher diese Daten in unser Netzwerk integriert, nach Zielproteinen gesucht, die direkt mit IEM assoziierten Proteinen interagieren und Anwendungsmöglichkeiten hierfür aufgezeigt.

Diese Studie zeigt erstmals die Anwendung von Netzwerkanalysen für IEM und soll als Grundlage für weitere Netzwerk-basierte Analysen im Feld der angeborenen Stoffwechselerkrankungen dienen.

Eigener Beitrag zum Artikel

Sämtliche *in-silico* Analysen dieses Artikels wurden von mir durchgeführt. Hierbei wurden Skripte zur automatischen Datenverarbeitung und Anwendung von netzwerkmedizinischen Algorithmen programmiert, sowie bereits bestehende und öffentlich zur Verfügung gestellte Skripte entsprechend modifiziert. Als Datengrundlage wurde die von der SSIEM zur Verfügung gestellte Einteilung angeborener Stoffwechselerkrankungen verwendet. Mittels automatischer OMIM Abfrage über die zur Verfügung gestellte API (application programming interface) konnten Proteinnamen und assoziierte Erkrankungen abgeglichen werden. Als nächster Schritt erfolgte die Suche nach Interaktionen für IEM-assoziierte Proteine im von Menche et al. zur Verfügung gestellten Interaktionsdatensatz. Dies geschah mit Hilfe von mir hierfür programmierter *Perl* Skripte. Zur Berechnung signifikanter *disease module* wurden die von Menche et al. bereitgestellten *Python* Skripte von mir angepasst, ebenso zur Berechnung der Distanzen zwischen IEM-assoziierten Proteinen und denen, die mit anderen Erkrankungen assoziiert sind. Die Integration der Daten zu Medikamenten und deren Zielproteine erfolgte mittels von mir dafür programmiertem *Perl* Skript, mit welchem die Daten aus einer Datenbank (drugbank) geladen und integriert wurden (Law *et al.*, 2014). Zur Veranschaulichung der Daten wurden gängige Netzwerkdarstellungen mittels der Software *Cytoscape* erzeugt (https://cytoscape.org). Die Ergebnisse wurden umfangreich diskutiert und führten letztendlich zur Publikation.

1.2 Beitrag zu Veröffentlichung II, iBRET screen of the ABCD1 peroxisomal network and mutation-induced network perturbations.

Ziel dieser Arbeit war es, BRET als empfindliche, automatisierte, informatik-gestützte Methode zur Untersuchung von Proteininteraktionen in lebenden Zellen zu etablieren (iBRET) und systematisch zu evaluieren. Der Nutzen von iBRET sollte an einem Anwendungsszenario demonstriert werden. Zunächst wurde der BRET-Assay mittels einer Tecan-Arbeitsplattform sowie bioinformatischen Methoden im 96-well Format automatisiert (iBRET). Mittels eines Referenzdatensatzes, der aus bekannt positiven (n=61) und zufällig ausgewählten nicht interagierenden Proteinpaaren besteht (n=60), wurde iBRET kalibriert und hinsichtlich seiner Leistungsfähigkeit validiert. BRET im Allgemeinen beruht auf der Detektion von Lumineszenz-Signalen durch Donor- und Akzeptorproteine, die an die zu untersuchenden Proteine angeheftet sind. In der Literatur wird häufig nur eine rekombinante Konfiguration getestet (z. B. N-Terminal angeheftete Donor- und Akzeptorproteine an die zu untersuchenden Proteine). Diese rekombinante Anheftung kann aber zur Unterdrückung von Proteininteraktionen, z. B. durch Blockade der Bindestellen, führen und wir konnten zeigen, dass eine Testung aller 8 möglichen Kombinationen, zwar zu einem erhöhten experimentellen Aufwand, jedoch auch zu einer deutlich verbesserten Sensitivität und Spezifität von iBRET führt. Die Klassifikation der gemessenen Proteinpaare erfolgt automatisiert und ermöglicht die

Einteilung in „interagierend“, „nicht-interagierend“, sowie in einen Graubereich. Proteinpaare, deren Messergebnis im Graubereich liegen, können mittels zusätzlichem BRET-Evaluationsexperiment klassifiziert werden. Wir konnten für iBRET eine True-Positive-Rate (TPR) von 58% und eine False-Positive-Rate (FPR) von 4% zeigen. Nimmt man die Ergebnisse der Evaluationsexperimente hinzu, so fanden wir eine Steigerung der TPR auf 72% bei einer zusätzlichen Erhöhung der FPR auf 12%.

Die Arbeit beschreibt zusätzlich die Anwendung von iBRET am Beispiel des Erkrankungsgens ABCD1. Dieses kodiert für einen peroxisomalen Transporter. Ein Defekt ist mit einer schweren Form der Adrenoleukodystrophie assoziiert (X-ALD). Die genauen Pathomechanismen sind bisher nicht eingehend verstanden. Der mittels iBRET durchgeführte Interaktionsscreen gegen 104 peroxisomale Proteine fand 9 neue Interaktionspartner. Um die biologische Relevanz dieser Interaktionen zu untersuchen, wurden 2 gefundene Interaktionen (PEX5, FAR1) gegen unterschiedliche ABCD1-Varianten getestet und zusätzliche Expressionsdaten aus einer Datenbank integriert (Human-Protein-Atlas). So zeigten unterschiedliche ABCD1-Varianten verschieden starke Störungen der Interaktion mit FAR1, die von einer reduzierten relativen Bindungsaffinität bis zu einem kompletten Ausfall der Interaktion reichen. Ein Phänomen welches als *edgetic perturbation* beschrieben ist und mit einer damit einhergehenden Pathogenität in Verbindung gebracht wird (Sahni *et al.*, 2013).

Zusammenfassend konnten wir zeigen, dass iBRET als zusätzliche hoch-sensitive Methode zur Untersuchung von Proteininteraktionen im mittleren Durchsatz (n=640 Proteinpaare pro Woche) eingesetzt werden kann. Die hohe Sensitivität von iBRET ist (i) auf die Detektion in lebenden Zellen, (ii) die Untersuchung aller 8-Kombinationen, sowie (iii) dem automatisierten informatikgestützten Setup zurückzuführen.

Eigener Beitrag zum Artikel

Ich war maßgeblich an der Etablierung des BRET-Assays auf einer Tecan-Arbeitsplattform beteiligt. Der komplette in der Arbeitsgruppe vorhandene BRET-Assay im Kuvettenformat wurde im 96-well Format auf dieser Arbeitsplattform automatisiert. Hierzu mussten ein 96-Well Elektroporationsgerät sowie ein 96-well Lumineszenz-Lesegerät in die Tecan-Arbeitsplattform integriert und die Pippetierabläufe implementiert werden. Zusätzlich war ich an den Programmierarbeiten zur Etablierung einer durch die Arbeitsplattform ansprechbaren Datenbank beteiligt. Alle notwendigen Arbeitsschritte – außer vorbereitender Zellkulturarbeiten – können somit automatisiert ablaufen. Die Validierung sowie Analyse der Leistungsfähigkeit des Assays erfolgte mit Hilfe eines Referenzdatensatzes. Die Zusammenstellung dieses Datensatzes und die experimentelle Vermessung habe ich zusammen mit A.S. Lotz-Havla durchgeführt. Der Assay wurde mehrfach optimiert und validiert insbesondere zur Erhöhung der Durchsatzzahlen z.B. durch Minimierung von Replikaten oder zur Verbesserung der Genauigkeit der einzelnen Pippetierschritte. Hieran war ich ebenfalls maßgeblich beteiligt. Zellkulturarbeiten wurden hauptsächlich durch A. Schultze, I. Dahmen und H. Noll-Puchta durchgeführt. Die Auswertung und statistische Validierung erfolgte durch mich in Zusammenarbeit mit C. Friedel und R. Zimmer des Instituts für Informatik der LMU. Ein nächster Schritt war die Anwendung von iBRET zur Untersuchung des peroxisomalen

ABCD1-Interaktionsnetzwerkes zusammen mit A.S. Lotz-Havla und P. Guder. Hier war ich erneut an der experimentellen Vermessung des Netzwerkes beteiligt. Die für die Wiederholungsexperimente der neu gefundenen ABCD1-Proteininteraktionen verwendeten Assays (BiFC und LUMIER) wurden durch P. Guder und A.S. Lotz-Havla etabliert. Die netzwerkmedizinische Auswertung der Daten sowie die zusätzliche Integration von bekannten Expressionsdaten zur Untermauerung der biologischen Relevanz der neu gefundenen Interaktionen wurde durch mich durchgeführt. Die Ergebnisse wurden mehrfach u.a. mit international anerkannten Experten auf dem Gebiet der Peroxisomen und der Adrenoleukodystrophie (R. Erdmann und S. Kemp) diskutiert und führten schließlich zusammen mit A.S. Lotz-Havla und P. Guder zur geteilten Erstautorenschaft.

2. Einleitung

2.1 Proteininteraktionen und deren Bedeutung hinsichtlich Krankheitsmechanismen

Leben wird durch Wechselwirkungen bestimmt. Dies gilt sowohl für unsere soziale Interaktion als auch auf zellulärer Ebene. Wechselwirkungen können Prozesse in die eine oder in die andere Richtung lenken. Die einzelnen zellulären Bestandteile müssen, perfekt aufeinander abgestimmt miteinander wechselwirken, um eine optimale zelluläre Funktion zu gewährleisten (Hartwell *et al.*, 1999; Stelzl *et al.*, 2005; Loscalzo and Barábasi, 2011; Luck *et al.*, 2017). So können zum Beispiel innerhalb eines Zellorganells Substrate von Enzymen gebunden werden, um weiter abgebaut zu werden. Ebenso müssen aber auch Zellorganelle miteinander wechselwirken, um verschiedene intermediäre Stoffwechselprodukte auszutauschen. Insbesondere Interaktionen zwischen Proteinen sind ein elementarer Bestandteil zellulärer Funktionen, zum Beispiel für die Signalübertragung innerhalb und zwischen Zellen oder die Ausübung vielfältiger molekularer Stoffwechselvorgänge. Seit Beendigung des humanen Genomprojektes wird der Protein-kodierende Teil unseres Genoms auf ca. 20.000-25.000 Gene geschätzt (Collins *et al.*, 2004). Folgt man der Online Mendelian Inheritance in Man (OMIM) Datenbank (Amberger *et al.*, 2015), so sind genetische Variationen in 4.668 Genen mit 7.220 Phänotypen assoziiert (https://www.omim.org/statistics/geneMap; Stand Aug 2022). Eine Zahl, die aufgrund der nun immer schneller und kostengünstiger zur Verfügung stehenden Exom- und Genom-Sequenzierungen weiter ansteigen wird. Trotz dieser steigenden Zahl an Genotyp-zu-Phänotyp Assoziationen sind die molekularen Mechanismen, die von einer genetischen Variation zu einer klinischen Ausprägung führen, selten linearer (ein Gen-ein-Enzym-eine-Funktion), sondern komplexerer Natur und meist nicht verstanden (Edwards *et al.*, 2013; Gstaiger and Aebersold, 2013; Sahni *et al.*, 2013; Zhang, Kuivenhoven and Groen, 2015; Piñero *et al.*, 2016). Nicht nur die korrekte zelluläre Funktion, sondern auch Fehlfunktionen sind das Resultat komplexer Wechselwirkungen (Zhang, Kuivenhoven and Groen, 2015). So kann eine genetische Variation nicht nur zum vollständigen Funktionsverlust des assoziierten Genproduktes und damit zum totalen Ausfall all seiner Wechselwirkungen führen, sondern auch zu einer vermehrten (gain-of-function) oder verminderten (loss-of-function) Bindung zu seinen Interaktionspartnern innerhalb eines zellulären Interaktionsnetzwerkes. In der wissenschaftlichen Literatur ist dieses Phänomen als *edgetic perturbation* beschrieben (Zhong *et al.*, 2009; Sahni *et al.*, 2013, 2015; Rolland *et al.*, 2014).

Beispielhaft sei hier eine Interaktionsstudie zum Genprodukt des CFTR-Gens angeführt (Pankow *et al.*, 2015). Das CFTR-Gen kodiert für einen Chlorid-Ionen Kanal und genetische Variationen sind mit zystischer Fibrose assoziiert. Im Rahmen der Studie von Pankow et al. wurde sowohl nach zellulären Proteininteraktionen (PPI) des Wildtyp-Proteins als auch nach Interaktionen der bei zystischer Fibrose am häufigsten vorkommenden p.Phe508del (F508del) Variante gesucht und die beiden Interaktionsnetzwerke miteinander verglichen. Es zeigte sich, dass die F508del Variante deutlich mehr Interaktionen zu Proteinen eingeht (gain-of-function), die für Störungen im

Trafficking sowie der Degradierung des Proteins im Vergleich zum Wildtyp verantwortlich sind. Somit kann ein Teil des klinischen Phänotyps mit Veränderungen des Interaktionsnetzwerk auf Proteinebene assoziiert werden (Pankow *et al.*, 2015).

Dieses Beispiel legt nahe, dass die Kenntnis sämtlicher Interaktionspartner aller Protein-kodierender Gene eine wichtige Voraussetzung für ein korrektes Verständnis der Mechanismen ist, die von einem Genotyp zu einem Phänotyp führen (Loscalzo and Barábasi, 2011; Vidal, Cusick and Barabási, 2011; Carter, Hofree and Ideker, 2013; Zhang, Kuivenhoven and Groen, 2015; Caldera *et al.*, 2017). Im Folgenden wird zunächst der Begriff humanes Interaktom eingeführt und dann auf experimentelle Methoden zur Detektion von PPI eingegangen.

2.2 Das humane Interaktom

Die Gesamtheit aller molekularen Wechselwirkungen einer humanen Zelle wird als humanes Interaktom bezeichnet (Koh *et al.*, 2012). Dabei ist mit der Definition im engeren Sinne häufig nur die Gesamtheit aller PPI einer Zelle gemeint. Wie in 2.1 bereits erläutert, ist die Kenntnis des Interaktoms einer Zelle für ein besseres Verständnis zellulärer Funktionen und Fehlfunktion von entscheidender Bedeutung. Daher wurde kürzlich die Erstellung eines humanen Refernzinteraktoms – ähnlich des humanen Referenzgenom – als Ziel ausgegeben (Vidal, 2016). Hierfür ist die systematische experimentelle Untersuchung aller möglichen PPI notwendig. Nimmt man ca. 20.000 – 25.000 Protein-kodierende Gene zur Grundlage, so ergeben sich mehr als 200 Millionen mögliche Proteinpaare und damit eine enorme Anzahl an durchzuführenden Experimenten (Vidal, 2016). Eine Abschätzung der wahrscheinlichen wirklichen Größe des humanen Referenzinteraktoms erscheint schwierig. Verschiedene Schätzungen reichen von 100.000 – 1 Million zellulärer PPI (Stumpf *et al.*, 2008; Venkatesan *et al.*, 2009; Vidal, 2016; Luck *et al.*, 2017).

Hierbei noch nicht mit berücksichtigt sind die durch alternatives Spleißen entstehenden unterschiedlichen Isoformen oder die zeitliche Variabilität einer PPI. Zieht man die Expressionstärke und damit die quantitative Verfügbarkeit von Proteinen in Abhängigkeit der unterschiedlichen Gewebe, spezialisierenden Zellen und der subzellulären Lokalisation zusätzlich in Betracht, so ergibt sich eine weiter wachsende Komplexität (Ellis *et al.*, 2012; Yang *et al.*, 2016).

Die in den letzten Jahren stetig weiterentwickelten experimentellen Methoden zur Detektion von PPI haben zur Veröffentlichung einiger großer systematisch vermessener Proteininteraktionsdatensätze geführt (Rual, Venkatesan and Hao, 2005; Stelzl *et al.*, 2005; Ewing *et al.*, 2007; Havugimana *et al.*, 2013; Rolland *et al.*, 2014; Hein *et al.*, 2015; Huttlin *et al.*, 2015). Zudem wurde ein Referenzinteraktom veröffentlicht, welches aus insgesamt 52.569 PPI zwischen 8.275 Proteinen besteht (Luck *et al.*, 2020). Diese Datensätze werden zusammen mit kleineren Datensätzen, die aus Studien mit spezifischen Fragestellungen entstehen, in Datenbanken abgelegt und für weitere Analysen zugänglich gemacht. Eine Übersicht über häufig verwendete Datenbanken gibt Tabelle 1. Das systematische Zusammenführen dieser Interaktionen aus den unterschiedlichen Datenbanken erlaubt das Erstellen eines humanen Interaktoms, basierend auf experimentellen

Daten. Der Vorteil des Zusammenführens unterschiedlicher Datensätze, insbesondere unter unterschiedlichen experimentellen Bedingungen entstandenen Daten, ist die damit erreichbare Komplementarität. Erst kürzlich wurde im Rahmen einer Studie ein solcher Datensatz eines humanen Interaktoms veröffentlicht. Dieser enthält Informationen über PPI aus 15 verschiedenen Datenbanken. Er besteht aus 16.677 Proteinen und 243.603 Interaktionen zwischen diesen (Cheng *et al.*, 2018). Mit Hilfe eines solchen Datensatzes ist es möglich, Erkrankungen und deren Zusammenhänge untereinander zu untersuchen und zu verstehen. Hierfür hat sich das Forschungsgebiet der Systemmedizin bzw. Netzwerkmedizin entwickelt, welches anhand bestehender Daten und mittels *in-silico* Modellierungen Hypothesen zu Krankheitsmechanismen, Biomarkern und therapeutischen Optionen generieren kann (Barabási, Gulbahce and Loscalzo, 2011; Loscalzo and Barábasi, 2011; Menche *et al.*, 2015). Der zweite Teil dieser Arbeit beschreibt die Anwendung eines solchen Ansatzes erstmals für die Gruppe der angeborenen Stoffwechselerkrankungen. Grundvoraussetzung hierfür ist ein möglichst komplettes humanes Interaktom. Der erste Teil dieser Arbeit hat sich daher mit der Etablierung einer weiteren experimentellen Methode zur Untersuchung von PPI befasst und zeigt exemplarisch, wie man diese Methode zur Untersuchung von Krankheitsgenen und deren Proteinen verwenden kann. Im Folgenden werden die wichtigsten Methoden zur Detektion von PPI erläutert, sowie im Anschluss die Grundlagen, der für diese Arbeit verwendeten Methode zur Detektion von PPI in lebenden Zellen vorgestellt.

Tabelle 1: Übersicht über Datenbanken, die Informationen zu bekannten PPI speichern und zur Verfügung stellen.

Datenbank	URL	Referenz
BioGRID	https://thebiogrid.org	Oughtred et al., 2021
HINT	http://hint.yulab.org	Das and Yu, 2012
HIPPIE	http://cbdm.uni-mainz.de/hippie/	Alanis-Lobato, Andrade-Navarro and Schaefer, 2017
HPRD	https://www.hprd.org	Prasad et al., 2009
HuRI	http://www.interactome-atlas.org	Luck et al., 2020
IID	http://iid.ophid.utoronto.ca	Kotlyar et al., 2019
InnateDB	https://www.innatedb.ca	Breuer et al., 2013
IntAct	https://www.ebi.ac.uk/intact/home	Orchard et al., 2014
MINT	https://mint.bio.uniroma2.it	Licata et al., 2012
MyProtein-Net	https://netbio.bgu.ac.il/myproteinnet2/	Basha et al., 2015
PIPS	http://www.compbio.dundee.ac.uk/www-pips/	McDowall, Scott and Barton, 2009
STRING	https://string-db.org	Szklarczyk et al., 2021
TissueNet	https://netbio.bgu.ac.il/tissuenet/	Basha et al., 2017

2.3 Methoden zur experimentellen Detektion von Proteininteraktionen

2.3.1 Übersicht über die wichtigsten Methoden

In den letzten Jahren haben sich einige Standardmethoden zur systematischen experimentellen Detektion von PPI im Hochdurchsatz etabliert. Tabelle 2 gibt eine Übersicht über die am weitesten verbreiteten Methoden. Generell lassen sich diese in Methoden unterscheiden hinsichtlich der Untersuchung von Interaktionen zwischen zwei Proteinen (binäre Detektion) oder um Proteinkomplexe zu detektieren. Die drei wichtigsten und in Studien am häufigsten verwendeten Methoden sind *Yeast-2-Hybrid* (Y2H) sowie zur Detektion von Proteinkomplexen Affinitätsaufreinigung mit anschließender Massenspektrometrie (*AP-MS*) oder Co-Fraktionierung mit anschließender Massenspektrometrie (CoFrac-MS) (Cafarelli *et al.*, 2017; Luck *et al.*, 2017).

Zwar wurde kürzlich ein erstes humanes Referenzinteraktom publiziert (Luck *et al.*, 2020). Ein Großteil der bereits erwähnten möglichen 200 Millionen Proteinpaare konnte jedoch bisher nicht systematisch und experimentell vermessen werden (Vidal, 2016). Als Grund hierfür wird die unbefriedigende Sensitivität der einzelnen Screeningmethoden angenommen (Luck *et al.*, 2017). Die aktuell am häufigsten verwendeten experimentellen Methoden bieten jeweils unterschiedliche Vor- und Nachteile und eignen sich daher unterschiedlich gut zur Detektion von PPI (Tabelle 2). Dies sei am Beispiel der *Y2H* Methode kurz verdeutlicht. Mit *Y2H* werden die Interaktionen der zu untersuchenden Proteine im Zellkern von Hefezellen detektiert und damit in einem für diese Proteine nicht-physiologischen Umfeld in einer nicht-humanen Wirtszelle. Eine Lokalisation im korrekten zellulären Kompartiment und mögliche für die Interaktion wichtige posttranslationale Modifikationen können somit nicht stattfinden. Dies kann die Detektion bestimmter PPI - wie zum Beispiel bei Membranproteinen - unmöglich machen. Es ist Konsens in der Literatur, dass die Entschlüsselung des kompletten humanen Interaktoms nur mit der steten Weiterentwicklung existierender Methoden und Entwicklung neuer Methoden sowie deren komplementären Einsatz gelingen kann. (Braun *et al.*, 2009; Lemmens, Lievens and Tavernier, 2010; Cafarelli *et al.*, 2017; Luck *et al.*, 2017).

2.3.2 Biolumineszenz Resonanz Energie Transfer zur Detektion von Proteininteraktionen in lebenden Zellen

Neben den oben erwähnten weit verbreiteten Methoden zur Detektion von PPI gibt es zusätzliche weniger im Hochdurchsatz angewendete experimentelle Verfahren. Zu diesen gehören Methoden, die auf Resonanz Energie Transfer (RET) beruhen, wie Fluoreszenz Resonanz Energie Transfer (FRET) und Biolumineszenz Resonanz Energie Transfer (BRET). Letztere wurde in unserer Arbeitsgruppe zur Detektion von PPI im Hochdurchsatz etabliert. Der BRET-Assay basiert auf einem natürlichen Phänomen, welches unter anderem in der Qualle *Aequoria victoria* und der Seefeder *Renilla reniformis* beobachtet werden kann. Hierbei kommt es zu einer strahlenlosen

Energieübertragung auf einen entsprechenden Akzeptor. Die Energie ist dabei durch die Reaktion eines Substrates mit einer Luciferase (Donor-Enzym) entstanden. Befinden sich sowohl Akzeptor als auch Donor innerhalb des Förster-Radius (< 10nm), so findet die Energieübertragung statt und der Akzeptor emittiert Licht einer bestimmten Wellenlänge. Der Einsatz von BRET zur Detektion von PPI wurde erstmals von Xu et al. beschrieben (Xu, Piston and Johnson, 1999). Die beiden zu untersuchenden Proteine werden dabei jeweils mit einem entsprechenden Donor- oder Akzeptor-Protein genetisch fusioniert und nach Transfektion in der gewünschte Zelllinie exprimiert. Interagieren die beiden zu untersuchenden Proteine miteinander, gelangen Donor und Akzeptor innerhalb des Förster-Radius. Läuft nun am Donor durch Zugabe eines Substrates eine Energieerzeugende Reaktion, so kann oben erwähnter strahlenloser Energietransfer auf den Akzeptor stattfinden. Der Akzeptor emittiert nun Licht, welches entsprechend detektiert und quantifiziert (BRET-Ratio) werden kann. Die Effizienz des Energietransfers und damit die Höhe der BRET-Ratio ist dabei invers proportional zur Entfernung von Donor und Akzeptor (Xu, Piston and Johnson, 1999; Pfleger and Eidne, 2006; Bacart *et al.*, 2008). Die Methode hat zwei wesentliche Vorteile: (i) Die Zellen müssen zur Zugabe des Substrates sowie zur anschließenden Detektion nicht lysiert werden und (ii) die Messungen finden im physiologischen Kontext statt, so dass auch post-translationale Modifikationen ablaufen können, die für PPI wichtig sein könnten (Bacart *et al.*, 2008; Kobayashi *et al.*, 2019).

BRET hat sich seit seiner Erstbeschreibung im Hinblick auf die verwendeten Donor- und Akzeptor-Enzyme stetig weiterentwickelt (Ayoub & Pfleger, 2010; Kobayashi et al., 2019; Kocan et al., 2008; Mo et al., 2015). In der vorliegenden Arbeit wurde eine optimierte Version des BRET[1] Systems verwendet, mit Luciferase (hRluc) als Donor und einem Derivat des grün fluoreszierenden Proteins (Venus-YFP) als Akzeptor (Tags) (Gersting, Lotz-Havla and Muntau, 2012). Die genetische Fusion dieser beiden Tags an die zu untersuchenden Proteine kann entweder am N- oder C-Terminus erfolgen. Für eine zuverlässige Detektion einer Proteininteraktion kann die Orientierung der Tags von Bedeutung sein (Bacart *et al.*, 2008). Zum Beispiel könnte eine Interaktion zweier Proteine über deren N-Terminus durch zwei N-Terminal fusionierte Tags verhindert werden. Um diese falsch-negativen Befunde zu minimieren, wurden in der vorliegenden Arbeit die BRET-Experimente mit sowohl N- als auch C-Terminal fusionierten Tags - also insgesamt 8 Kombinationsmöglichkeiten pro Proteinpaar - durchgeführt. Zur Klassifikation eines Proteinpaares als interagierendes Proteinpaar wird die BRET-Ratio herangezogen. Diese ist neben der oben bereits erwähnten Distanz der Tags auch vom Expressionsverhältnis abhängig (Akzeptor/Donor Ratio). Auch zwei nicht interagierende Proteine könnten durch zufällige Begegnung ihrer Tags zu einem BRET Signal führen (Bacart *et al.*, 2008). Dies kann man sich mit Hilfe von Sättigungskurven, sog. BRET Evaluationsexperimente, zu Nutze machen, um falsch-positive BRET Signale von richtig positiven zu unterscheiden. Eine falsch-positive Interaktion wird bei steigender Akzeptor/Donor Ratio zu einem linearen Anstieg der BRET-Ratio führen, wohingegen eine tatsächlich stattfindende Interaktion gesättigt werden kann.

BRET wird bisher insbesondere zur Detektion von Membran-PPI sowie als Hochdurchsatz-Screening-Methode für pharmakologisch einsetzbare Moleküle verwendet, jedoch bisher nicht zur automatisierten Hochdurchsatz-Detektion von PPI (Hamdan *et al.*, 2005; Couturier and Deprez, 2012; Robinson, Yang and Zhang, 2014). Ziel der vorliegenden Promotionsarbeit war es daher, BRET hinsichtlich des Einsatzes als Screeningmethode für PPI zu validieren und anzuwenden. Um einen erhöhten Durchsatz zu generieren, sollte der BRET-Assay mittels einer eigens dafür konzipierten Tecan-Arbeitsplattform zunächst automatisiert werden. Die so entstehenden Interaktionsdaten können mittels systemmedizinischer Ansätze ausgewertet und Rückschlüsse über zu Grunde liegende Pathomechanismen gewonnen werden. Das folgende Kapitel gibt eine kurze Einführung in die in dieser Arbeit verwendeten systemmedizinischen Ansätze.

Tabelle 2: Übersicht über häufig verwendete Methoden zur Detektion von Proteininteraktionen.

Methode	Prinzip	Vorteile	Nachteile	Referenzliteratur / Anwendungen
Y2H and variations	Zwei Fragmente eines Transkriptionsfaktors (TF) fusioniert an Proteinpaar; bei Interaktion erfolgt Rekonstitution des TF und Transkription eines Reporters; Interaktion muss im Zellkern der Hefezelle stattfinden	günstig; hochdurchsatzfähig;	Hefezellkern; keine PTM	Hamdi and Colas, 2012; Rolland et al., 2014; Vidal and Fields, 2014; Choi et al., 2019; Luck et al., 2020
MYTH / MaMTH	Zwei Fragmente Ubiquitin fusioniert an Proteinpaar; Bei Interaktion, Abspaltung eines TF dieser aktiviert Expression eines Reporters	günstig; Membran Proteine	MYTH: Detektion in Hefezelle; MaMTH: ein Protein mit Membran assoziiert	Snider et al., 2010; Petschnigg et al., 2014
LUMIER	Proteinpaar fusioniert mit Antibody-Flag sowie Luciferase; Co-IP basierte Detektion des Proteinpaares; Readout mittels Fluoreszenz	hochdurchsatzfähig	Zelllyse vor Detektion	Braun et al., 2009; Blasche and Koegl, 2013; Barrios-Rodiles et al., 2017
wNAPPA	Proteinpaar Flag-tagged und in ein "transcription-translation-system" gegeben; ein Protein kann über Flag-Tag an eine membran binden; interagieren beiden, so kann der Partner via immunoaffinity detektiert werden	unterschiedliche Bedingungen wie pH oder Salz-Konzentration einstellbar	in vitro	Braun et al., 2009; Rolland et al., 2014
Biomolecular Complementation Assays: PCA, BiFC, BiLC	Reporterprotein in zwei Fragmente geteilt und fusioniert an je ein zu untersuchendes Protein; Bindung der beiden Proteine führt Fragmente zusammen und aktiviert Reporterprotein	nicht auf Zellkern limitiert wie Y2H	keine Erfassung der Bindungsaffinität	Tarassov et al., 2008; Rolland et al., 2014; Miller et al., 2016; Li, Wang and Di, 2019
MAPPIT / MASPIT /KISS	Ein Protein fusioniert an C-Terminus eines Zytokin-Rezeptors, das zweite an "STAT3-recruitment site"; Bindung führt zu Aktivierung von STAT3 und induziert Transkription eines Rezeptors	Messung in Säugerzellen; small-molecule screening möglich (MASPIT)	Interaktion muss an Zytoplasma-Membran stattfinden	Lievens et al., 2009, 2016; Venkatesan et al., 2009; Rolland et al., 2014; Vyncke et al., 2019

Tabelle 2 Fortsetzung

Methode	Prinzip	Vorteile	Nachteile	Referenzliteratur / Anwendungen
FRET/BRET	Proteinpaar wird fusioniert mit Donor/Akzeptorpaar und exprimiert; Strahlenloser Energietransfer wenn innerhalb 10nm nach Anregung; Detektion mittels Fluoreszenz (FRET) oder Lumineszenz (BRET);	Messung von PPI in "Echtzeit"; lebende Zellen; Interaktionsdynamik / Affinität quantifizierbar	FRET: Hintergrund Fluoreszenz kann zu falsch-Negativen führen; FRET/BRET: Fusionsproteine notwendig; Überexpression	Boute, Jockers and Issad, 2002; Hamdan et al., 2005; Bacart et al., 2008; Kobayashi et al., 2019
AP-MS	Bindung eines Proteins an spez. Antikörper, anschließende Aufreinigung und Detektion möglicher Interaktionspartner mittels MS	hochdurchsatzfähig; Detektion von Proteinkomplexen; Überexpression nicht zwingend erforderlich	Zelllyse vor Detektion; falsch-Negative insb. transienter PPI; falsch Positive bei zufälliger Co-Elution gerichtete MS	Hein et al., 2015; Gordon et al., 2020; Huttlin et al., 2021
CoFrac-MS	Komplexe aus Zelllysaten werden durch Chromatographie/Gel-Ekektrophorese fraktioniert und mittels MS analysiert	hochdurchsatzfähig; ungerichtete MS; keine Überexpression notwendig	falsch Positive bei zufälliger Co-Elution möglich	Havugimana et al., 2013; Moutaoufik et al., 2019
BioID-MS	Biotinligase wird fusioniert mit Protein und exprimiert; direkte / indirekte Proteinpartner werden biotinyliert (10nm Radius) und mittels Affinitätsaufreiniung und MS detektiert	detektiert direkte/indirekte/transiente Interaktoren	Proteine in der "Nähe" (10nm) werden biotinyliert, falsch Positive; Überexpression	(Gupta *et al.*, 2015; Antonicka *et al.*, 2020)

Y2H: Yeast-Two-Hybrid; **MYTH**: Membrane Yeast-Two-Hybrid; **MaMTH**: Mammalian-Membrane Yeast-Two-hybrid; **LUMIER**: Luminescence-based Mammalian Interactome; **wNAPPA:** Well-based Nucleic Acid Programmable Protein Arrays; **PCA**: Protein-complementation Assay; **BiFC**: Biomolecular Fluorescence Assay; **BiLC**: Bimolecular Luminescence Complementation; **MAPPIT**: Mammalian Protein-Protein Interaction Trap; **MASPIT**: Mammalian Small Molecule Protein Interaction Trap; **KISS**: Kinase Substrate Sensor; **FRET**: Fluorescence Resonance Energy Transfer; **BRET**: Bioluminescence Resonance Energy Transfer; **AP-MS**: Affinity Purification combined with Mass Spectrometry; **CoFrac-MS**: Co-Fractionation coupled with Mass Sepectrometry; **BioID-MS**: Proximity-dependent Biotin Identification coupled with Mass Spectrometry

2.4 Einführung in Systemmedizin

Verschiedenste interne und externe Einflüsse haben Auswirkungen auf die Funktionen eines Systems. Der menschliche Körper lässt sich als solch ein dynamisches System betrachten und das Gebiet der Systemmedizin untersucht die Zusammenhänge und Auswirkungen der auf den Körper bzw. dessen Bestandteile wirkenden unterschiedlichen Einflüsse. Genetische, epigenetische oder metabolische Wechselwirkungen der einzelnen Bestandteile des Systems, wie z.B. Proteine, lassen sich mittels Graphen bzw. Netzwerken darstellen. Das Gebiet der mathematischen Graphentheorie liefert hierfür die Grundlagen. Die zu untersuchenden Objekte werden dabei als Knoten gesehen, welche über Linien miteinander verbunden sind, sofern eine Beziehung zwischen diesen besteht. Diese Netzwerke können dann weiter hinsichtlich des Zusammenhangs der einzelnen Knoten oder zum Beispiel der Topologie des Netzwerkes untersucht werden, um zum Beispiel die Fehleranfälligkeit eines solchen Netzwerkes auf interne oder externe Einflüsse zu berechnen.

Die Netzwerkmedizin bezeichnet dabei ein Untergebiet der Systemmedizin, das diese Graphentheorien auf biologische Netzwerke - wie das humane Interaktom – anwendet (Sonawane *et al.*, 2019; Conte *et al.*, 2020). Sie betrachtet dabei Krankheiten als Folge von Störungen in den zu Grunde liegenden biologischen Netzwerken. Die Anwendung der Netzwerkmedizin ermöglicht es, Krankheiten systematisch zu untersuchen und so Erkenntnisse zur Entstehung sowie zu möglichen neuen Therapieansätzen zu finden. Vielmehr noch bieten sich durch die Erstellung individueller Netzwerke - zum Beispiel durch Einbindung von Transcriptomics Daten - Chancen auf eine „personalisierte Medizin“ (Yadav, Vidal and Luck, 2020).

Der Begriff Netzwerkmedizin wurde von Albert-Laszlo Barabasi, einem ungarischen Physiker, geprägt (Barabasi, 2007; Loscalzo and Barábasi, 2011). Er fand, dass fast alle Netzwerke „skalenfrei“ sind, eine Eigenschaft, die Netzwerke robust gegenüber zufälligen Fehlern macht. Diese Theorie besagt, dass einige wenige Knoten eines Netzwerks viele Verbindungen zu anderen Knoten ausbilden – sogenannte *hub* Knoten –, während ein Großteil der Knoten jedoch nur wenige Verbindungen hat (Abb. 1a) (Barábasi and Albert, 1999; Barabasi and Bonabeau, 2003).

Das humane Interaktom lässt sich als Netzwerk darstellen (Barábasi and Oltvai, 2004). Hierbei repräsentieren Knoten Proteine und die jeweiligen Linien stellen nachgewiesene Wechselwirkungen zwischen diesen dar. Auch hier finden sich *hub* Knoten. Diese *hub* Proteine sind zum Beispiel häufiger mit Tumorerkrankungen assoziiert und korrelieren häufiger mit essenziellen Proteinen (Piñero *et al.*, 2016). Der Ausfall eines *hub* Proteins kann zum kompletten Zusammenbruch des Netzwerks führen. Zudem konnte gezeigt werden, dass Proteine nicht zufällig im humanen Interaktom verteilt sind, sondern eine höhere Wahrscheinlichkeit haben, miteinander verbunden zu sein, wenn sie an der Entstehung der gleichen Erkrankung beteiligt sind (Loscalzo and Barábasi, 2011). Solche Cluster werden *disease module* genannt. Wobei ein *disease module* definiert ist als Gruppe interagierender Netzwerkknoten, die zu einer gemeinsamen zellulären Funktion beitragen (funktionelle Cluster) und dessen Störung zu einer Erkrankung führt (Oti and Brunner, 2007; Loscalzo and Barábasi, 2011). Hierauf aufbauend konnte gezeigt werden, dass es trotz

unvollständigem humanen Interaktom möglich ist, *disease module* und deren Beziehungen zueinander im Netzwerkkontext zu identifizieren und dass sich durch diese Beziehungen Hypothesen zu gemeinsamen Pathomechanismen der zu Grunde liegenden Erkrankungen generieren lassen (Abb. 1b) (Ghiassian, Menche and Barabási, 2015; Menche *et al.*, 2015). Zudem beschrieben Menche et al., dass Proteine, die mit einem ähnlichen klinischen Phänotyp assoziiert sind, häufiger in der gleichen Netzwerknachbarschaft liegen und dass auch *disease modules* mit ähnlichen klinischen Symptomen näher beieinander liegen (Menche *et al.*, 2015). *Disease modules* können sich in unterschiedlichen Geweben hinsichtlich ihres Aufbaus unterscheiden und werden häufiger in den jeweiligen für die Erkrankung typischen Geweben exprimiert (Kitsak *et al.*, 2016). Mit Hilfe dieser Theorien ist die Einordnung von Kandidatengenen aus Genom- oder Exom-Sequenzierungen möglich (Menche *et al.*, 2015). Mit Hilfe von netzwerkmedizinischen Methoden konnten zum Beispiel neue Hypothesen zur Krankheitsentstehung für Asthma, COPD, Autismus und kardiovaskuläre Erkrankungen generieren (Li *et al.*, 2014; Sharma *et al.*, 2015, 2018; Lempiäinen *et al.*, 2018). Weitere Möglichkeiten ergeben sich durch die zusätzliche Integration von Omics-Daten, wie zum Beispiel Informationen über Medikamente und deren jeweilige Zielproteine. Hiermit können sich für bereits etablierte Medikamente neue Indikationen ergeben (*drug-repurposing*) (Aguirre-Plans et al., 2018; Cheng et al., 2018; Wang and Loscalzo, 2018). Netzwerkmedizinische Methoden wurde bisher nicht für die Gruppe der angeborenen Stoffwechselerkrankungen angewendet (Argmann *et al.*, 2016), die einen Fokus unserer Arbeitsgruppe darstellen. Gerade für diese klinisch sehr heterogene und bisher nicht komplett verstandene Erkrankungsgruppe könnten sich durch die Anwendung solcher Methoden Hypothesen über Erkrankungsmechanismen sowie therapeutischen Optionen ergeben.

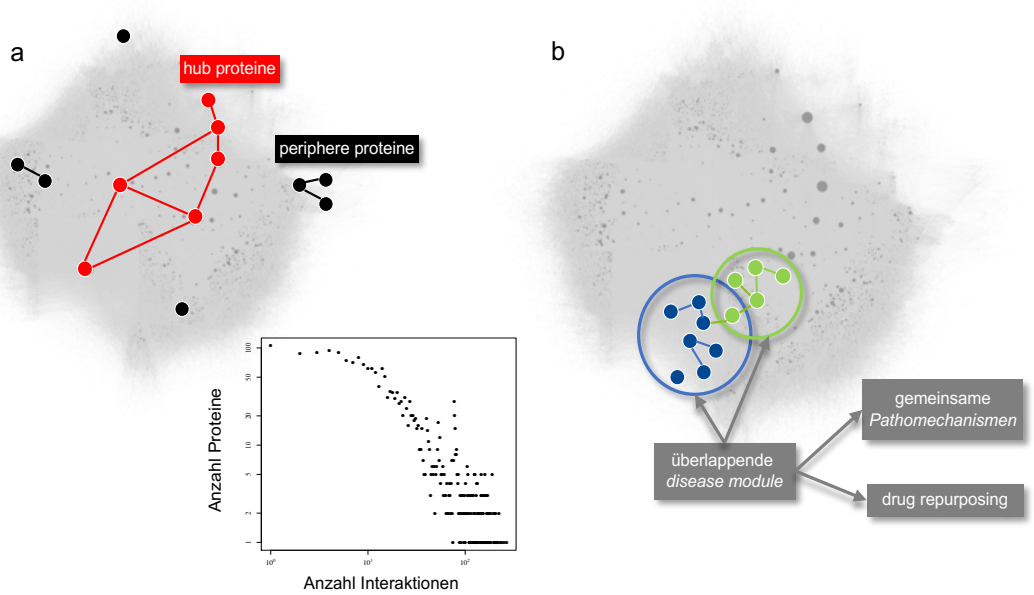


Abbildung 1: Netzwerkdarstellung eines humanen Interaktomdatensatzes (Daten aus Cheng et al. 2018). Knoten stellen dabei jeweils Proteine dar, eine Verbindung zwischen diesen kennzeich-

net eine experimentell nachgewiesene Interaktion. (a) Schematisch dargestellt sind hier rote Knoten als zentral liegende *hub Proteine* sowie schwarze Knoten als peripherer liegende Proteine. Ein skalen-freies Netzwerk - wie das humane Interaktom - hat viele *periphere Proteine*, die im Mittel weniger Interaktionen zu anderen Proteinen ausbilden. Im Gegensatz hierzu gibt es wenige *hub Proteine*, die durch viele Interaktionen charakterisiert sind. Die Verteilung der Proteine in Abhängigkeit von der Anzahl ihrer jeweiligen Interaktionen ist im Inlay rechts unten dargestellt. (b) Schematische Darstellung zweier sich im Netzwerk überlappender *disease modules* jeweils in grün und blau. Eine enge Nachbarschaft im Netzwerk lässt Rückschlüsse auf mögliche gemeinsame molekulare Pathomechanismen sowie auf die Übertragung bestehender Therapien einer Erkrankung auf die sich überlappende Erkrankung zu (drug repurposing).

3. Ziele der Arbeit

Eingehende Kenntnisse über die einzelnen Interaktionen zellulärer Bestandteile sind eine wichtige Voraussetzung für ein besseres mechanistisches Verständnis der Abläufe einer Zelle. Erkrankungen können aus Störungen auf unterschiedlichen Ebenen in diesem komplexen Zusammenspiel resultieren. Eine wichtige Ebene für die korrekte Funktionsausübung einer Zelle sind Proteininteraktionen (PPI). Zur experimentellen Detektion dieser PPI haben sich in den letzten Jahren einige Standardmethoden etabliert, welche aufgrund unterschiedlicher Vor- und Nachteile nicht ausreichen werden, um systematisch alle PPI - das humane Interaktom - zu detektieren. Im Rahmen der hier durchgeführten Arbeiten sollte zunächst eine weitere Methode zur Untersuchung von PPI in lebenden Zellen (Biolumineszenz-Resonanz-Energie-Transfer) mittels einer Tecan-Arbeitsplattform automatisiert, validiert (iBRET) und deren Anwendung zur Untersuchung von Erkrankungsgenen gezeigt werden.

Die somit erhaltenen Daten können als Interaktionsnetzwerke dargestellt werden. Diese ermöglichen *in-silico* mittels mathematischer Netzwerkanalysen neue Erkenntnisgewinne über Krankheitsentstehung, potenzielle neue Biomarker sowie Therapieansätze. Krankheiten werden hierbei nicht mehr als einzelne Entität, sondern im systemischen Kontext gesehen. Unsere Arbeitsgruppe beschäftigt sich mit der Erforschung von neuen Therapieansätzen für genetische Erkrankungen mit Fokus auf angeborene Stoffwechselerkrankungen. Im zweiten Teil dieser Arbeit sollten daher systemmedizinische Algorithmen genutzt werden, um diese erstmals für die Gruppe der angeborenen Stoffwechselerkrankungen anzuwenden und so eine Basis für weitere Analysen sowie für ein besseres Verständnis der Pathomechanismen von angeborenen Stoffwechselerkrankungen zu schaffen.

4. Schlussbetrachtung

Im Rahmen der hier vorgelegten kumulativen Promotionsarbeit habe ich mich mit der Automatisierung und Validierung der Detektion von PPI mittels iBRET, der Erstellung von Interaktionsnetzwerken, sowie der Anwendung von system- bzw. netzwerkmedizinischen Methoden für ein besseres mechanistisches Verständnis von genetisch bedingten Erkrankungen befasst.

Zunächst wurde der bereits in unserer Arbeitsgruppe etablierte BRET Assay mittels einer Tecan-Arbeitsplattform automatisiert. Hierfür mussten sämtliche Arbeitsschritte von Transfektion, Inkubation und Detektion in ein 96-Well Format überführt werden. Dies ermöglicht die automatisierte Untersuchung von bis zu 640 Proteinpaaren in lebenden Zellen pro Woche.

Schätzungen zufolge sind erst knapp 20% aller humanen PPI experimentell vermessen (Caldera *et al.*, 2017). Dies bedeutet, dass entweder zusätzliche Screeningmethoden notwendig sind oder bestehende Methoden modifiziert werden müssen (Choi *et al.*, 2019). Die Anwendung von iBRET bietet die Möglichkeit, PPI in einem deutlich physiologischeren Milieu zu untersuchen, als dies zum Beispiel mittels der Y2H-Methode möglich ist. iBRET ist hierbei nicht als Ersatz, sondern als Ergänzung zu den bereits bestehenden Screeningmethoden zu sehen.

Anhand eines Referenzdatensatzes, bestehend aus 61 gut dokumentierten bekannt positiven sowie 60 zufällig gezogenen Proteinpaaren konnten wir zeigen, dass iBRET eine hervorragende Sensitivität (58%) sowie Spezifität (96%) zur Detektion von PPI in lebenden Zellen aufweist. Wir konnten außerdem belegen, dass die verbesserte Performance neben dem Screening im physiologischen Kontext lebender Zellen auch aus der Testung aller 8 möglichen Tag-Kombinationsmöglichkeiten eines Proteinpaares resultiert. Durch die automatisierte Klassifikation der Ergebnisse in interagierend, nicht-interagierend sowie in einen Graubereich mit anschließender Nachtestung durch Evaluationsexperimente ist es weiter möglich, die Sensitivität von iBRET von 58% auf 72% zu steigern. Die Anwendung eines Graubereiches mit anschließender Nachtestung hat zwar einen erhöhten experimentellen Aufwand zur Folge, kann aber je nach Anwendungsszenario – Screening vs. gezieltem Testen – die Detektionsrate deutlich anheben.

Eine Anwendungsmöglichkeit von iBRET konnte in Kooperation mit der Arbeitsgruppe Baiker des Max von Pettenkofer-Institutes der LMU München aufgezeigt werden. Hier wurde untersucht, welche Bedeutung das virale Protein ORF25 auf die Replikation des Varizella Zoster Virus hat. Im Rahmen dieser Arbeit wurde iBRET sowie Y2H und eine weitere Methode (LuMPIS) zur Untersuchung von PPI eingesetzt. Verglichen mit Y2H - hier als Goldstandard -, zeigte sich für iBRET eine Sensitivität von 47,82% und eine Spezifität von 78,57%. Dies bestätigte die bereits vorher ermittelten guten Detektionsraten. Die Ergebnisse führten zu einer Ko-Autorenschaft (Vizoso Pinto et al., 2011). Es sei anzumerken, dass im Anschluss iBRET weiter hinsichtlich Anzahl der Replikate (Reduktion von Triplikaten auf Duplikate) sowie Automatisierung angepasst wurde.

Die Kenntnisse über PPI ermöglicht die Erstellung von Interaktionsgraphen bzw. -netzwerken. iBRET eignet sich hervorragend zur Untersuchung kleinerer Interaktionsnetzwerke, zum Beispiel

von einzelnen Erkrankungsgenen. In der vorliegenden Arbeit wurde ein auf peroxisomale Proteine begrenztes Interaktionsnetzwerk für das Erkrankungsgen ABCD1 vermessen. Dabei wurden für das Wildtyp-Protein bisher unbekannte Interaktionspartner gefunden. Zusätzlich konnte gezeigt werden, dass diese neuen Partner mit bestimmten pathogenen Missense-Varianten von ABCD1 nicht mehr interagieren können. Dies zeigt, dass pathogene Missense-Varianten zu Störungen im Interaktionsnetzwerk von ABDC1 führen, und lässt zusätzliche Pathomechanismen der X-ALD vermuten.

iBRET kam während meiner Promotionstätigkeit für weitere Interaktionsstudien zu Erkrankungsgenen zur Anwendung (Hillebrand *et al.*, 2012; Nold-Petry *et al.*, 2015; Frixel *et al.*, 2016), bei denen meine Mitarbeit zu weiteren Ko-Autorschaften geführt hat (Lotz-Havla et al. 2021, Guder et al., 2018).

In den letzten Jahren konnte gezeigt werden, dass sich Interaktionsnetzwerke mittels allgemein gültiger mathematischer Algorithmen untersuchen lassen. Die Betrachtung von Krankheiten als Folge von Störungen in diesen Netzwerken - Netzwerkmedizin - bietet die Möglichkeit, rasch *in-silico* neue Hypothesen zu Krankheitsentstehung, Diagnostik und Therapie zu generieren (Caldera *et al.*, 2017; Silverman *et al.*, 2020). In der vorliegenden Arbeit konnten wir erstmals die Anwendung von netzwerkmedizinischen Methoden auf die Erkrankungsgruppe der angeborenen Stoffwechselerkrankungen zeigen. Mittels Berechnung netzwerkbasierter Distanzen von Erkrankungsgruppen konnten Verbindungen zwischen metabolischen Erkrankungen und anderen nichtmetabolischen jedoch gut beschriebenen Erkrankungsgruppen gefunden werden. Diese Verbindungen lassen Rückschlüsse auf gemeinsame Pathomechanismen zu und helfen, gut verstandene Therapieansätze der einen Erkrankung auf die andere Erkrankung zu übertragen. Um Medikamente und Therapieansätze übertragen zu können, ist es notwendig zu wissen, wie nah sich die entsprechenden *drug targets* den Erkrankungs-spezifischen *disease modulen* sind. Ein solches *drug target network* wurde im Rahmen dieser Arbeit erstmals für metabolische Erkrankungen erstellt. Dieses *drug target network* soll als Ausgangspunkt für weitere Untersuchungen zu neuen therapeutischen Optionen für Stoffwechselerkrankungen dienen.

Zusammenfassend liefert die vorliegende Promotionsarbeit eine validierte, hochsensitive Methode zur Detektion von PPI in lebenden Zellen – iBRET. iBRET eignet sich mit seinem Durchsatz und seiner Sensitivität hervorragend zur Erweiterung bestehender Interaktionsnetzwerke, sowie zur Untersuchung des Einflusses genetischer Variationen auf einzelne PPI. Diese PPI lassen sich als Interaktionsnetzwerke darstellen und mittels netzwerkmedizinischer *in-silico* Algorithmen die Konsequenzen von Störungen in diesen Netzwerken berechnen. Die vorliegende Arbeit zeigt erstmalig die Anwendung dieser Algorithmen auf die Gruppe der angeborenen Stoffwechselerkrankungen und liefert eine Grundlage für weitere Analysen.

5. Abstract (English):

Cells consist of multiple interacting cellular components. As such, proteins, besides genes, rna or organelles represent a major part, whose interactions result in diverse cellular functions. Mapping these interactions provides the basis to unravel associated function and the molecular mechanisms of disease. The human interactome is defined as the entire network of human protein interactions. Despite huge efforts, the human interactome is still limited with respect to coverage and accuracy. In recent years it has become evident that a full interactome coverage is only achievable by combining additional complementary experimental methods. However, several studies demonstrated, that the current interactome datasets already enable *in-silico* network-based approaches to study human disease. These studies revealed the modular nature of the human interactome yielding the definition of disease modules. These disease modules assist in elucidating common underlying cellular mechanisms between different disease. Moreover, the combination of highly reliable experimental screening methods and *in-silico* approaches will speed up the discovery of additional biomarkers and therapeutic options.

This thesis firstly describes the set-up, validation, and application of an automated informatics-aided bioluminescence resonance energy transfer-based method (BRET) for the detection of protein interactions in living mammalian cells (iBRET). Extensive benchmarking of iBRET and its application to map the peroxisomal interaction network of ABCD1 demonstrates its added value in screening protein interactions at a reasonable scale. Secondly, this thesis demonstrates - for the first time - the application of a network medicine approach to study the clinically very heterogenous group of inborn errors of metabolism (IEM). We found that IEM associated proteins are organized as linked modules within the human interactome and several shared pathways between IEM but also between IEM and non-IEM disease groups. These links may provide interesting hints to new drug targets.

Altogether, the results of this thesis show that the combination of iBRET with network medicine approaches will facilitate and accelerate elucidation of the functional context of disease associated proteins and of molecular processes leading to clinical phenotypes.

6. Veröffentlichung I: Inborn errors of metabolism and the human interactome: a systems medicine approach.

Journal of Inherited Metabolic Disease (2018) 41:285–296
https://doi.org/10.1007/s10545-018-0140-0

METABOLOMICS

Inborn errors of metabolism and the human interactome: a systems medicine approach

Mathias Woidy[1] · Ania C. Muntau[1] · Søren W. Gersting[2]

Received: 8 October 2017 / Revised: 1 January 2018 / Accepted: 10 January 2018 / Published online: 5 February 2018

Abstract

The group of inborn errors of metabolism (IEM) displays a marked heterogeneity and IEM can affect virtually all functions and organs of the human organism; however, IEM share that their associated proteins function in metabolism. Most proteins carry out cellular functions by interacting with other proteins, and thus are organized in biological networks. Therefore, diseases are rarely the consequence of single gene mutations but of the perturbations caused in the related cellular network. Systematic approaches that integrate multi-omics and database information into biological networks have successfully expanded our knowledge of complex disorders but network-based strategies have been rarely applied to study IEM. We analyzed IEM on a proteome scale and found that IEM-associated proteins are organized as a network of linked modules within the human interactome of protein interactions, the IEM interactome. Certain IEM disease groups formed self-contained disease modules, which were highly interlinked. On the other hand, we observed disease modules consisting of proteins from many different disease groups in the IEM interactome. Moreover, we explored the overlap between IEM and non-IEM disease genes and applied network medicine approaches to investigate shared biological pathways, clinical signs and symptoms, and links to drug targets. The provided resources may help to elucidate the molecular mechanisms underlying new IEM, to uncover the significance of disease-associated mutations, to identify new biomarkers, and to develop novel therapeutic strategies.

Keywords Inborn errors of metabolism · Interactome · Network medicine · Disease module · Protein-protein interaction

Introduction

The human genome consists of ~25,000 protein coding genes and most proteins carry out cellular functions by interacting with other small molecules or macromolecules. In the network resulting from all these interactions—the human interactome—the cellular components, such as proteins, RNA, or metabolites, serve as nodes and their interactions are commonly depicted as edges (Schadt 2009; Barabási et al 2011; Caldera et al 2017). Inborn errors of metabolism (IEM) are the consequence of genetic variation, but the disease phenotype does not only result from functional alteration of the affected gene product, perturbations rather spread along the links of the underpinning cellular networks (Schadt 2009; Barabási et al 2011; Menche et al 2015; Piñero et al 2016; Caldera et al 2017).

Proteins are the drivers of cellular function and the entirety of protein-protein interactions forms a large sub-network within the human interactome (Vidal et al 2011). A complete high-quality map of the protein interaction network is of fundamental importance for the understanding of diseases (Vidal et al 2011; Luck et al 2017). The analysis of these networks with the emerging tools of network medicine may help to understand the cellular mechanisms underlying IEM, the relationships between different IEM or between IEM and other diseases.

Biological networks are not random but governed by organizing principles. They are scale-free, *i.e.*, there are many nodes with few neighbors and few nodes with many neighbors, which are called hubs (Barabási et al 2011; Pavlopoulos

Communicated by: Ron A Wevers

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10545-018-0140-0) contains supplementary material, which is available to authorized users.

✉ Søren W. Gersting
soeren.gersting@med.lmu.de

[1] University Children's Hospital, University Medical Center Hamburg Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany

[2] Department of Molecular Pediatrics, Dr. von Hauner Children's Hospital, Ludwig-Maximilians-University Munich, Lindwurmstrasse 4, 80336 Munich, Germany

et al 2011). Disease-associated proteins are not scattered randomly in the interactome, but tend to interact with each other (Goh et al 2007; Feldman et al 2008; Caldera et al 2017). A cluster of disease-associated proteins in the same network neighborhood forms a subgraph, which constitutes a disease module (Barabási et al 2011; Menche et al 2015). These modules may be tissue-specific (Kitsak et al 2016) and personalized gene expression pools that relate to these modules impact individual disease expression (Menche et al 2017). Within a disease module, different diseases may arise from common mechanisms and display overlapping phenotypes as shown for complex diseases, such as inflammation, asthma, and cardiovascular diseases (Menche et al 2015; Sharma et al 2015; Ghiassian et al 2016).

However, network medicine approaches have not been used to systematically study IEM (Argmann et al 2016) as a heterogeneous group of disorders affecting virtually all human biochemical pathways and impairing the function of many organs. A hierarchical classification of IEM provided by the Society for the Study of Inborn Errors of Metabolism (SSIEM) lists 612 diseases with MIM numbers (www.omim.org) in 15 disease groups (www.ssiem.org/resources/IEC.asp). By using the emerging tools of network medicine, we aimed to investigate whether the entity of IEM is indeed organized as disease modules in the interactome of protein interactions and how these relate to each other. We approached IEM on a proteome scale and established an IEM-specific interactome (IEMi), which comprised 298 of 427 IEM-related proteins. To further improve the clinical relevance of the strategy we investigated relations of the IEMi to other non-IEM diseases, integrated database information such as functional annotations from the gene ontology, phenotypic features, and links to drugs or biologically active compounds.

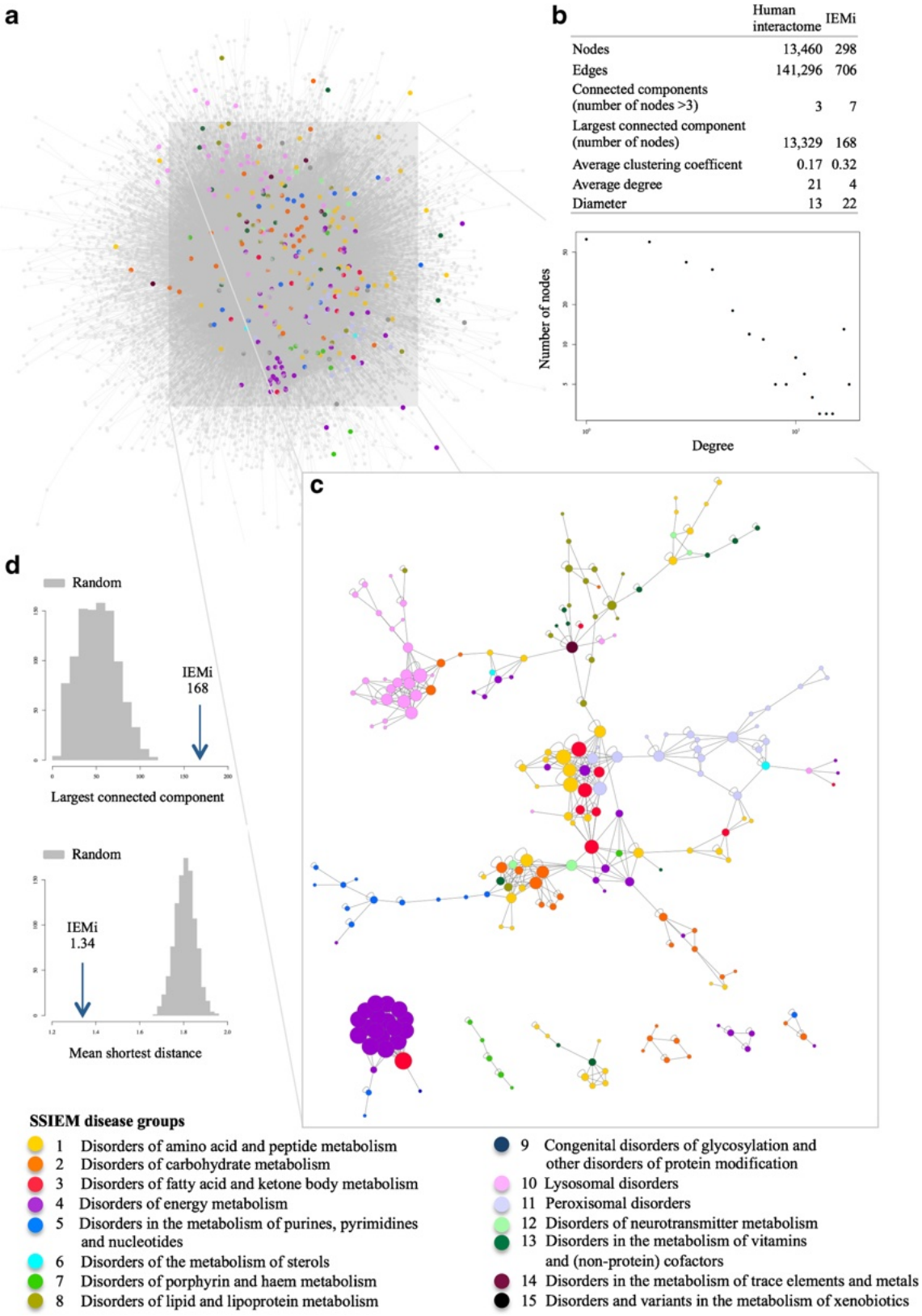
Results

The IEM interactome

As a platform for network medicine analyses, Menche et al (2015) compiled a human interactome (Fig. 1a) that consists of 13,460 proteins connected by 141,296 high-confident interactions including binary protein-protein interactions, regulatory interactions and metabolic pathway interactions (Fig. 1b). Following the SSIEM classification, 427 disease genes are known to be associated with an IEM (Suppl. Table 1) and 376 of their gene products, the IEM-associated proteins, were identified in the human interactome. Within the set of 376 IEM-associated proteins, 298 proteins established 706 interactions with each other, resulting in the IEM interactome (IEMi) (Fig. 1c, Suppl. Fig. 1). The IEMi comprised seven connected components of more than three nodes, where the largest connected component contained 168 IEM disease proteins linked to 14 out of 15 IEM disease groups. The second largest connected component with 22 nodes mainly consisted of proteins associated with disorders of energy metabolism (group 4), while the nine proteins that formed the third largest connected component are associated with disorders of amino acid and peptide metabolism (group 1) and disorders in the metabolism of vitamins and (non-protein) cofactors (group 13).

Fig. 1 From the human interactome to the IEM interactome. **a** Network representation of the interactome of all human protein-protein interactions. This interactome represents a complete map of all known interactions between proteins and is a part of the full human interactome including interactions between proteins and DNA, RNA or metabolites. Nodes are proteins and they are connected by an edge when two proteins interact. Colored nodes are associated with an inborn error of metabolism (IEM) and colors depict different IEM disease groups. IEM-associated proteins that interact with other IEM-associated proteins define the IEM interactome (IEMi) within the human interactome. **b** Network parameters of the complete human interactome and the IEMi. Log-log plot of the degree distribution for all nodes of the IEMi against the number of nodes points to a scale-free network. **c** The inset shows the seven largest connected components ($n > 3$ nodes) of the IEMi. The size of a node is proportional to the number of its neighbors, which are interaction partners (the number of neighbors defines the degree of a node within a network). For a detailed scalable view including protein names, see Suppl. Fig. 1. **d** The distribution of largest connected component sizes and mean shortest distances for randomly chosen sets of proteins from the human interactome compared to the IEMi (arrows). The number of randomly chosen proteins is similar to that of the IEMi. The arrows show the observed values for largest connected component (z-score, 5.32) and mean shortest distances (z-score, −10.09) of the IEMi, which are both significantly different from random expectation

The *degree* of a node within a network describes the number of connections to other nodes, in the case of a protein interaction network the number of interactions a given protein establishes with other proteins. As previously shown for other biological networks (Barabási et al 2011), the *degree* distribution of the IEMi followed a power law, indicating a scale-free biological network with many proteins with low *degree* and few proteins with high *degree* (Fig. 1b). The analysis of the IEMi showed that, on average, IEM-associated proteins have only four interaction partners, whereas the average *degree* of the human interactome is 21. Highly connected nodes, the hub proteins, take over specific biological roles and can be subdivided by their dynamical behavior in networks (Pavlopoulos et al 2011; Seebacher and Gavin 2011). They can either constitute a central node within a module or connect several subgraphs within a network. Within the IEMi we identified 45 nodes with a degree ≥10. The *clustering coefficient* of a node describes the tendency with which the neighbors of this node also interact with each other. The average *clustering coefficient* (cl) of the IEMi was 0.32; thus, larger than for the human interactome (cl, 0.17). The *diameter* is the largest distance between any two pairs of nodes in a network. The *diameter* was 22 for the IEMi and 13 for the human interactome. The *largest connected component* of the IEMi (168, *z*-



score 5.3) was larger than the random expectation (52.4), whereas the network-based *mean shortest distance* of IEM-associated proteins was smaller (1.34, *z-score* − 10.1) than the random expectation (1.81) (Fig. 1d). Taken together, the results from the analysis of the network-based measures *largest connected component* and *mean shortest distance* support the

notion that IEM-associated proteins tend to locate to the same network neighborhood in the human interactome. In addition, we observed two different patterns of clustering within the IEMi. Proteins associated with certain disease groups built rather homogenous clusters and consequently were separated from other IEM-associated proteins (lysosomal disorders; peroxisomal disorders; disorders in the metabolism of purines, pyrimidines and nucleotides; disorders in the metabolism of lipids and lipoproteins; disorders of energy metabolism). Proteins associated with other disease groups either contributed to heterogeneous clusters central to the largest connected component or they were distributed over remote areas of the network.

IEM disease modules

To test the disease modules hypothesis for each of the IEM disease groups following the SSIEM classification, we calculated the network-based measures *largest connected component* and *mean shortest distances* for all disease groups (Fig. 2a). The analysis was performed for all IEM-associated proteins mapping to the human interactome (*n*, 376). Disease groups 6 (disorders of the metabolism of sterols) and 15 (disorders in the metabolism of xenobiotics) were excluded from the analysis due to low numbers of associated proteins or lack of interactions. The IEM disease groups were compared to sets of nodes with the respective numbers of nodes that represented the general characteristics of the human interactome. We computed *z*-scores for each comparison and observed that 13 disease groups established significant disease modules (*largest connected component*, *z*-scores >1.6; *mean shortest distance*, *z*-scores <1.6). The largest disease modules were identified for lysosomal disorders (25 nodes), peroxisomal disorders (22 nodes), and disorders of energy metabolism (18 nodes), where the latter formed a highly interlinked motif. Three medium-sized modules arising from the disease groups 1, 5, and 8 were less interlinked. The remaining modules either were not interlinked (disease group 7) or were very small (disease groups 12–14).

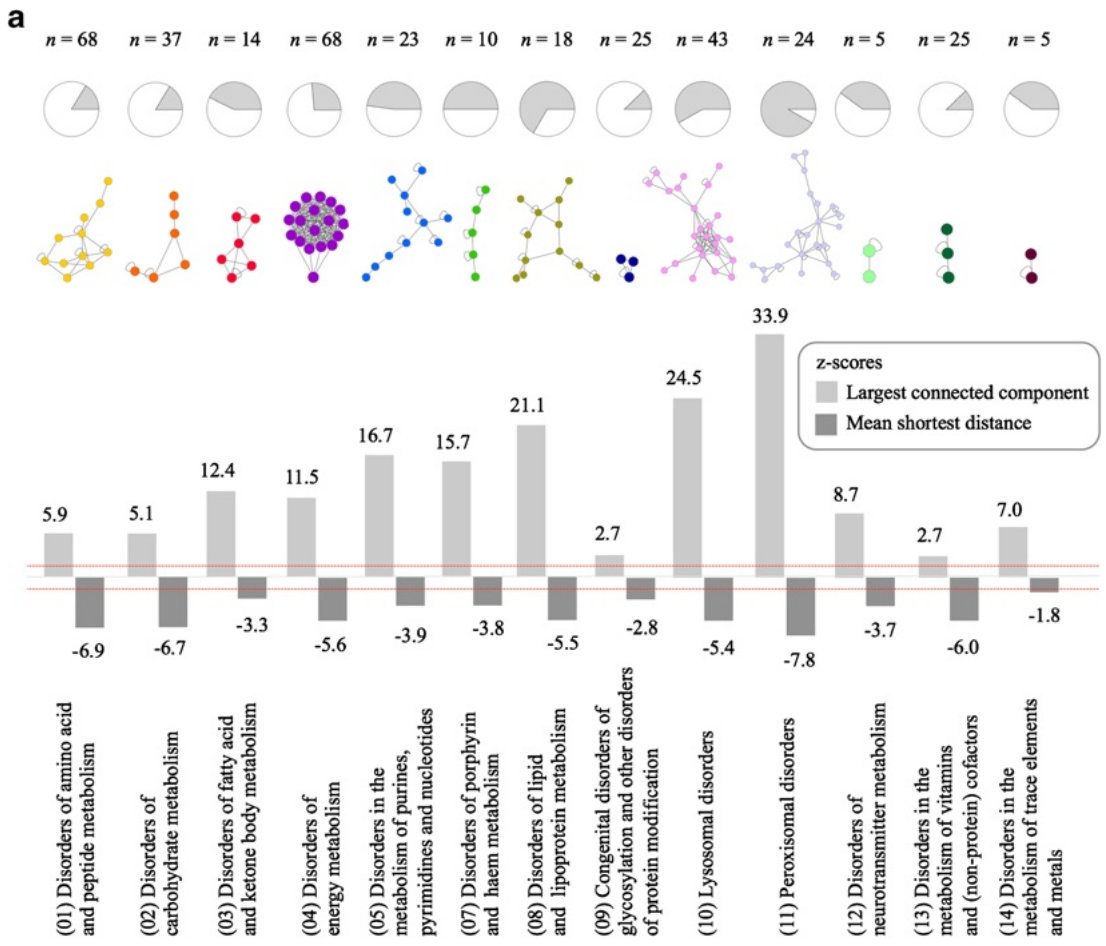
Next, we aimed to investigate the overlap between the IEM disease groups based on protein interactions (Fig. 2b) and found 211 interactions between proteins belonging to different disease groups. Disorders of amino acid and peptide metabolism (disease group 1) had a central position within this network. This group formed a subnetwork together with disorders of carbohydrate metabolism (group 2), disorders of fatty acid and ketone body metabolism (group 3), and disorders of energy metabolism (group 4) that accounted for 44% of all interactions between the disease groups. In addition, lysosomal disorders (group 10) shared 6% of interactions with disorders of carbohydrate metabolism (group 2). Peroxisomal disorders (group 11) showed significant overlap with disorders of amino acid and peptide metabolism (group 1) as well as disorders of fatty acid carbohydrate metabolism (group 3). Congenital disorders of glycosylation and other disorders of protein modification (group 9) were, except for one interaction to disorders of fatty acid and ketone body metabolism (group 3), disconnected from the other IEM disease groups.

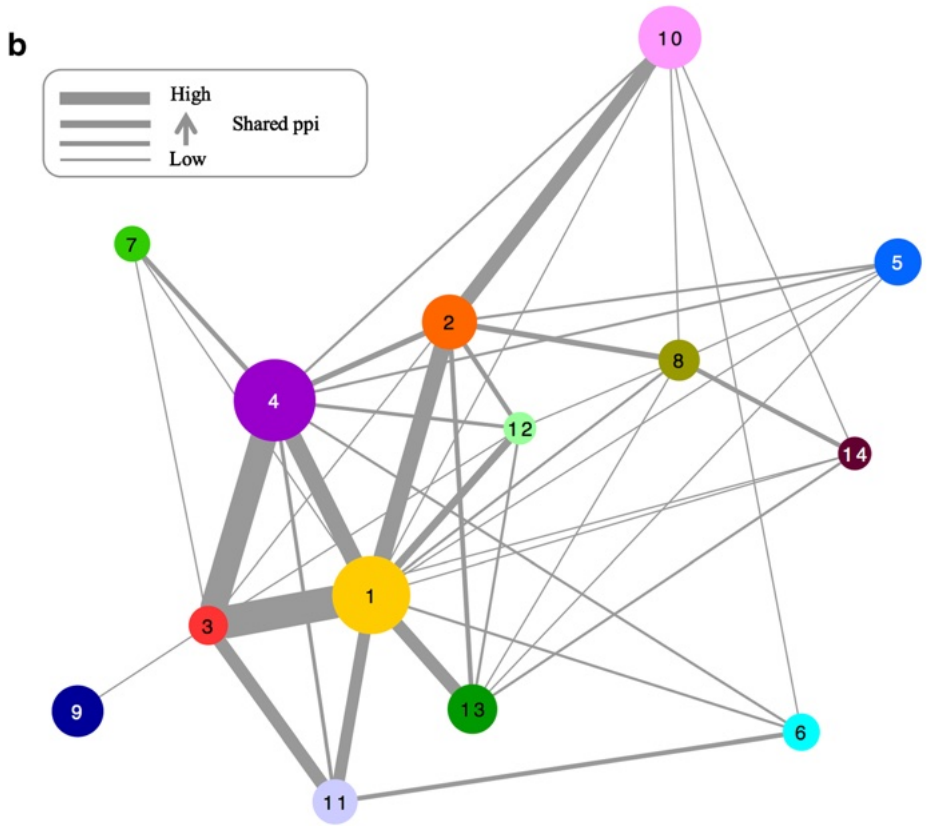
Disease groups with a high degree of self-organization, *i.e.*, many nodes belonging to the same disease group are organized within the specific module (groups 5, 7, 8, 10, 11), established few interactions to other disease groups. In the case of disease groups showing a low degree of self-organization, *i.e.*, only a small share of the nodes is organized within the specific module, the degree of interconnection with other disease groups increased (groups 1, 2, 4). Taken together, we identified disease groups with a high degree of interaction-mediated overlap to other IEM, whereas others constituted more distinct entities based on their interaction behavior. Disease group 3 (disorders of fatty acid and ketone body metabolism) represented a special case. Here we observed a compact specific disease module where half of the proteins belonging to the disease group contributed to the module. In addition to this rather high degree of self-organization strong interactions were established with disease groups 1 (disorders of amino acid and peptide metabolism), 4 (disorders of energy metabolism), and 11 (peroxisomal disorders) and weak interactions were established with disease groups 2, 7, 9, and 12.

Fig. 2 IEM disease modules. **a** Identification of disease modules within IEMi that are specific for IEM disease groups. The *z*-scores for the observed values of the largest connected components (light gray) and mean shortest distances (dark gray) are shown. The red lines indicate *z*-scores of +1.6 and −1.6 as thresholds for the level of significance (*p*-value <0.05). The pie charts give the share of proteins involved in the largest connected component (gray area) compared to the number of proteins associated with the respective IEM disease group. Disease groups 6 (disorders of the metabolism of sterols) and 15 (disorders in the metabolism of xenobiotics) were excluded from the analysis due to low numbers of associated proteins or lack of interactions. **b** IEMi disease group network. Nodes represent IEM disease groups. The thickness of edges is proportional to the number of shared interactions. The node size corresponds to the number of proteins associated with the respective IEM disease group

Relationship between IEM and non-IEM pathways and diseases

To uncover disease-disease relationships of IEM with other non-IEM disease classes we sought to expand the IEMi. We pursued the hypothesis that the location of IEM disease modules in the neighborhood to other disease genes in the human interactome implies shared biological pathways, similar clinical signs and symptoms, and common disease mechanisms. We extracted binary interactions of IEM-associated proteins with non-IEM proteins in the human interactome, retrieved 1994 first order interaction partners and additionally included interactions within this set of proteins (Suppl. Table 2). This


a
n = 68
n = 37
n = 14
n = 68
n = 23
n = 10
n = 18
n = 25
n = 43
n = 24
n = 5
n = 25
n = 5
z-scores
Largest connected component
Mean shortest distance
5.9
-6.9
5.1
-6.7
12.4
-3.3
11.5
-5.6
16.7
-3.9
15.7
-3.8
21.1
-5.5
2.7
-2.8
24.5
-5.4
33.9
-7.8
8.7
-3.7
2.7
-6.0
7.0
-1.8
(01) Disorders of amino acid and peptide metabolism
(02) Disorders of carbohydrate metabolism
(03) Disorders of fatty acid and ketone body metabolism
(04) Disorders of energy metabolism
(05) Disorders in the metabolism of purines, pyrimidines and nucleotides
(07) Disorders of porphyrin and haem metabolism
(08) Disorders of lipid and lipoprotein metabolism
(09) Congenital disorders of glycosylation and other disorders of protein modification
(10) Lysosomal disorders
(11) Peroxisomal disorders
(12) Disorders of neurotransmitter metabolism
(13) Disorders in the metabolism of vitamins and (non-protein) cofactors
(14) Disorders in the metabolism of trace elements and metals


b
High
Shared ppi
Low
10
7
5
2
8
4
12
14
1
3
9
13
6
11

resulted in an expanded IEM interactome (eIEMi) of 2370 proteins linked by 39,916 interactions (Fig. 3a). The eIEMi consisted of one connected component of more than three nodes (*largest connected component*, 2362), the average *degree* increased to 33 as compared to 4 for the IEMi and the *diameter* of the eIEMi decreased to 8 (Fig. 3b). Again, the degree distribution followed a power law, indicating a scale-free biological network with few highly connected hub proteins. We characterized eIMEi hub proteins with a *degree* > 100, which are important for the topology and function of the network, with respect to their cellular distribution and observed overrepresentation of proteins located in the ribosome (fold enrichment (FE), 48.01; *p*-value, 1.81E-90), cytosol (FE, 18.92; *p*-value, 1.13E-66), nucleolus (FE, 12.34; p-value, 1.18E-12), and cytoplasm (FE, 4.05; *p*-value, 3.36E-33). Membrane proteins were underrepresented (FE, 0.2; *p*-value, 0.0029) (Fig. 3c).

Next, we performed a pathway enrichment analysis for pathways associated with non-IEM proteins in the eIEMi. Besides enrichment for metabolic pathways and immune signaling pathways, we found significant enrichments for angiogenesis (FE, 2.97; p-value, 5.51E-8), apoptosis signaling pathway (FE, 3.86; p-value, 2.76E-10), and for the two diseases Parkinson disease (FE, 3.44; p-value, 1.95E-6) and Huntington disease (FE, 2.18; p-value, 0.034) (Fig. 3d and Suppl. Table 3). The gonadotropin releasing hormone receptor pathway, the CCKR signaling pathway, and angiogenesis were associated with ≥4 IEM disease groups. A high number of different pathways was enriched for non-IEM proteins interacting with IEM-associated proteins linked to disorders of carbohydrate metabolism, energy metabolism or neurotransmitter metabolism (groups 2, 4, 12), whereas no pathway enrichments were observed for non-IEM proteins associated with disease groups 6, 7, 9, 11, 13.

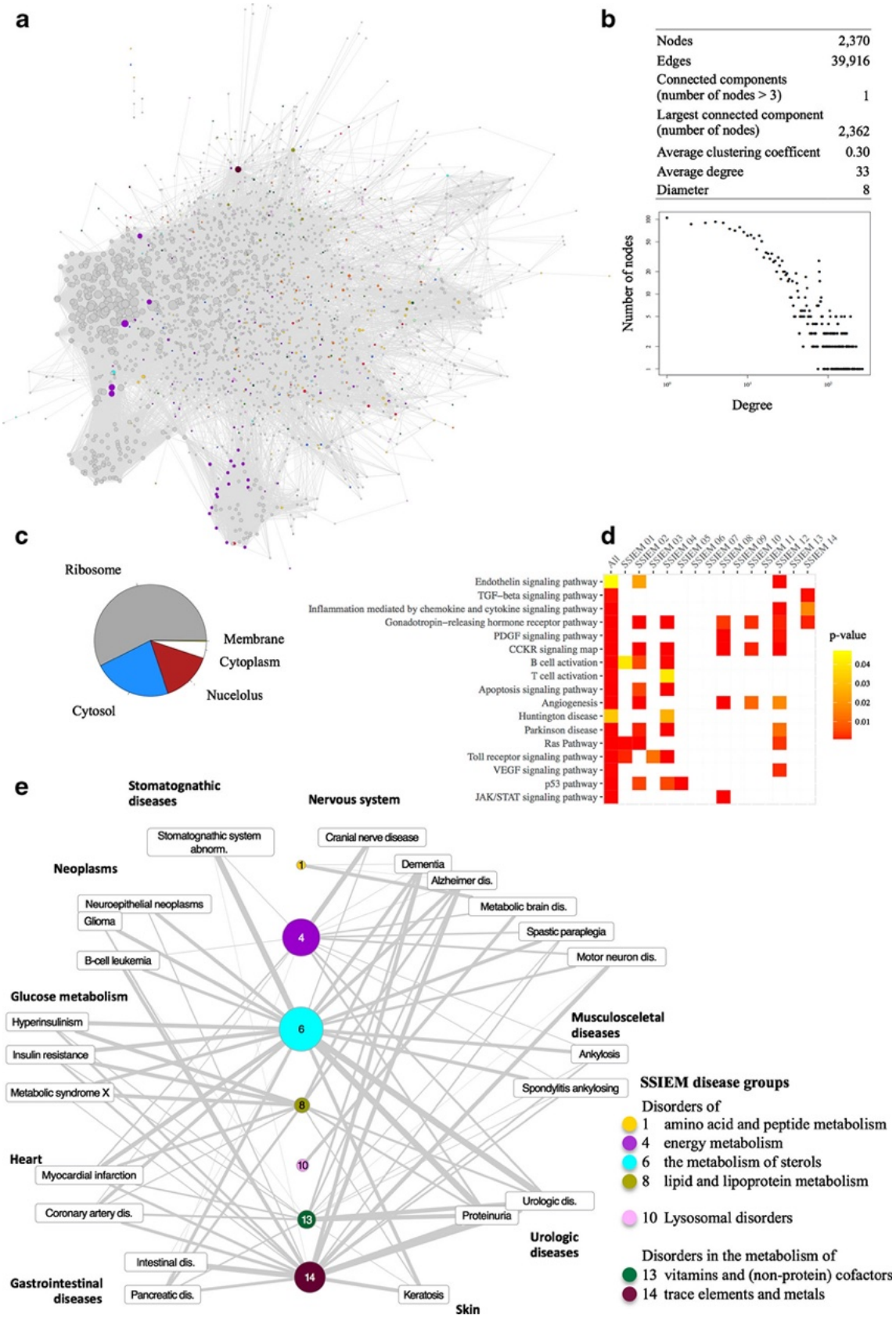
In order to evaluate the overlap of IEM disease groups with non-IEM diseases, we analyzed their network-based distances using a separation score, which is a network-based measure for pathobiological and clinical similarity (Menche et al 2015; Caldera et al 2017), and identified the related phenotypic features (MeSH terms). We calculated separation scores for IEM disease groups 1–14 against a set of 299 non-IEM diseases and ranked separation scores for each of the 4186 resulting pairs (Fig. 3e and Suppl. Table 4). Disorders of energy metabolism and metabolism of sterols (groups 4, 6) shared the most pronounced overlap to other diseases, reflected by the highest number of established links. The closest similarities among diseases, reflected by low separation scores, were observed between disorders of sterols metabolism and urologic diseases (−0.080), stomatognathic system abnormalities (−0.054), and chronic B-cell leukemia (0.054), respectively. Disorders in the metabolism of trace elements and metals were closely linked to intestinal (0.003) as well as urologic diseases (0.013). Additionally, disorders of energy metabolism were strongly connected to cranial nerve diseases (0.066). Dementia, Alzheimer disease (AD), and motor neuron diseases displayed similarities with IEM disease groups 6, 8, 13, and 14. In particular, we found associations of disorders of metabolism of sterols (group 6) with AD.

Fig. 3 The expanded IEM interactome. **a** Expanded IEM interactome (eIEMi). The IEMi was expanded to include first-order interactions of IEM-associated proteins (colored nodes) to non-IEM proteins (gray nodes). The size of a node corresponds to the number of its neighbors (degree). **b** Network parameters of the eIEMi; a log-log plot of the degree distribution of all nodes of the eIMEi against the number of nodes points to a scale-free network. **c** Pie chart of the calculated fold enrichment of cellular compartment distribution for proteins with degree >100 of the eIEMi. Ribosomal, cytosolic, and nuclear proteins are overrepresented. Membrane proteins are underrepresented. **d** Heat map of enriched PANTHER pathways of non-IEM proteins interacting with proteins of the different IEM disease groups. Cells are colored according to their respective p-value. Metabolic pathways were excluded from this matrix representation. **e** The eIMEi disease network reveals associations of IEM disease groups with non-IEM diseases based on protein interactions. Rectangle nodes are MeSH terms of non-IEM diseases. Colored nodes represent IEM disease groups, the size is proportional to the shared links with general disease groups. The width of an edge is proportional to the network-based relationship, *i.e.*, the edge width negatively correlates with a calculated separation score

The association between serum cholesterol levels and cerebral amyloidosis in the pathogenesis of AD (Puglielli et al 2003; Vaya and Schipper 2007; Reed et al 2014) is not likely due to similar genetic predispositions (Proitsi et al 2014). However, there is evidence for a mechanistic link between β-amyloid protein toxicity and mitochondrial cholesterol trafficking (Barbero-Camps et al 2014), and the HMG-CoA-reductase inhibitor Simvastatin has been demonstrated to affect β-amyloid peptides (Zandl-Lang et al 2017). In addition, we found associations of disorders in the metabolism of trace elements and metals (group 14) with AD (Fig 3e). Recently, polymorphisms in the ATP7B gene, which is responsible for Wilson disease and a part of IEM-disease group 14, have been associated with increased risk of AD (Mercer et al 2017). These findings underlined how the identification of such network-based connections and the transfer of knowledge between associated disease groups may help to elucidate molecular mechanisms of disease and to elaborate new therapeutic strategies.

IEM drug–target network

The observed similarities as to underlying biological pathways, disease mechanisms, and phenotypes between IEM-associated proteins and non-IEM proteins in the eIMEi network allow for the hypothesis of the effectiveness of common treatment strategies. Consequently, a drug may have impact on proteins that are linked to the drug target protein in a protein interaction network. We included drug-target information from the drugbank database (Law et al 2014) into the eIMEi,


a
b
Nodes 2,370
Edges 39,916
Connected components (number of nodes > 3) 1
Largest connected component (number of nodes) 2,362
Average clustering coefficent 0.30
Average degree 33
Diameter 8
Number of nodes
Degree
c
Ribosome
Membrane
Cytoplasm
Nucelolus
Cytosol
d
All
SSIEM 01
SSIEM 02
SSIEM 03
SSIEM 04
SSIEM 05
SSIEM 06
SSIEM 07
SSIEM 08
SSIEM 09
SSIEM 10
SSIEM 11
SSIEM 12
SSIEM 13
SSIEM 14
Endothelin signaling pathway
TGF–beta signaling pathway
Inflammation mediated by chemokine and cytokine signaling pathway
Gonadotropin–releasing hormone receptor pathway
PDGF signaling pathway
CCKR signaling map
B cell activation
T cell activation
Apoptosis signaling pathway
Angiogenesis
Huntington disease
Parkinson disease
Ras Pathway
Toll receptor signaling pathway
VEGF signaling pathway
p53 pathway
JAK/STAT signaling pathway
p-value
0.04
0.03
0.02
0.01
e
Stomatognathic diseases
Nervous system
Stomatognathic system abnorm.
Cranial nerve disease
Neoplasms
Dementia
Alzheimer dis.
Neuroepithelial neoplasms
Glioma
Metabolic brain dis.
Spastic paraplegia
B-cell leukemia
Motor neuron dis.
Glucose metabolism
Hyperinsulinism
Musculosceletal diseases
Insulin resistance
Ankylosis
Metabolic syndrome X
Spondylitis ankylosing
Heart
Myocardial infarction
Urologic dis.
Coronary artery dis.
Proteinuria
Urologic diseases
Intestinal dis.
Gastrointestinal diseases
Pancreatic dis.
Keratosis
Skin
1
4
6
8
10
13
14
SSIEM disease groups
Disorders of
1 amino acid and peptide metabolism
4 energy metabolism
6 the metabolism of sterols
8 lipid and lipoprotein metabolism
10 Lysosomal disorders
Disorders in the metabolism of
13 vitamins and (non-protein) cofactors
14 trace elements and metals

identified 634 approved drugs or active compounds that target 316 non-IEM proteins, and mapped their protein interactions with IEM-associated proteins in the eIEMi space. The resulting IEM drug-target network consisted of 538 proteins and 845 interactions (Fig. 4 and Suppl. Table 5). Besides a share of otherwise unlinked binary interactions between non-IEM drug targets and IEM-associated disease proteins, the IEM drug-target network showed a large connected component with 465 nodes. This sub-network was organized by several highly interlinked hub proteins, both IEM and non-IEM proteins. Hence, different groups of drug targets were interconnected by IEM-associated proteins or IEM disease modules were interconnected by non-IEM drug targets. To exemplify approaches toward the evaluation of therapeutic

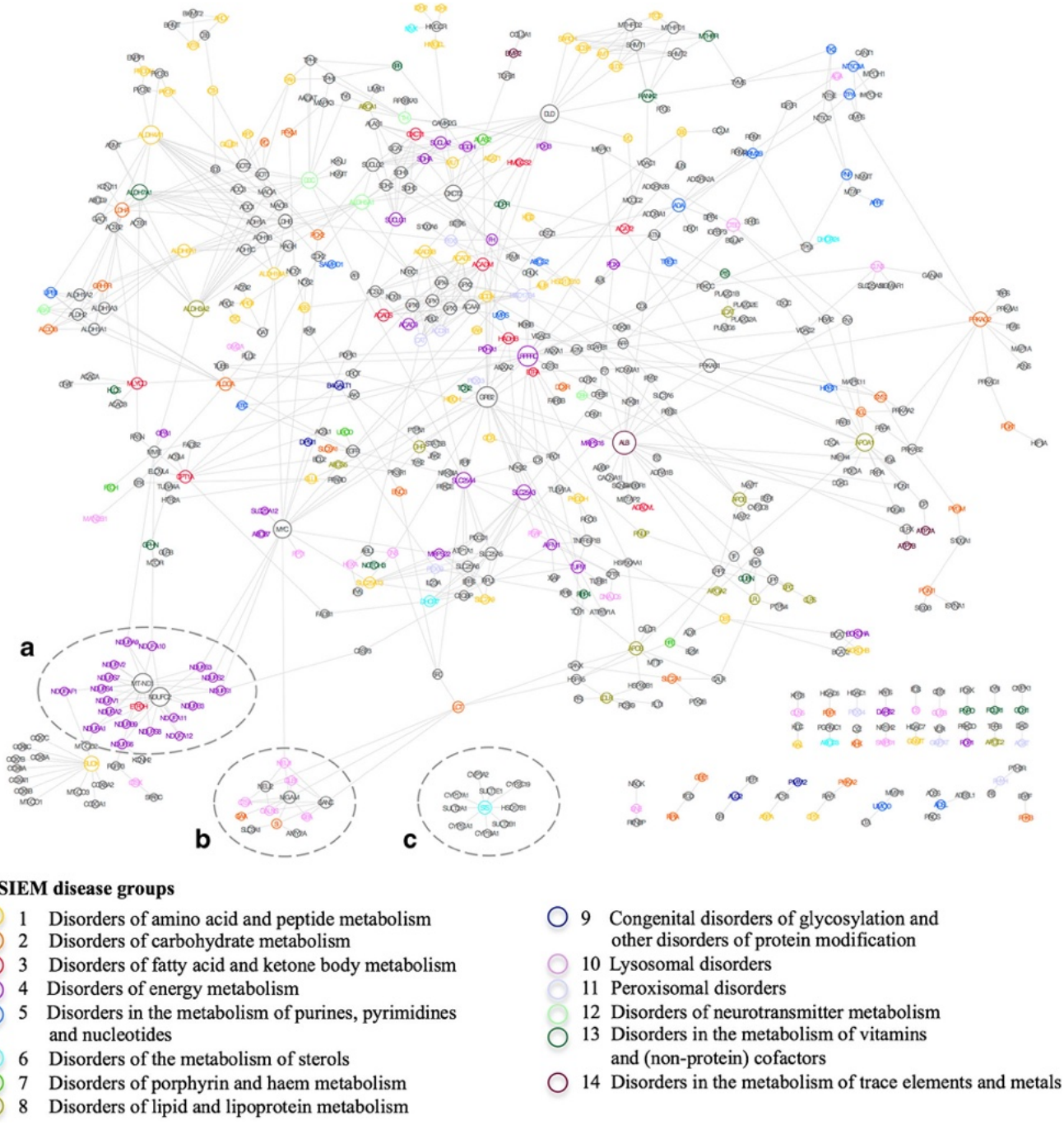


Fig. 4 The eIEMi drug target network. The eIEMi drug target network shows interactions of IEM-associated proteins with non-IEM proteins, which are listed as drug targets in the drug bank (www.drugbank.ca). For all proteins, gene names are given as short names (UniProt), nodes are colored according to the IEM disease group and gray colored nodes represent non-IEM proteins. The size of a node is proportional to the number of its neighbors (degree)

concepts and identification of novel treatment options for IEM, we discuss three scenarios. (A) An IEM disease module linked by non-IEM drug targets. The IEM disease module for disorders of energy metabolism (group 6) was interconnected by two hub proteins, which are subunits of the mitochondrial respiratory chain complex I, NADH-ubiquinone oxidoreductase (MT-ND1), and a NADH dehydrogenase (NDUFC2). MT-ND1 is listed in the drugbank database as target of volatile anesthetics, such as isoflurane and halotane, whereas NDUFC2 is an off-target of the β-blocker carvedilol. (B) Multiple links between IEM-associated proteins and non-IEM drug targets within a disease module. A disease module consisting of IEM-associated proteins linked to lysosomal disorders and disorders of carbohydrate metabolism contains different non-IEM drug targets. Here, intestinal maltase-glucoamylase (MGAM) and neutral alpha-glucosidase C (GANC) are targeted by alpha-glucosidase inhibitors. In addition, sialidase 1 is a target of low-molecular weight heparins and sialidase 2 of the neuramidase inhibitor oseltamivir. (C) An IEM-associated protein linked to multiple non-IEM drug targets. The steroid sulfatase A, which is deficient in X-linked ichthyosis, is central to a drug-target module addressed by imidazole or nicotinic acid derivatives (rifampicin, isoniazid, metronidazole).

Binary interactions of IEM-associated proteins and non-IEM drug targets may represent mechanistic links on the molecular level, and hence give rise to novel therapeutic approaches. The solute carrier family 25 member 4 gene (SLC25A4) encodes for adenine nucleotide translocase type 1 (ANT1) that interacts with mitochondrial peptidyl-prolyl cis-trans isomerase F (PPIF) in the IEM drug target network. Mutations in SLC25A4 are associated with different mitochondrial disorders including phenotypes of mitochondrial myopathy, exercise intolerance, cardiomyopathy, and progressive external ophtalmoplegia (Kaukonen et al 2000; Strauss et al 2013; Thompson et al 2016; Tosserams et al 2017). Cyclosporine A targets the mitochondrial permeability transition pore (mPTP) via the mPTP regulator PPIF and CypD-ANT1 and has been shown to improve mitochondrial function (Merlini et al 2008; Qiu et al 2014). In patients, cyclosporine A acts as a neuroprotective agent against stroke (Osman et al 2011) and is cardioprotective in events of ischemia (Hausenloy et al 2012). The interaction of SLC25A4 with PPIF in our drug-target network points to a potential therapeutic effect of cyclosporine A in the treatment of mitochondrial disorders linked to SLC25A4.

Discussion

The group of IEM displays a marked heterogeneity and IEM can affect all functions and organs of the human organism. The currently most applied classification defines IEM disease groups according to three different criteria (i) the metabolites pertained by the genetic variation such as disorders of carbohydrate metabolism; (ii) the affected biological pathways such as disorders of glycosylation; (iii) the cellular localization of the IEM-associated proteins such as lysosomal disorders. A considerable number of IEM have been discovered over the past 100 years and particularly the more frequent among them are well studied. However, also due to the recent advances in genomics, many new, partly very rare IEM have been discovered, for which satisfying knowledge about diagnostics, care, and treatment is often not available.

Besides their heterogeneity, IEM share that their associated proteins function in metabolism. The latter is a huge functional unit that, as all extended functional pathways, is organized in biological networks. Network-based studies gave rise to the disease module hypothesis (Barabási et al 2011; Vidal et al 2011) where diseases with an overlap in biological networks show significant symptom similarity and common disease mechanisms, whereas diseases residing in separated network neighborhoods are phenotypically distinct. In a systematic investigation to uncover relationships between human diseases, Menche et al identified peroxisomal disorders as a distinct disease module in the human interactome (Menche et al 2015). Systematic approaches toward human disease that integrate database information and -omics data into biological networks have proven useful to further our understanding of common or complex disorders (Menche et al 2015; Sharma et al 2015; Ghiassian et al 2016). However, network-based strategies have been rarely applied to study the complex nature of IEM (Argmann et al 2016). Our study aimed to evaluate the expediency of such approaches using the human interactome and other existing biological or disease-associated datasets.

We used publicly available protein-protein interaction data between IEM-associated proteins and showed that IEM-associated proteins tend to locate to the same neighborhood within the human interactome. We termed the resulting subnetwork IEM interactome (IEMi). The identification of IEM-specific disease modules revealed that the IEMi indeed contained allocations that correspond to the IEM-disease groups. Certain disease modules were highly interlinked to each other, on the other hand, we identified disease modules in the IEMi that consisted of proteins from many different disease groups. Notably, some IEM-associated proteins were not part of the IEMi and based on our network medicine approach did not show any overlap with other IEM. Therefore, the IEMi may offer a platform to systematically explore not only the molecular complexity of a particular disorder, but also the relationships among apparently distinct pathophenotypes of different IEM disease groups. In our study, we used a high-quality and comprehensive human interactome that has been curated recently (Menche et al 2015). However, it is estimated that current high-throughput methods cover less than 20% of

all potential pairwise interactions in the human cell (Caldera et al 2017). Therefore, the lack of associations between certain IEM-disease groups has to be evaluated carefully considering this incompleteness.

The central position of metabolism in the functional network of human biology implies a plethora of links that IEM-associated proteins can establish with non-IEM proteins including related cellular pathways and pathomechanisms in human disease. The expanded IEMi (eIEMi) explored the overlap between IEM and the human interactome through integration of database information on biological pathways, phenotypic features, and drugs or other active compounds. The graphical workup provided in this study exemplified network medicine approaches for IEM, offered resources in the supplementary material, and may be useful for the investigation of shared molecular functions, disease mechanisms, and the evaluation of existing or novel therapeutic concepts for IEM.

The transfer of knowledge within single disease groups, between related IEM disease modules or between IEM and non-IEM diseases will help to elucidate the molecular consequences associated with new disease genes, to uncover the significance of disease-associated mutations, to identify new biomarkers, and to expand the druggable space.

Materials and methods

Curation of disease genes associated with IEM

A classification of IEM is provided by the SSIEM. We downloaded the classification file at http://www.ssiem.org/resources/IEC.asp on March 15, 2015. Mapping of IEM diseases to associated disease genes was done in a web-based manner using the OMIM API (https://www.omim.org/help/api). This resulted in 427 IEM-related disease genes each assigned to the respective SSIEM disease group (Suppl. Table 1). We found only one disease protein associated with SSIEM class 15.

Constructing the IEM interactome (IEMi) and the expanded IEM interactome (eIEMi)

We used protein-protein interaction data from a previously curated high-quality interactome (Menche et al 2015). In order to map protein-protein interaction data to the curated IEM-related proteins we used Entrez Gene ID retrieved from the HGNC database (http://www.genenames.org, downloaded Feb 5, 2017). For the construction of the IEMi, we extracted only protein interactions that directly linked IEM-related proteins resulting in 708 IEM-to-IEM protein interactions between 298 different proteins. To expand the IEMi we included protein interactions of first order proteins. These are non-IEM proteins that directly interact with IEM-associated disease proteins. The eIEMi contains IEM-to-IEM, protein interactions to first order proteins and protein interactions between first order proteins, in total 39,916 links. The 1994 first order proteins are listed in Suppl. Table 2.

Calculation of *z*-score

Z-scores for the observed values V of largest connected component size and mean shortest distances were calculated as follows:

$$z\text{–}score = \frac{V - mean(V_{Random})}{stand(V_{Random})},$$

where mean(V_{Random}) and stand(V_{Random}) indicate the mean value and standard deviation of the random expectation.

Enrichment analysis

Enrichment analysis was performed using the web-interface provided by PANTHER (http://pantherdb.org, Panther Annotation version 12 released 10 July, 2017, Reference List *Homo sapiens*). We excluded metabolic pathways for the heat map presentation of enriched pathways. A full list of results is provided in Suppl. Table 3.

Comparison to non-IEM diseases

To perform a network-based comparison to other disease classes we used a previously described measure (Menche et al 2015). The separation score compares the mean shortest distance of protein pairs of the same disease to the mean shortest distance between protein pairs of different diseases. We calculated the separation score for every pair of a IEM disease group to a non-IEM disease. For non-IEM diseases we used a set of 299 diseases based on MeSH terms and compiled by Menche et al.

Drug target information

We downloaded drug target information from the DrugBank database as of July 16, 2017 (https://www.drugbank.ca, DrugBank Release Version 5.0.7). In more detail, we used the *approved Target Drug-UniProt* sheet to retrieve only approved drug target information for our proteins. We mapped UniProt IDs to Entrez Gene IDs and found 634 approved drugs including small molecules and biotech drugs targeting 316 non-IEM binding proteins.

Software

We used Cytoscape V.3.5.1 for the drawing and coloring of all networks. Network layouts were obtained using the built-in layout algorithm *organic*. The network layout for Fig. 3e was

done manually. Basic network measures were calculated using the Cytoscape plugin *NetworkAnalyzer*.

Details of funding SWG was supported by a grant from the Prinz Lennart von Hohenzollern foundation.

Compliance with ethical standards

Conflict of interest M. Woidy, A. C. Muntau, and S. W. Gersting declare that they have no conflict of interest.

Documentation of approval from the Institutional Committee for Care and Use of Laboratory Animals Not applicable.

References

Argmann CA, Houten SM, Zhu J, Schadt EE (2016) Perspective a next generation multiscale view of inborn errors of metabolism. Cell Metab 23:13–26. https://doi.org/10.1016/j.cmet.2015.11.012

Barabási A-L, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 12:56–68. https://doi.org/10.1038/nrg2918

Barbero-Camps E, Fernández A, Baulies A et al (2014) Endoplasmic reticulum stress mediates amyloid β neurotoxicity via mitochondrial cholesterol trafficking. Am J Pathol 184:2066–2081. https://doi.org/10.1016/j.ajpath.2014.03.014

Caldera M, Buphamalai P, Müller F, Menche J (2017) Interactome-based approaches to human disease. Curr Opin Syst Biol 3:88–94. https://doi.org/10.1016/j.coisb.2017.04.015

Feldman I, Rzhetsky A, Vitkup D (2008) Network properties of genes harboring inherited disease mutations. Proc Natl Acad Sci 105:4323–4328. https://doi.org/10.1073/pnas.0701722105

Ghiassian SD, Menge J, Chasman DI et al (2016) Endophenotype network models: common core of complex diseases. Sci Rep 6:27414. https://doi.org/10.1038/srep27414

Goh K, Cusick M, Valle D et al (2007) The human disease network. Proc Natl Acad Sci 104:8685–8690

Hausenloy DJ, Boston-Griffiths EA, Yellon DM (2012) Cyclosporin A and cardioprotection: from investigative tool to therapeutic agent. Br J Pharmacol 165:1235–1245. https://doi.org/10.1111/j.1476-5381.2011.01700.x

Kaukonen J, Juselius JK, Tiranti V et al (2000) Role of adenine nucleotide translocator 1 in mtDNA maintenance. Science 289:782–785. https://doi.org/10.1126/science.289.5480.782

Kitsak M, Sharma A, Menche J, Guney E (2016) Tissue specificity of human disease module. Nat Sci Rep:1–12. https://doi.org/10.1038/srep35241

Law V, Knox C, Djoumbou Y et al (2014) DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res 42:1091–1097. https://doi.org/10.1093/nar/gkt1068

Luck K, Sheynkman GM, Zhang I, Vidal M (2017) Proteome-scale human Interactomics. Trends Biochem Sci 42:342–354. https://doi.org/10.1016/j.tibs.2017.02.006

Menche J, Sharma A, Kitsak M et al (2015) Uncovering disease-disease relationships through the human interactome. Science 347:1257601. https://doi.org/10.1126/science.1257601

Menche J, Guney E, Sharma A et al (2017) Integrating personalized gene expression profiles into predictive disease-associated gene pools. Syst Biol Appl 3:10. https://doi.org/10.1038/s41540-017-0009-0

Mercer SW, Wang J, Burke R (2017) In vivo modeling of the pathogenic effect of copper transporter mutations that cause Menkes and Wilson diseases, motor neuropathy, and susceptibility to Alzheimer's disease. J Biol Chem 292:4113–4122. https://doi.org/10.1074/jbc.M116.756163

Merlini L, Angelin A, Tiepolo T et al (2008) Cyclosporin A corrects mitochondrial dysfunction and muscle apoptosis in patients with collagen VI myopathies. Proc Natl Acad Sci 105:5225–5229. https://doi.org/10.1073/pnas.0800962105

Osman MM, Lulic D, Glover L et al (2011) Cyclosporine-A as a neuroprotective agent against stroke: its translation from laboratory research to clinical application. Neuropeptides 45:359–368. https://doi.org/10.1016/j.npep.2011.04.002

Pavlopoulos GA, Secrier M, Moschopoulos CN et al (2011) Using graph theory to analyze biological networks. BioData Min 4:10. https://doi.org/10.1186/1756-0381-4-10

Piñero J, Berenstein A, Gonzalez-Perez A et al (2016) Uncovering disease mechanisms through network biology in the era of next generation sequencing. Sci Rep 6:1–12. https://doi.org/10.1038/srep24570

Proitsi P, Lupton MK, Velayudhan L et al (2014) Genetic predisposition to increased blood cholesterol and triglyceride lipid levels and risk of Alzheimer disease: a Mendelian randomization analysis. PLoS Med. https://doi.org/10.1371/journal.pmed.1001713

Puglielli L, Tanzi RE, Kovacs DM (2003) Alzheimer's disease: the cholesterol connection. Nat Neurosci 6:345–351. https://doi.org/10.1038/nn0403-345

Qiu Y, Yu T, Wang W et al (2014) Curcumin-induced melanoma cell death is associated with mitochondrial permeability transition pore (mPTP) opening. Biochem Biophys Res Commun 448:15–21. https://doi.org/10.1016/j.bbrc.2014.04.024

Reed B, Villeneuve S, Mack W et al (2014) Associations between serum cholesterol levels and cerebral Amyloidosis. JAMA Neurol 71:195. https://doi.org/10.1001/jamaneurol.2013.5390

Schadt EE (2009) Molecular networks as sensors and drivers of common human diseases. Nature 461:218–223. https://doi.org/10.1038/nature08454

Seebacher J, Gavin A-C (2011) SnapShot: protein-protein interaction networks. Cell 144:1000–1000.e1. https://doi.org/10.1016/j.cell.2011.02.025

Sharma A, Menche J, Chris Huang C et al (2015) A disease module in the interactome explains disease heterogeneity, drug response and captures novel pathways and genes in asthma. Hum Mol Genet 24:3005–3020. https://doi.org/10.1093/hmg/ddv001

Strauss KA, DuBiner L, Simon M et al (2013) Severity of cardiomyopathy associated with adenine nucleotide translocator-1 deficiency correlates with mtDNA haplogroup. Proc Natl Acad Sci 110:3453–3458. https://doi.org/10.1073/pnas.1300690110

Thompson K, Majd H, Dallabona C et al (2016) Recurrent de novo dominant mutations in SLC25A4 cause severe early-onset mitochondrial disease and loss of mitochondrial DNA copy number. Am J Hum Genet 99:860–876. https://doi.org/10.1016/j.ajhg.2016.08.014

Tosserams A, Papadopoulos C, Jardel C et al (2017) Two new cases of mitochondrial myopathy with exercise intolerance, hyperlactatemia and cardiomyopathy, caused by recessive SLC25A4 mutations. Mitochondrion. https://doi.org/10.1016/j.mito.2017.08.009

Vaya J, Schipper HM (2007) Oxysterols, cholesterol homeostasis, and Alzheimer disease. J Neurochem 102:1727–1737. https://doi.org/10.1111/j.1471-4159.2007.04689.x

Vidal M, Cusick ME, Barabási A-L (2011) Interactome networks and human disease. Cell 144:986–998. https://doi.org/10.1016/j.cell.2011.02.016

Zandl-Lang M, Fanaee-Danesh E, Sun Y et al (2017) Regulatory effects of simvastatin and apoJ on APP processing and amyloid-β clearance in blood-brain barrier endothelial cells. Biochim Biophys Acta Mol Cell Biol Lipids. https://doi.org/10.1016/j.bbalip.2017.09.008

7. Veröffentlichung II: iBRET screen of the ABCD1 peroxisomal network and mutation-induced network perturbations.

https://doi.org/10.1021/acs.jproteome.1c00330

Journal of proteome research

pubs.acs.org/jpr

Article

iBRET Screen of the ABCD1 Peroxisomal Network and Mutation-Induced Network Perturbations

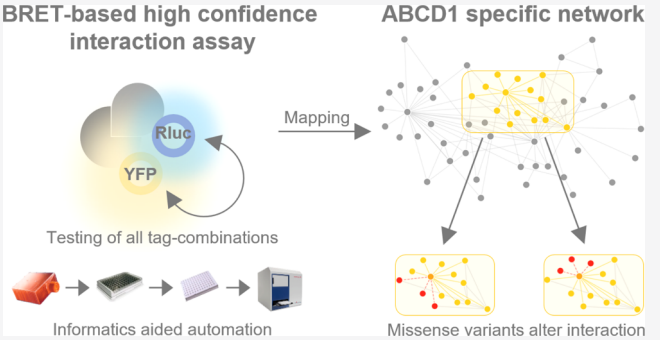
Amelie S. Lotz-Havla,* Mathias Woidy, Philipp Guder, Caroline C. Friedel, Julian M. Klingbeil, Ana-Maria Bulau, Anja Schultze, Ilona Dahmen, Heidi Noll-Puchta, Stephan Kemp, Ralf Erdmann, Ralf Zimmer, Ania C. Muntau, and Søren W. Gersting*

Cite This: https://doi.org/10.1021/acs.jproteome.1c00330

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Mapping the network of proteins provides a powerful means to investigate the function of disease genes and to unravel the molecular basis of phenotypes. We present an automated informatics-aided and bioluminescence resonance energy transfer-based approach (iBRET) enabling high-confidence detection of protein−protein interactions in living mammalian cells. A screen of the ABCD1 protein, which is affected in X-linked adrenoleukodystrophy (X-ALD), against an organelle library of peroxisomal proteins demonstrated applicability of iBRET for large-scale experiments. We identified novel protein−protein interactions for ABCD1 (with ALDH3A2, DAO, ECI2, FAR1, PEX10, PEX13, PEX5, PXMP2, and PIPOX), mapped its position within the peroxisomal protein−protein interaction network, and determined that pathogenic missense variants in *ABCD1* alter the interaction with selected binding partners. These findings provide mechanistic insights into pathophysiology of X-ALD and may foster the identification of new disease modifiers.



KEYWORDS: *bioluminescence resonance energy transfer, BRET, protein−protein interaction, screening, interactome, lipid droplets, FAR1, living cells, ABCD1, X-ALD, fatty acids*

INTRODUCTION

A mechanistic understanding of biological processes emerges from studying how proteins interconnect with each other and form a dynamic biological network that drives cellular function: the human interactome.[1−5] Disease phenotypes, in particular in genetic disorders, often arise from perturbations caused in the human interactome.[5,6] As a consequence, detailed analyses of these perturbations will guide future drug development.[7]

The human interactome of protein−protein interactions is estimated to comprise 130,000−1,000,000 binary interactions.[8−10] The recently published first draft of the human reference interactome utilized a set of different yeast-two-hybrid assays to uncover 52,569 interactions between 8275 proteins.[11] When the size of this reference interactome is related to the abovementioned estimates, it becomes evident that the majority of protein interactions remains to be discovered.

Over the past decades protein interaction mapping revealed that a full interactome coverage is only achievable by combining different experimental methodologies, all bearing specific advantages and disadvantages. Depending on the method applied, the experimental setup may cause artifacts such as alterations of subcellular localization, alterations of post-translational modification, tag-induced steric hindrances, or conformational changes.[12−18] Besides the development of new experimental methods, it has been suggested that existing assays should be improved by increasing the number of different assay setups.[19]

To address this issue, we established an automated platform for the detection of protein−protein interactions based on bioluminescence resonance energy transfer (BRET) with the aim to provide both high confidence data and reasonable throughput. Since its first description in 1999,[20] several versions of BRET using different substrates and varying combinations of energy acceptors and donors have been used.[21−23] Despite its potential for application in screening approaches, BRET-based high-throughput studies focused on protein−protein interactions as drug targets[23−29] rather than addressing the identification of protein−protein interactions at

Received: April 21, 2021

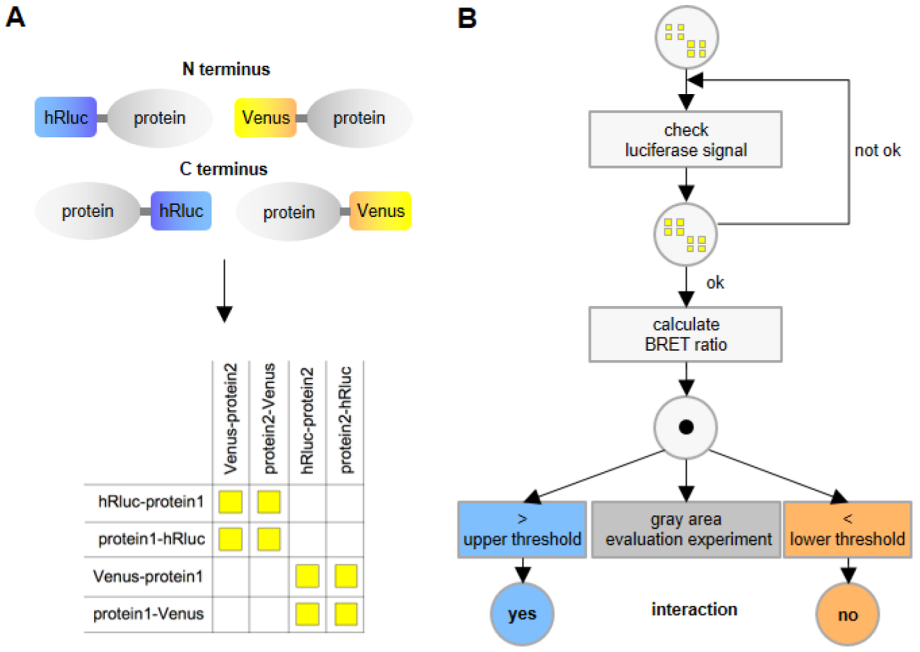


Figure 1. Strategy of an automated BRET-based method for the detected of protein interactions in living cells. (A) Each protein of interest is N- and C-terminally tagged to a humanized version of Renilla luciferase (hRluc) and a stabilized variant of yellow fluorescent protein (Venus). Testing all tag combinations per protein pair results in eight BRET experiments, depicted by the matrix. (B) Bioinformatics-based approach for high-quality and high-throughput iBRET. The process flow integrates proteomics into a multistep bioinformatics algorithm. First, the quality of the hRluc signal is analyzed. For hRluc signals above a defined cutoff, a ratio between the Venus signal and the hRluc signal is calculated (BRET ratio). Interactions are subsequently classified based on the BRET ratio. If the BRET ratio exceeds the upper threshold for at least one tag combination, the interaction is classified as positive; if the BRET ratios for all tag combinations are below the lower threshold, the interaction is classified as negative; and if no BRET ratio exceeds the upper threshold and at least one BRET ratio exceeds the lower threshold (gray area), an evaluation experiment is performed. This procedure is applied in the iBRET software package when "classification of measured pairs" is performed (Supporting Information Material S1). Cutoff and thresholds were determined during the "training" method of the iBRET software on a carefully designed gold-standard set (Supporting Information Figure S6 and Supporting Information Material S1). Note that cutoff and thresholds are specific to the PRS, RRS, and instrumental setup of this study. The iBRET software was designed to be generally applicable to BRET screens independent of reference data sets and technical variation.

a large scale. The process flow of our experimental setup integrates techniques suitable for high-throughput (recombinational cloning and multiwell electroporation), an automated *in vivo* BRET protein−protein interaction assay, and bioinformatics operations (informatics-based BRET, iBRET). To achieve high accuracy, iBRET by default tests eight different combinations of BRET tags.

We applied iBRET to map a specific interactome of a disease gene product. The *ABCD1* gene encodes a peroxisomal transporter that imports very long-chain fatty acids (VLCFAs) for β-oxidation.[30,31] A deficiency of the ATP-binding cassette transporter ABCD1 leads to the accumulation of VLCFAs resulting in X-linked adrenoleukodystrophy (X-ALD, ORPHA:43). X-ALD is the most common peroxisomal disorder and one of the most puzzling and devastating inborn errors of metabolism affecting the nervous system. X-ALD is characterized by a highly divergent clinical picture, ranging from childhood cerebral adrenoleukodystrophy and chronic progressive adult-onset adrenomyeloneuropathy (AMN) to cerebral AMN.[32] The same variants in *ABCD1* can lead to different clinical phenotypes and the underlying mechanisms of the disease are poorly understood.[33−35] Despite its function as a peroxisomal transporter, we previously showed that ABCD1 is part of a cytosolic but peroxisome-associated multiprotein complex involved in fatty acid synthesis (ACLY, FASN, and ACC) and fatty acid activation (FATP4).[36] In addition, recent investigations showed that ABCD1 forms a tethering complex to maintain peroxisome-lipid droplet membrane contact sites.[37] These contact sites are important for different cellular and metabolic pathways.[38] Taken together, these findings suggest additional functions of ABCD1 and facilitate new investigations to enhance our understanding of X-ALD pathology. By means of our iBRET technique and subsequent investigations of mutation-induced disruption of the ABCD1 protein−protein interactions, we demonstrate the applicability of our approach in a screening setup. Moreover, we provide further evidence for an extended function of ABCD1, specifically in maintaining the tight connections between peroxisomes and lipid droplets.

EXPERIMENTAL SECTION

Plasmids

For iBRET, Gateway destination vectors were designed based on the pcDNA 6.2. DEST general Gateway vector (Invitrogen). To generate BRET destination vectors with different tag orientations, the coding sequence of hRluc (phRG-TK) and Venus (pEYFP-N1),[39] modified as previously described (Supporting Information Table S1), was amplified without stop codon and fused 5′ in frame with the Gateway ccdB-cassette for the generation of N-terminal fusion proteins or with stop codon 3′ in frame with the ccdB-cassette for the generation of C-terminal fusion proteins, respectively.

Full-length ORFs without and with stop codons were obtained as Gateway Entry clones from the PlasmID database

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

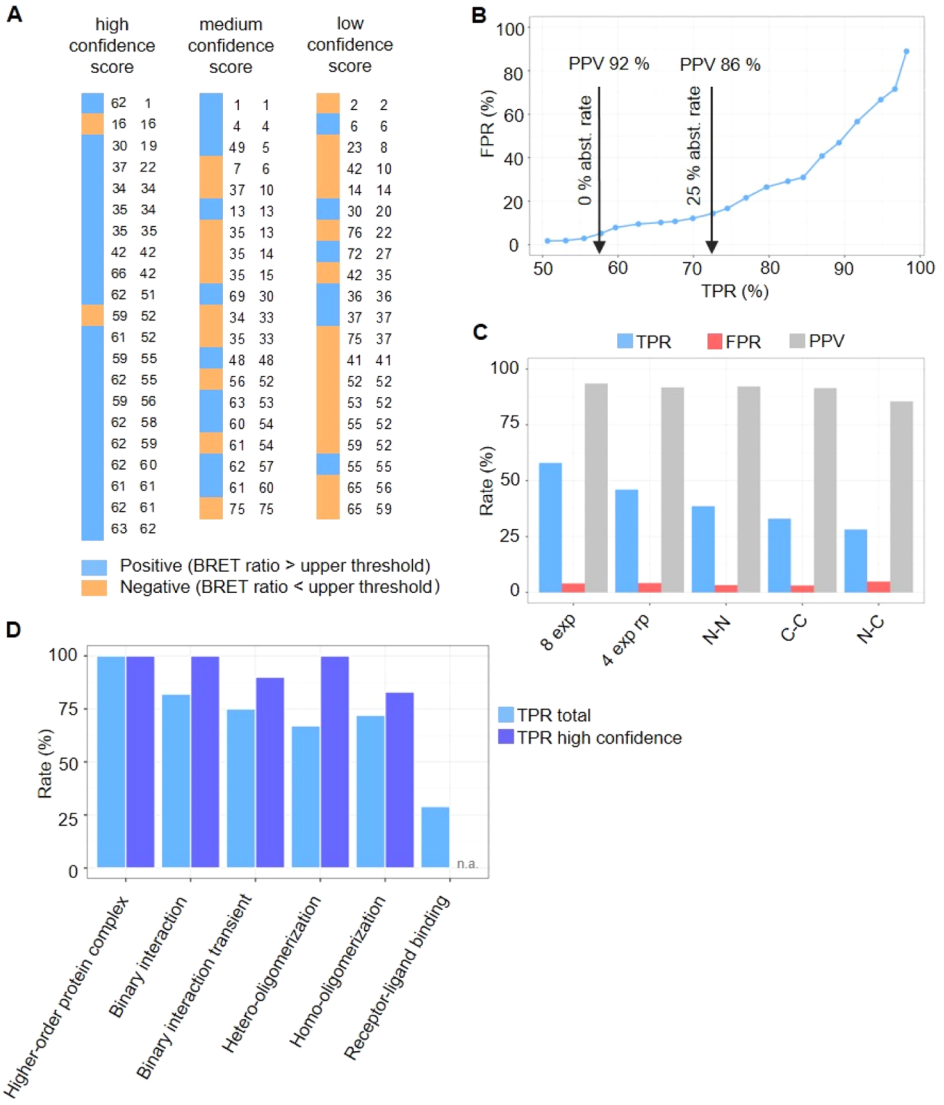


Figure 2. Analysis of iBRET assay performance. (A) Assay performance without evaluation experiments according to different confidence scores of interactions in the PRS. Positive and negative results are indicated, and numbers denote ID of tested proteins (Supporting Information Table S6). (B) Comparison of the TPR and FPR of the iBRET approach using the gray area (25% abstention rate) against an approach using only one threshold to classify interactions as positive. Evaluation of all possible values for this threshold showed a reduced PPV and increased FPR compared to the method using the gray area. The TPR using only the upper threshold determined by cross-validation is also indicated (0% abstention rate). (C) Performance of iBRET assays without evaluation experiments tested on all eight combinations of tag orientations (8 exp) and compared to different tag combinations: randomly chosen four combinations (4 exp rp), N-terminal only (N−N), C-terminal only (C−C), and N−C-terminal (N−C). (D) Performance of iBRET assays with evaluation experiments tested for different classes of protein interactions with respect to the confidence score. For receptor-ligand binding PPI, no data with a high confidence score were available (n.a.).

(http://plasmid.med.harvard.edu/PLASMID/) and from a copy of the mammalian gene collection or amplified by single-step PCR from target open reading frames and introduced into an appropriate Gateway Donor vector. Sequences of full-length ORFs (see Supporting Information Table S6 for a list of all coding genes) were verified by DNA sequence analysis (Eurofins MWG Operon). Mutations in *ABCD1* were introduced by conventional site-directed mutagenesis. BRET expression clones were generated by Gateway-based recombination of Gateway Entry clones for genes of interest without stop codon and the C-terminally fused (*hRluc*/*Venus*) destination vectors or Gateway Entry clones with stop codon and the N-terminally fused destination vectors, respectively, resulting in four expression vectors for each protein of interest. To generate a BRET control plasmid coding for a Venus−hRluc fusion protein, the coding sequence of Venus was amplified without stop codon by single-step PCR, introduced into a Gateway Donor vector, which was then recombined with the C-terminally fused *hRluc* destination vector.

For LUMIER, BRET expression vectors carrying N-terminal hRluc-tag were exploited. His-tagged expression clones were generated by Gateway-based recombination of Entry clones for respective genes of interest with stop codon and the general Gateway destination vector pEF-DEST 51 carrying N-terminal V5/6xHis (Invitrogen).

For BIFC, Gateway destination vectors were designed based on the MBP pCR3.1 rf B (Invitrogen) Gateway vector. The coding sequence of YFP fragments (YFP1 and YFP2, Nyfeler *et al.* 2005 PNAS)[40] carrying linker sequences containing a Kozak

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

sequence was amplified without stop codon and fused 5′ in frame with the Gateway ccdB-cassette for the generation of N-terminal fusion proteins. Expression clones were again generated by Gateway-based recombination of Gateway Entry clones for genes of interest with stop codon and the destination vectors. Plasmids were transformed into *Escherichia coli* DH5 alpha cloning strains, and DNA was purified using the PureYield plasmid maxiprep system (Promega).

Curation of Reference Data Sets PRS and RRS

The reference datasets were based on a plasmid library. Sixty-one database entries for positive protein−protein interactions (positive reference set, PRS) of our plasmid library were derived from PubMed and from HPRD, STRING, BIND, and BioGRID databases. All curation reports were verified and no filters were applied; hence, interactions between membrane proteins, ligand−receptor pairs, and interactions depending on post-translational modification were all included (Supporting Information Figures S4 and S5). A confidence score was manually defined based on a literature review, with a confidence score of three representing interactions confirmed by at least three complementary methods or in at least three independent studies (n = 21), a confidence score of two representing interactions confirmed by two complementary methods or in two independent studies (n = 20), and a confidence score of one representing interactions only shown by one method in one study (n = 20) (Figure 2A and Supporting Information Table S4). The random reference set (RRS) was composed of 60 randomly drawn protein pairs of our cDNA library (Figure 2A and Supporting Information Table S5) with random drawing executed by a Perl script. First, two proteins were independently drawn from the library and redrawn if a positive interaction was listed in databases. Next, the cellular component of each protein was checked based on the Gene Ontology database,[41] and the pair was redrawn if proteins carried similar gene ontology tags with this respect. If a randomly drawn protein pair passed the first two steps, it was added to our RRS. This procedure was repeated until 60 randomly drawn protein pairs were collected.

Cell Culture

COS-7 cells and HEK293 cells were cultured under conventional monolayer cell culture conditions (37 °C, 5% CO_2) in HYPERFlask cell culture vessels (Corning).

iBRET Process

The whole process of DNA preparation, 96-well electroporation, incubation, and luminescence detection was integrated on an automated liquid handling platform (Tecan Freedom EVO200). Transfection of cells for BRET experiments as part of the iBRET process was performed in 96-well format, as previously described in detail.[42] COS-7 cells were cotransfected by electroporation (Amaxa 96-well shuttle system, Lonza, see Supporting Information Table S3 for transfection of other cell lines) with the respective two genes of interest either fused to *hRluc* (energy donor) or *Venus* (energy acceptor). Cells were transfected at an acceptor to donor ratio of 3:1 with a total of 1 μg DNA for all tag combinations. For BRET evaluation experiments, increasing acceptor to donor ratios at a constant donor level (0.15 μg DNA) were applied.[43−45] All liquid handling steps for DNA preparation and electroporation were performed by independent 8-channel and 96-channel pipetting arms of the liquid handling platform. Plates were moved by a robotic manipulator arm. Subsequent to electroporation, cells were transferred to white-wall 96-well cell culture plates (Corning) and automatically stored in a 44 position incubator (Liconic STR44). After 24 h incubation at 37 °C, the cell culture medium was replaced by fresh medium and plates were automatically transferred to a 96-well microplate luminometer (LUMIstar OPTIMA, BMG Labtech). In this device, ViviRen (Promega) was added to the living cells (60 μM) by automated injection and light emission was collected for 10 s at 475 nm (hRluc signal) and 535 nm (BRET signal).

In order to optimize speed and reliability of the process (Supporting Information Table S2), an operating procedure was designed based on the EVOware software (Tecan) with integration of additional visual basic scripts to control external devices (electroporation and data detection), allowing for simultaneous execution of multiple processes. Here, the different processes of electroporation, incubation, and luminescence data collection were scheduled yielding similar treatment of all 64 plates within a 1 week run (Supporting Information Figure S2). All data are automatically evaluated (see iBRET algorithm), and raw data as well as results are saved to a MySQL database. A lab intern webinterface enables users to view and further explore our interaction data. Integration of BioGRID, OMIM, and Cytoscape additionally supports the analysis of results.

Calculation of BRET Ratios

The BRET ratio R was calculated as $R = I_A/I_D - cf$, where I_A is the intensity of light emission at 535 nm, I_D is the intensity of light emission at 475 nm, and cf is a correction factor $(I_A/I_D)_{control}$ with the control being the expression of a hRluc fusion protein in the absence of a Venus acceptor.

iBRET Approach

Binary protein−protein interactions were tested in duplicates and three individual sets of experiments for all possible combinations of two proteins of interest either N- or C-terminally fused to hRluc or Venus yielding eight different pairs of fusion proteins.

Controls were included on every 96-well transfection plate and consisted of a plasmid coding for a Venus−hRluc fusion protein, an hRluc BRET expression clone (hRluc-PEX19) in the absence of a Venus acceptor giving the correction factor (cf), and a background control with nontransfected cells. The suitability of hRluc as an energy donor for BRET was confirmed, as it does not dimerize with Venus (Supporting Information Figure S1B).[44,46]

Upon data acquisition, data were first analyzed with regard to the controls. As a positive control, the Venus−hRluc fusion protein was expected to result in BRET ratios of ∼1, and the cf representing bleed through of I_D into I_A should not exceed a ratio of 0.5. If variation of more than 10% of these values was observed, analysis of the respective plates was repeated. Raw data were then analyzed with regard to the hRluc signal of every individual tag combination per protein pair. A BRET ratio was calculated only for those tag combinations exceeding the cutoff determined using cross-validation (13,884 RLU, see below). The BRET ratios for each combination were determined as an average of three independent experiments. For classification of interactions, BRET ratios of all eight possible combinations (four in the case of self-interactions) were compared to upper (0.0915, see below) and lower (0.0373, see below) thresholds defining the gray area. Interactions were classified as positive if the BRET ratio of

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

at least one tag combination exceeded the upper threshold; if BRET ratios for all tag combinations were below the lower threshold, the interaction was classified as negative; and if no BRET ratio exceeded the upper threshold but the BRET ratio of at least one tag combination was at least as high as the lower threshold, the interaction was assigned to the gray area and a BRET evaluation experiment was performed (see Figure 1B and Supporting Information Figure S3).

Identification and Evaluation of Thresholds for the iBRET Method

Cutoffs to determine hRluc signals of insufficient intensity, as well as the thresholds to classify an interaction as positive (upper threshold of the gray area), and to classify an interaction as negative (lower threshold of the gray area) were calculated from the PRS/RRS pairs as follows. First, the smallest hRluc signal cutoff was determined, such that the number of positive (=PRS) interactions with a BRET ratio higher than any negative (=RRS) interaction was maximized (13,884 RLU). For this purpose, only average values between two consecutive hRluc signal values in the ordered sequence of all observed values were considered. This reduced runtime, and any other possible threshold would yield the same results in the training step as one of the thresholds considered. The upper threshold was then set as the average of the highest BRET ratio of a negative interaction and the next higher BRET ratio of a positive interaction. The lower threshold was then determined, such that a predefined fraction of interactions (25%) falls within the gray area. This is the only parameter of the training method. As possible lower thresholds, only the average between two consecutive BRET ratios was evaluated in the sorted order of BRET ratios below the upper threshold for the given hRluc cutoff.

Evaluation of threshold determination (Supporting Information Figure S6) was performed using 10-fold cross-validation repeated 100 times (Supporting Information Material S1). For this, we utilized the same data sets (PRS/RRS) as for the threshold determination. Final lower and upper thresholds (0.0373 and 0.0915, respectively) and luciferase signal cutoff (13,884 RLU) were determined as the median values obtained across all cross-validation repeats. The number of true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), positive interactions abstained (PA), and negative interactions abstained (NA) was calculated separately for each cross-validation and averaged across repeats.

Evaluation of iBRET Performance

iBRET performance was estimated calculating the true positive rate (TPR), the false positive rate (FPR), and the positive predicted value (PPV). TPR is defined as the fraction of positive interactions, which can be recovered and was used to assess the sensitivity of the method [TPR = TP/(TP + FN + PA)]. The FPR is defined as the fraction of negative interactions predicted wrongly as positive interactions [FPR = FP/(TN + FP + NA)]. This was used to assess the specificity of the approach (specificity = 1 − FPR). The PPV is defined as the fraction of correct predictions among the predicted interactions, that is, the probability that an identified interaction is actually a true interaction [PPV = TP/(TP + FP)].

Application of iBRET

Overall experiments were performed as mentioned in the method section for the iBRET process. However, in terms of customization, experiments were performed in HEK293 (see Supporting Information Table S3). Cells were transfected at an acceptor to donor ratio of 3:1 with a total of 0.6 μg DNA for all tag combinations. For BRET evaluation experiments, increasing acceptor to donor ratios at a constant total DNA amount (2 μg DNA) were applied.[43−45] As a substrate, coelenterazine (PJK) was added to the living cells (30 μM). For luminescence detection a by then optimized, most sensitive multimode reader enabling simultaneous dual emission was applied (PHERAstar, BMG Labtech). Because the experimental conditions were thus changed, the reference dataset was measured again and lower and upper BRET thresholds (0.026 and 0.0427, respectively) and luciferase signal cutoff (21,550 RLU) were determined as described above. For experiments of screens II and III, cells were incubated for 48 h at 30 °C subsequent to electroporation. For these changed conditions, a detection limit for the luciferase signals (dl) was calculated as follows

$$dl = \mathrm{mean}(blank) + 9^{*}\mathrm{sd}(blank)$$

where *blank* is the detected hRluc signal of nontransfected cells, mean(*blank*) is the average hRluc signal across all experiments, and sd(*blank*) is the respective standard deviation.

BiFC

COS-7 cells were cotransfected by electroporation (Amaxa 96-well shuttle system, Lonza) with the respective two genes of interest either N-terminally fused to *YFP* fragment 1 or *YFP* fragment 2. 200,000 cells were transfected at a ratio of 1:1 with a total of 0.8 μg DNA. After 48 h incubation at 30 °C, the YFP signal was detected in a 96-well microplate reader (PHERAstar OPTIMA, BMG Labtech) at 535 nm. After subtracting the background, signals were normalized to a positive control (bJun-bFos[47]). The resulting ratio was then classified as positive, if it exceeded at least the normalized sum of mean background signals and threefold background standard deviation for each corresponding experiment. As a negative control, cells were transfected with respective *YFP1* or *YFP2* tagged gene of interest alone.

LUMIER

COS-7 cells were cotransfected by electroporation (Amaxa 96-well shuttle system, Lonza) with the respective two genes of interest either N-terminally fused to *6xHis* or *hRluc*. 200,000 cells were transfected at a ratio of 1:2 with a total of 0.3 μg DNA. After 48 h incubation at 30 °C, cells were lysed with IGEPAL CA-630 (Sigma-Aldrich) and immunoprecipitated by application of Protino Ni-NTA (Machery-Nagel). Immunoprecipitation was done by means of an automated protocol set up on a liquid handling platform (Tecan Freedom EVO200). Subsequent 30 μM coelenterazine was added to the immunoprecipitate, and hRluc signals were measured in a 96-well microplate reader (PHERAstar OPTIMA, BMG Labtech) at 475 nm. Signals were corrected by subtraction of the mean background signal and normalized to the signal of a positive control (homodimerization of PAH, Supporting Information Table S7). As a negative control, cells were transfected with respective hRluc or His tagged gene of interest alone.

Coexpression Analysis

Data for coexpression analysis were obtained from The Human Protein Atlas as of 25th of October 2020.[48] To retrieve blood

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

cell-specific transcription levels, we downloaded the provided .tsv file from the section "RNA HPA blood cell gene data", which is based on The Human Protein Atlas version 19.3 and Ensembl version 92.38. We calculated z-scores to better compare the NX expression values of different genes. To quantify the expression of gene i in a blood cell b, the average expression mean$(E(i))$ and the standard deviation sd$(E(i))$ of a gene's expression across all considered blood cells were calculated. The z-score is defined as

$$z\text{-score}(i, b) = (E(i, b)) - \text{mean}(E(i))/\text{sd}(E(i))$$

A z-score >0 for gene i in blood cell b indicates that the gene's expression is higher compared to its expression across all other blood cells. Note that a z-score <0 does not necessarily mean that the gene is not expressed in this cell type.

RESULTS

Automation of the BRET Assay in Living Cells

The iBRET method setup was designed in consideration of adaptability toward the input (gene of interest) and read-out (static or dynamic information) and suitability for straightforward implementation. A recombinational cloning strategy ensured rapid integration of genes of interest into the BRET system (Supporting Information Figure S1A). Gateway entry clones obtained from public repositories (PlasmID database, mammalian gene collection) or cDNAs amplified *via* single-step PCR using target open reading frames from different sources were introduced into BRET expression constructs. In order to yield the optimal luminescence signal and BRET efficiency, we applied the optimized BRET[1] system[42,49] with luciferase (hRluc) as an energy donor[24] and a stabilized variant of yellow fluorescent protein (Venus) as an energy acceptor[39] (Figure 1A, Supporting Information Table S1). In order to minimize tag-induced steric hindrances and conformational alterations, each protein of interest was fused to N-terminal and C-terminal hRluc- and Venus-tags, respectively. This resulted in four different BRET expression vectors per individual protein and eight combinations of BRET experiments per protein pair (Figure 1A, Supporting Information Figure S1B). Physiological subcellular distribution of constructs carrying BRET-tags has been demonstrated before.[50,51]

For the biotechnological implementation of a fully automated BRET assay, a 96-cuvette electroporation device, an automated incubator, and a luminescence multiwell reader were integrated into a liquid handling platform (Supporting Information Figure S1C). We included a scheduling software module to take into account the 24 h incubation time following transfection, which enabled an interlaced operation of 64 plates per week in 8 runs of 8 individual plates per run (Supporting Information Figure S2 and Table S2). This setup allowed for a throughput of 5120 BRET experiments per week, which correspond to 640 individual protein pairs without duplicate testing. In the present study, COS-7 cells were cultured in reusable 10-layer flasks yielding more than 1.0×10^9 cells. However, any adherent or suspension cell line can be utilized if high transfection efficiency and post-transfection viability are attainable (Supporting Information Table S3).

iBRET Process and Assay Performance

The iBRET approach combines automated interaction screening of a target protein library with a bioinformatics procedure for integrated assessment of process quality and classification of test results. The bioinformatics procedure identified clearly interacting or clearly noninteracting protein pairs and undecided pairs for subsequent retesting (gray area) (Figure 1B). First, the algorithm categorized the raw data of the BRET screen (Supporting Information Table S7) with regard to hRluc signals of all individual tag combinations per protein pair to exclude tag combinations below a method-specific cutoff. Second, the BRET ratio, which is the control corrected ratio of light intensity at 535 nm (Venus) and 475 nm (hRluc), classified a protein pair as interacting (positive), if the BRET ratio of at least one tag combination per protein pair exceeded an upper threshold. A protein pair was classified as non-interacting (negative), if the BRET ratio of all tag combinations per protein pair was below a lower threshold. Third, the interval between upper and lower thresholds defined a gray area containing the uncertain candidates for retesting in BRET saturation experiments (hereafter referred to as evaluation experiments). Evaluation experiments allow us to distinguish positive and negative interactions. Positive interactions show a nonlinear hyperbolic behavior. Negative interactions show a linear increase in BRET ratios, which is caused by random collision of proteins (bystander BRET).[44,52] The evaluation experiments were performed using varying ratios of hRluc (donor) over Venus (acceptor) tagged protein pairs, and BRET ratios were analyzed as a function of the acceptor to donor ratio (Supporting Information Figure S3). To test a sufficient expression of BRET constructs and reliability of the luciferase assay and to check variation in luminescence detection, control experiments on each plate were analyzed upon data acquisition.

For parameter training and evaluation of assay performance, we compiled two data sets consisting of documented known interactions PRS and a RRS, covering seven different subcellular compartments (Supporting Information Tables S4−S6 and Figures S3 and S4). All 61 interactions of the PRS were derived from databases and manually curated to ensure high reliability of the data set.[53] For lack of a gold standard data set for noninteracting proteins,[54,55] the RRS was generated based on the cDNA library of our PRS. Sixty protein pairs were drawn randomly, and protein pairs with interactions documented in databases were excluded. This approach is well accepted to create control data sets.[53,56]

All protein pairs in the PRS and RRS were assayed in three independent runs of iBRET screening, testing eight combinations of BRET tags, and BRET ratios for the respective tag combinations were averaged across repeats. Repeated 10-fold cross-validation (100 repeats) on the PRS and RRS was then used to determine the optimal cutoff for hRluc signals, the thresholds for BRET ratios, and the assay performance (Supporting Information Figure S6, Supporting Information Material S1). The TPR, or sensitivity, differed markedly between high-confidence (TPR, 90%) versus low-confidence (TPR, 30%) interactions of the PRS (Figure 2A, Supporting Information Table S4). When a setup without the gray area was applied (0% abstention rate) to the full PRS, the TPR was 58%, the FPR was 4%, and the PPV was 94% (Figure 2B, Supporting Information Figure S7). In contrast, the TRP was 70.1%, the FPR was 3.5%, and the PPV was 93.2% when only high- and medium-confidence interactions were considered. Thus, calculation of assay performance strongly depends on quality and reliability of the reference data sets used.[12,56−59]

When the gray area was set to an abstention rate of 25% for the full PRS, 30 interactions resulted in BRET ratios between the upper and lower thresholds. BRET saturation evaluation

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

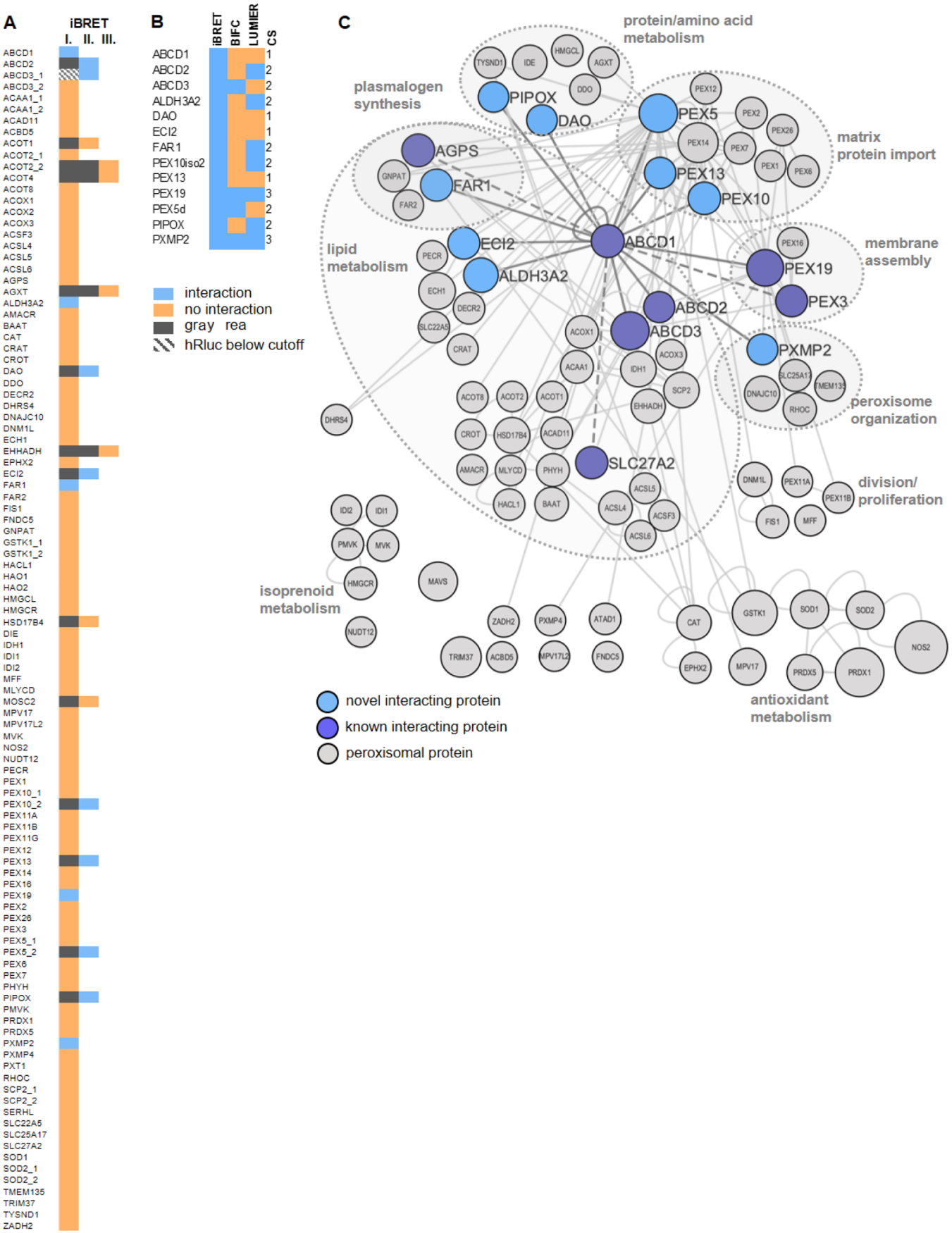


Figure 3. Application of iBRET to a whole-organelle peroxisomal screen. (A) BRET-based interaction screen of ABCD1 against a peroxisomal protein library (n = 96, additional eight isoforms, Supporting Information Table S9) confirms four known interactions (Supporting Information Table S10) and identifies nine novel interactions. Interactions with hRluc signals below the cutoff in all eight tested tags are indicated in gray-white-

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

Figure 3. continued

stripped. First screen performed at 37 °C (I); second screen performed at 30 °C (II); and BRET saturation evaluation experiments (III) (Supporting Information Figure S7). (B) Confirmation experiments for ABCD1 interactions by BiFC and LUMIER. All 13 interactions of ABCD1 found in the iBRET screen were tested with BiFC and LUMIER. Results are depicted for iBRET, BiFC, and LUMIER. The confidence score (cs) is given depending on the number of methods used in this study (iBRET, BiFC, and LUMIER) confirming the interaction. (C) Interaction network of the peroxisomal interactome. The emphazised ABCD1 subnetwork (highlighted gray edges) comprises all nodes (proteins) and edges (interactions) that are directly linked to ABCD1. All other peroxisomal proteins are depicted with their respective known interactions (thin gray edges). The size of a node depicts the number of interaction partners of this node within the human interactome. Proteins are grouped with respect to peroxisomal function.

experiments correctly confirmed 9 of 14 positive and 11 of 16 negative interactions. Application of this setup yielded a TPR of 72%, a FPR of 12%, a PPV of 86%, and a specificity (true negative rate) of 88%. Thus, allowing 25% of protein pairs to be retested increased the TPR from 58 to 72% at the cost of an increase in the FPR from 4 to 12%.

Integrated testing of eight combinations of N-terminal and C-terminal BRET tags requires a higher number of BRET experiments per individual interaction as compared to a standardized N−N tagging, for instance. Therefore, we evaluated all possible setups of N-terminal and C-terminal BRET-tag combinations with respect to the TPR, FPR, and PPV (Figure 2C). The highest TPR was observed using the iBRET test protocol with eight combinations ($P < 0.01$), whereas no marked differences were observed for the FPR and PPV among different setups of tag combinations. A random selection of four combinations of N- or C-terminal tagging resulted in a higher TPR as compared to fixed single N−N, C−C, or N−C setups. In conclusion, the TPR correlated with the number, but not orientation, of tested tag combinations.

Biological Aspects of Assay Performance

Assessment of iBRET performance with respect to different classes of protein−protein interactions (Figure 2D, Supporting Information Table S8) revealed a TPR of 100% (TP = 2/2) for proteins involved in higher-order protein complexes. For binary interactions, the TPR was 82% (TP = 9/11), for transient binary interactions, the TPR was 75% (TP = 15/20), and hetero- or homo-oligomerization was detected with TPRs of 67% (TP = 2/3) and 72% (TP = 13/18). For all classes of interactions mentioned, the TPR was markedly higher when the evaluation was confined to protein−protein interactions assigned to a high confidence score [binary interactions, 100% (TP = 2/2); transient binary interactions, 90% (TP = 9/10); hetero-oligomerization, 100% (TP = 2/2); homo-oligomerization, 83% (TP = 5/6)]. The TPR for the detection of receptor−ligand binding was 29% (TP = 2/7); however, none of these interactions was assigned to a high confidence score (Supporting Information Table S4). Taken together, iBRET shows high sensitivity for the detection of all tested classes of protein−protein interactions, but not for receptor−ligand binding. Sequestration of ligands to the extracellular space would result in a loss of this protein due to a washing step in the current protocol and may add to the low TPR observed for this class of interaction.

The BRET method has particularly proven useful for the study of membrane proteins.[60] We have also studied the performance of our approach in isolation for protein pairs in which both partners are denoted by a gene ontology tag for the membrane; the TPR was 79% (23/29).

Screening ABCD1 against a Whole-Organelle Peroxisomal Library

Next, we demonstrated the applicability of iBRET for *de novo* detection of protein−protein interactions, by screening the gene product of a disease gene against a target protein library. To investigate the position of ABCD1 in the context of larger functional networks, we mapped the ABCD1 protein−protein interaction network within the peroxisome. We tested ABCD1 against a library covering a large proportion of all proteins annotated with peroxisomal localization (n = 96 proteins, 8 additional isoforms, Supporting Information Table S9).[41,61,62] Based on our iBRET setup, all 104 protein pairs were tested for eight tag combinations. In addition, the ABCD1 homo-oligomerization was tested for four tag combinations. All protein pairs were tested in duplicates resulting in 1672 BRET experiments. This process required automated pipetting of 21 assay plates in 5 runs, which was performed in 3 days. In addition, we cataloged all known interactions of ABCD1 with peroxisomal proteins based on a combined database and literature research and assigned these interactions to a confidence score of 1−3 according to the number of studies and/or methods linked to database entries (n = 7 interactions, Supporting Information Table S10). In the first round of screening, protein−protein interactions were assayed in all eight possible tag combinations (Figure 3A). We confirmed the known ABCD1 homo-oligomerization and the interaction with PEX19. In addition, three previously unknown interactions of ABCD1 with ALDH3A2, FAR1, and PXMP2 were identified. Following our bioinformatics-based iBRET approach (Figure 1B), all tested protein pairs with hRluc signals above the cutoff and BRET ratios below the lower threshold (n = 85) were classified as negative. All protein pairs with hRluc signals above the cutoff and BRET ratios below the upper threshold but above the lower threshold (n = 14) were assigned to the gray area. However, one protein pair remained below the hRluc cutoff, precluding a reliable classification. Low hRluc signals point toward inefficient expression of the respective protein of interest. Expression and trafficking of peroxisomal proteins in cultured cells often benefit from reduction of the incubation temperature to 30 °C.[63] Hence, a second screen at 30 °C (Figure 3A) retested the interaction of ABCD1 with the protein that displayed hRluc signals below the cutoff and all protein pairs that were assigned to the gray area. This confirmed interactions of ABCD1 with ABCD2 and ABCD3 and classified six previously unknown interactions as positive (DAO, ECI2, PEX10, PEX13, PEX5, and PIPOX). For protein pairs assigned to the gray area in the second screen (n = 4), BRET evaluation experiments were performed in a third screen (Figure 3A, Supporting Information Figure S8). None of these protein pairs was classified as a positive protein−protein interaction in the third screen. By means of iBRET analyses in living cells, we confirmed four out of five known

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

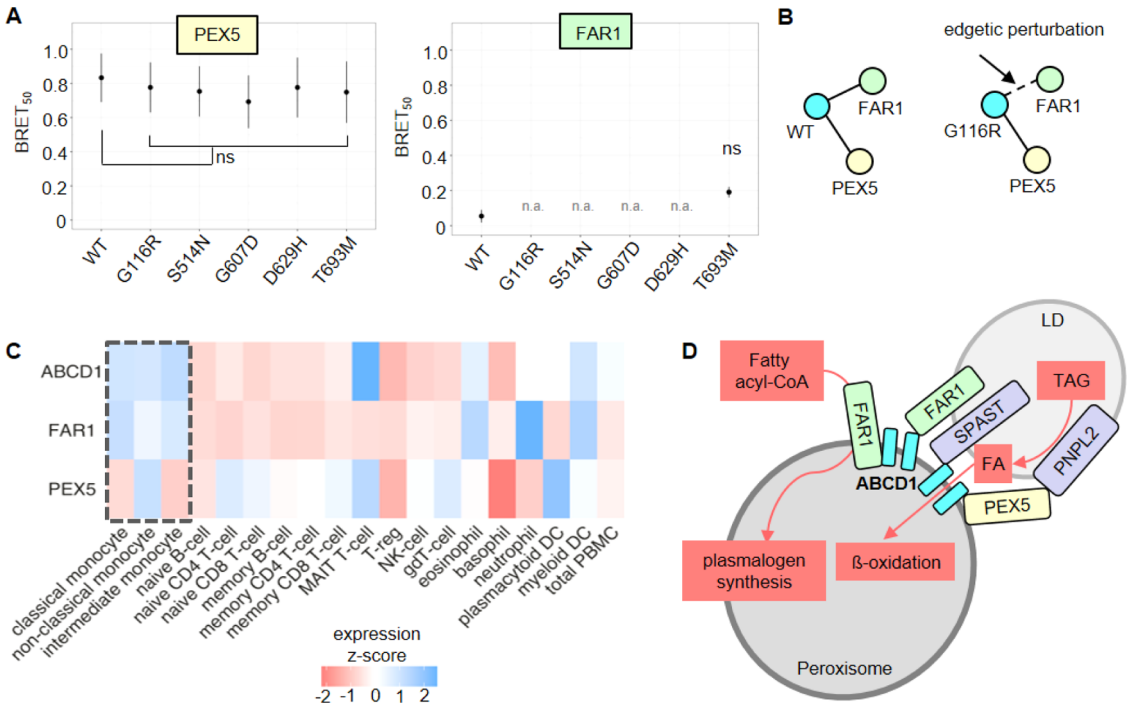


Figure 4. Functional aspects of ABCD1 protein interactions. (A) Variations of ABCD1 gene mutations on selected interactions of ABCD1. Wild-type ABCD1 (WT) was compared to ABCD1 constructs representing disease causing missense mutations (G116R, S514N, G607D, D629H, and T693M) and interactions with PEX5 and FAR1 were determined. BRET50 values as a measure of relative binding affinities (acceptor to donor ratio at a half-maximum BRET ratio) were calculated by nonlinear regression analysis; n.a., nonlinear regression analysis not applicable. Data are given as means ± SD of $n = 4$ independent experiments. The statistical significance of differences between individual BRET50 values was analyzed by ANOVA (ns, not significant with $P > 0.05$). (B) Edgetic perturbation of ABCD1 proteins with FAR1 interactions. Specific genetic variations in the ABCD1 gene (e.g., G116R) lead to loss of the interaction with FAR1. (C) Heatmap representation of expression z-scores for ABCD1, FAR1, and PEX5 in different blood cell types. Expression data were obtained from the human protein atlas. Note that in monocytes, ABCD1 and FAR1 are strongly coexpressed as indicated by positive z-scores for both genes. Additionally, PEX5 is also coexpressed with ABCD1 and FAR1 in nonclassical monocytes. (D) Hypothetical model how the identified interactions of FAR1 and PEX5 with ABCD1 might engage in peroxisome and lipid droplet (LD) metabolism. The known interaction of ABCD1 with SPAST tethers peroxisomes to LDs. FAR1 converts fatty acyl-CoA to fatty alcohol, which is then imported into peroxisomes. Fatty alcohol is needed for the peroxisomal plasmalogen synthesis. FAR1 is not only a peroxisomal tail-anchored membrane protein but also locates to the membrane of LD. ABCD1 binds FAR1 either in the peroxisomal membrane or in LD and hence may contribute to peroxisome-LD contact sites. PEX5 transiently binds PNPL2 and shuttles it to LD. PNPL2 is a lipase, which hydrolysis triglycerides (TAG) to fatty acids (FA) and diacylglycerol. Long-chain fatty acids then enter the peroxisomes via ABCD1 for subsequent β-oxidation. ABCD1 could serve as an anchor for PEX5. The interaction of ABCD1 with PEX5 may contribute to the interplay between peroxisomes and LDs.

protein−protein interactions of ABCD1 that were assigned to a confidence score of ≥2 (Supporting Information Table S10). Moreover, we provided evidence for nine not previously described interactions of ABCD1 with other peroxisomal proteins. We retested the 13 interactions of ABCD1 with peroxisomal proteins identified in the iBRET screen by means of bimolecular fluorescence complementation (BiFC) and a coaffinity purification (LUMIER) assay (Figure 3B). The BiFC assay resulted in positive classification for two of five known ABCD1 interactions with a confidence score of ≥2 (ABCD3, PEX19) and two of nine newly identified interactions (PEX5 and PXMP2). The LUMIER assay resulted in positive classification for two of five known ABCD1 interactions with a confidence score of ≥2 (ABCD2 and PEX19) and five of nine newly identified interactions (ALDH3A2, FAR1, PEX10, PIPOX, and PXMP2). The sequential application of both methods confirmed 69% of protein−protein interactions identified by iBRET, with an overlap of 15% between BiFC and Lumier. In summary, mapping ABCD1 against a whole-organelle peroxisomal library in a three-step iBRET screening approach followed by two-step sequential retesting yielded 13 binary protein−protein interactions with varying confidence scores (Figure 3B).

Next, we expanded the peroxisomal interaction network of ABCD1 in order to analyze ABCD1 function beyond its well-known role in lipid metabolism. For the 96 proteins of our peroxisomal library, we retrieved all known protein−protein interactions between them using a recently compiled interactome data set[64] and created an ABCD1-centered network (Figure 3C). In addition, we determined functional annotations of all ABCD1 interaction partners with respect to the gene ontology term ⟨biological process⟩[41] and to the peroxisomal database term ⟨function⟩.[61] Interaction partners of ABCD1 are involved in five distinct nonoverlapping processes (lipid metabolism, protein/amino acid metabolism, matrix protein import, peroxisome organization, and membrane assembly), where ABCD1 establishes seven interactions with proteins not associated with lipid metabolism. In addition to its well-documented role in fatty acid transport, ABCD1 also interacted with proteins that are involved in other pathways of peroxisomal lipid metabolism such as fatty acid and PUFA synthesis (ECI2, ALDH3A2) and the regulation of the Acyl-CoA/CoA ratio and plasmalogen synthesis (FAR1).

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

To further investigate a potential pathophysiological relevance of the newly identified protein−protein interactions, we analyzed the impact of variants in the *ABCD1* gene on ABCD1 interactions. If a variant leads to the total lack of the corresponding protein, all interactions of this protein will be erased (node removal). In the presence of residual proteins, a variant-induced modification or disruption of a single protein−protein interaction can occur (edgetic perturbation). We analyzed whether five pathogenic missense variants in the *ABCD1* gene known to be associated with residual protein stability[65] would induce edgetic perturbations. BRET saturation experiments were performed to investigate an influence of the *ABCD1* variants on the newly identified protein−protein interactions of ABCD1 with ALDH3A2, ECI2, FAR1, and PEX5. We specifically selected ALDH3A2, ECI2, FAR1, and PEX5, as we suspected a pathophysiological relevance of these interactions in X-ALD.[66−68] For ALDH3A2 and ECI2, the interaction with wild-type ABCD1 showed highly variable affinities, compatible with instability of the respective interaction. To our understanding, testing edgetic perturbations on instable protein−protein interactions is not feasible. Hence, subsequent investigations focused on the interaction of ABCD1 with FAR1 and PEX5. The latter were confirmed by BRET saturation experiments, both resulting in nonlinear regression fits with stable affinities. The interaction between ABCD1 and PEX5 was not affected by any of the variants, showing comparable affinities to the wild-type interaction (Figure 4A,B). Four pathogenic variants (G116R, S514N, G607D, and D629H) led to a loss of interaction with FAR1 (Figure 4A,B). No significant influence on the binding affinity of ABCD1−FAR1 was observed for T693M (Figure 4A).

To demonstrate the biological relevance of these interactions, we examined coexpression of ABCD1, FAR1, and PEX5 in blood cells. For this, we retrieved normalized transcript expression levels per gene in 18 different blood cell types from the human protein atlas database. Monocytes are one of the most affected cell types in cerebral ALD.[69] Interestingly, when comparing expression values across different blood cell types, we found coexpression for ABCD1 and FAR1 in all types of monocytes, indicated by positive *z*-scores (Figure 4C). For PEX5, a relevant coexpression with ABCD1 was observed only in nonclassical monocytes (Figure 4C).

DISCUSSION

Mapping protein−protein interaction networks with the aim to improve the understanding of cellular functions of disease genes and their involvement in pathological conditions requires complementary screening technologies with high accuracy and coverage. To address these needs, we established an automated high-throughput BRET-based assay. Performance and versatility of iBRET originated from (i) detection of interactions in living cells, (ii) testing of eight BRET-tag combinations, (iii) embedded bioinformatics algorithms, and (iv) automation of the workflow.

The application of the iBRET method in living mammalian cells aimed to achieve high sensitivity and specificity. iBRET minimizes false positives and false negatives resulting from exposition of target proteins to an artificial environment during cell lysis and sample preparation. As one consequence, the method is particularly well suited to determine interactions of membrane proteins. We did not reliably detect interactions based on receptor−ligand binding in this study. However, we implemented successful detection of extracellular receptor−ligand binding by omitting the washing step prior to the addition of the hRluc substrate.[70] Most methods applied for protein−protein interaction analyses make use of tags that are genetically fused to overexpressed target proteins. In BRET experiments, overexpression may increase the rate of false positive results by means of random collision (bystander BRET).[44,71] Conservative calibration of the threshold used for classification of positive interactions and the application of a gray area for undecided test results were the means we used to control the FPR in iBRET setups. Moreover, application of a live cell substrate (ViviRen)[42] allowed for exclusion of unspecific protein aggregation in nonvital cells.

The high sensitivity observed for iBRET was in part driven by the fact that iBRET used all possible orientations and combinations of tags. Testing a higher number of tag combinations is thought to promote the identification of a greater spectrum of protein−protein interactions.[56,72] In addition, we showed that the TPR can be improved by expanding the gray area and by increasing the number of BRET evaluation experiments. However, this scenario leads to an increase in the number of experiments per protein pair. Depending on the focus of a given screening approach, the number of tag combinations can be reduced to increase the throughput, however, at the cost of sensitivity. To reemphasize, attributed to the possible throughput, iBRET enables to perform additional experimental setups increasing the sensitivity of a protein−protein interaction screen.

We included a bioinformatics procedure for integrated control of assay quality, which accounts for variance observed in live cell assays and instrument-specific fluctuation of signal detection. In addition, the bioinformatics procedure evaluated test results based on parameters validated and trained on reference interactions. Raw data of experiments and controls are fed into an algorithm that decides on (a) repetition of experiments, (b) assignment of a protein pair to an evaluation experiment, or (c) final classification as interacting/noninteracting protein pair. Training and validation of parameters for the evaluation of assay quality and classification of test results require high-quality reference interaction sets.[12,58,59] Thorough analysis of control data sets may uncover potential bias in estimation of assay performance. We observed a significant increase in sensitivity for iBRET in a subset of manually curated high-confidence interactions (TPR, 90%; confidence score, 3) as compared to the full PRS (TPR, 58%; confidence score, 1.9) or a subset of low-confidence interactions (TPR, 30%; confidence score, 1). Low values for the TPR determined for low-confidence interactions may originate from false positives that are enriched in database entries with low confidence scores. On the other hand, manually curated high-confidence interactions may be enriched for easily detectable protein−protein interactions. Data sets of randomly selected noninteracting protein pairs (RRS) are well accepted for use as control cohorts;[53,56] however, this approach may lead to a bias in the estimation of prediction accuracy.[73]

Despite recent advances in screening technologies and the rapid growth of the human interactome, it has become clear that additional complementary assays are needed in order to generate high-confidence protein−protein interaction data sets.[11,19,53,74,75] With regard to its high sensitivity and specificity as well as efficient throughput, iBRET is distinctly eligible for replication studies of interactions found in high-throughput screens and provides the potential to screen

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

confined data sets such as signaling pathways or organelle interactomes *de novo*. Automated detection of protein−protein interactions in living cells sets the basis for future demands such as mapping of dynamic networks[76] or tissue-specific networks.[77] When considering protein−protein interactions as potential pharmaceutical targets,[78,79] application of iBRET in compound library screens may constitute a significant added value to detect molecules that stabilize or disrupt protein complexes.[80]

Applicability of iBRET in protein−protein interaction assays was further demonstrated in a whole-organelle peroxisomal screen. The peroxisomal membrane transporter ABCD1, which is deficient in X-ALD, was screened against a library of 96 proteins with additional 8 isoforms. Confidence scores were assigned to nine novel ABCD1 interactions that were identified by iBRET with sequential analysis by BiFC and LUMIER. Based on known interactions of ABCD1 with peroxisomal proteins and novel interactions identified in this study, we applied a network-based approach to explore the molecular complexity of X-ALD.

The ABCD1-centered peroxisomal network showed that all newly identified interaction partners of ABCD1 are assigned to functional processes that have not been linked to ABCD1 function before. Of these, several hypotheses can be derived. (i) The interactions with the peroxisomal matrix protein import machinery components PEX10 and PEX13 and with the pore-forming membrane protein PXMP2 may constitute an association of ABCD1 function with processes of peroxisomal organization. (ii) PIPOX is involved in peroxisomal lysine metabolism, which was demonstrated to play a functional role in osmotic stress resistance,[81] and oxidative stress is suggested as a pathogenic mechanism in X-ALD. (iii) Our data point to a broader function of ABCD1 in pathways of the fatty acid metabolism. ALDH3A2 is involved in the last step of the conversion of ω-hydroxy-VLCFA into the VLC-dicarboxylic acids.[82] Even though there are clear indications that VLC-dicarboxylic acids are most preferentially transported by the half-transporter PMP70 into the peroxisome,[83] the interaction of ABCD1 with ALDH3A2 may suggest a role of ABCD1 in the VLC-dicarboxylic metabolism. (iv) FAR1 is a tail-anchored peroxisomal membrane protein that supplies fatty alcohols required for plasmalogen synthesis. The interaction between ABCD1 and FAR1 could point to a function of ABCD1 in plasmalogen synthesis. When compared to healthy controls, Herzog *et al.* showed reduced plasmalogen levels in ABCD1-deficient skin fibroblasts.[84] The authors speculate that accumulated VCLFA might interfere with plasmalogen synthesis. In light of existing knowledge and data of our study, we hypothesize that the interaction of ABCD1 with FAR1 affects the enzyme function of FAR1 and, as a consequence, intracellular plasmalogen synthesis. Further studies are required to dissect the molecular consequences of this physical protein interaction. This may be of special interest considering that pathogenic variants in ABCD1 disturb the interactions with FAR1, which in turn could lead to reduced plasmalogen levels. Further evidence for the biological importance of this interaction is given by the fact that both proteins are specifically expressed in monocytes, which are known to be important players in X-ALD disease pathology (Figure 4D).[69] (v) The interactions of ABCD1 with FAR1 and PEX5 allow us to speculate about the role of ABCD1 in the interplay of lipid droplets and peroxisomes. Peroxisomes and lipid droplets share contact sites to facilitate fatty acid metabolism, established by the interaction of ABCD1 with Spastin.[37] A knockout model of ABCD1 in *Caenorhabditis elegans* showed an increased size of intracellular lipid droplets.[85] Recently, a lipid metabolism-dependent localization of FAR1 to lipid droplets has been suggested.[86] PEX5 is a peroxisomal protein that targets peroxisomal matrix enzymes into peroxisomes. A recent investigation revealed that under fasting conditions, PEX5 is able to shuttle the lipid droplet-associated lipase PNPL2 to the surface of lipid droplets.[87] PNPL2 hydrolyses triglycerides to fatty acids and diacylglycerols and provides long-chain fatty acids, which can enter the peroxisomes by ABCD1 for subsequent ß-oxidation. Considering these observations, the interactions of ABCD1 with PEX5 and FAR1 may contribute to the tight interplay between peroxisomes and lipid droplets, for example, ABCD1 could serve as a linker for PEX5 (Figure 4D). However, in contrast to the ABCD1−FAR1 interaction, we did not observe any direct alterations of the ABCD1−PEX5 interaction by selected disease causing ABCD1 pathogenic missense variants.

In conclusion, robust protein−protein interaction data in combination with network analyses serve as powerful tools to further our understanding of the function and dysfunction of proteins encoded by disease genes. iBRET provides high sensitivity and efficient throughput for the study of protein−protein interactions and respective mutation-induced perturbations in live-cell screening approaches. The identification of novel interaction partners for ABCD1 expanded the functional context of X-ALD and may in future provide targets for alternative strategies for the treatment of patients suffering from this disease.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.1c00330.

Strategy and automation of a BRET-based method for the detection of protein interactions in living cells, iBRET process flow, BRET evaluation experiments, generation of the reference data set, subcellular distribution of candidate proteins, supervised approach; comparison of the TPR and FPR with respect to different thresholds for positive prediction, BRET evaluation experiments for ABCD1, coding sequence of BRET tags, iBRET throughput, cell lines tested for high-throughput transfection, PRS, RRS, cDNA library, BRET raw data for PRS and RRS, classes of protein interactions of the PRS, library of genes encoding peroxisomal proteins, protein−protein interactions of ABCD1 with peroxisomal proteins, and iBRET software with algorithms for training and classification (softwarepackage/.zip) (PDF)

BRET evaluation experiments (PDF)

iBRET software (ZIP)

Raw data of BRET experiments for the positive reference set (XLS)

Raw data of BRET experiments for the random reference set (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Amelie S. Lotz-Havla − *Dr. von Hauner Children's Hospital, Ludwig-Maximilians-Universität München, 80337 Munich,*

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

Germany; orcid.org/0000-0002-6058-3645; Phone: +49 89 4400 52811; Email: amelie.lotz@med.uni-muenchen.de

Søren W. Gersting − *University Children's Research, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany*; Phone: +49 40 7410 55330; Email: gersting@uke.de

Authors

Mathias Woidy − *University Children's Research, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany*

Philipp Guder − *University Children's Research, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany*

Caroline C. Friedel − *Institute of Informatics, Ludwig-Maximilians-Universität München, 80538 Munich, Germany*

Julian M. Klingbeil − *Dr. von Hauner Children's Hospital, Ludwig-Maximilians-Universität München, 80337 Munich, Germany*

Ana-Maria Bulau − *Dr. von Hauner Children's Hospital, Ludwig-Maximilians-Universität München, 80337 Munich, Germany*

Anja Schultze − *Dr. von Hauner Children's Hospital, Ludwig-Maximilians-Universität München, 80337 Munich, Germany*

Ilona Dahmen − *Dr. von Hauner Children's Hospital, Ludwig-Maximilians-Universität München, 80337 Munich, Germany*

Heidi Noll-Puchta − *Dr. von Hauner Children's Hospital, Ludwig-Maximilians-Universität München, 80337 Munich, Germany*

Stephan Kemp − *Department of Clinical Chemistry, Laboratory Genetic Metabolic Diseases, Amsterdam UMC, Amsterdam Neuroscience, Amsterdam Gastroenterology & Metabolism, University of Amsterdam, 1105 WX Amsterdam, The Netherlands*

Ralf Erdmann − *Systems Biochemistry, Medical Faculty, Ruhr-University Bochum, 44801 Bochum, Germany*

Ralf Zimmer − *Institute of Informatics, Ludwig-Maximilians-Universität München, 80538 Munich, Germany*

Ania C. Muntau − *University Children's Hospital, University Medical Center Hamburg Eppendorf, 20246 Hamburg, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.1c00330

Author Contributions

A.S.L.-H., M.W., and P.G. contributed equally to this work. S.W.G., R.Z., and A.C.M. designed the research. S.W.G., A.S.L.-H., and A.C.M. designed the iBRET platform. S.W.G. and M.W. set up the automated process. A.S.L.-H., M.W., J.M.K., P.G., A.-M.B., A.S., H.N.-P., and I.D. performed protein−protein interaction analyses. C.C.F., M.W., and R.Z. performed bioinformatics analyses. A.S.L.-H., M.W., A.C.M., and S.W.G. wrote the manuscript. A.S.L.-H., M.W., P.G., C.C.F., S.K., S.W.G., R.Z., R.E., and A.C.M. contributed to the final version of the manuscript.

Funding

This work was supported by the Bavarian Genome Research Network (BayGene) and an LMUexcellent grant 42595-6 to A.C.M., by a Prinz Lennart von Hohenzollern Stiftung grant to A.C.M and S.W.G. as well as a Fritz Thyssen Stiftung scholarship to A.S.L.-H.

Notes

The authors declare no competing financial interest.
Code availability: provided in Supporting Information Material S1.

■ ACKNOWLEDGMENTS

We wish to thank Horst Nöcker for contribution to the setup of the final automated process and Wolfgang Schliebs for valuable scientific discussion of reference data sets. We wish to acknowledge Ulrich Schatz for construction of BRET vectors. We wish to thank Anna Heckel-Pompey for technical assistance.

■ REFERENCES

(1) Ideker, T.; Sharan, R. Protein networks in disease. *Genome Res.* **2008**, *18*, 644−652.
(2) Vidal, M.; Cusick, M. E.; Barabási, A.-L. Interactome networks and human disease. *Cell* **2011**, *144*, 986−998.
(3) Novarino, G.; Fenstermaker, A. G.; Zaki, M. S.; Hofree, M.; Silhavy, J. L.; Heiberg, A. D.; Abdellateef, M.; Rosti, B.; Scott, E.; Mansour, L.; Masri, A.; Kayserili, H.; Al-Aama, J. Y.; Abdel-Salam, G. M. H.; Karminejad, A.; Kara, M.; Kara, B.; Bozorgmehri, B.; Ben-Omran, T.; Mojahedi, F.; Mahmoud, I. G. E. D.; Bouslam, N.; Bouhouche, A.; Benomar, A.; Hanein, S.; Raymond, L.; Forlani, S.; Mascaro, M.; Selim, L.; Shehata, N.; Al-Allawi, N.; Bindu, P. S.; Azam, M.; Gunel, M.; Caglayan, A.; Bilguvar, K.; Tolun, A.; Issa, M. Y.; Schroth, J.; Spencer, E. G.; Rosti, R. O.; Akizu, N.; Vaux, K. K.; Johansen, A.; Koh, A. A.; Megahed, H.; Durr, A.; Brice, A.; Stevanin, G.; Gabriel, S. B.; Ideker, T.; Gleeson, J. G. Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. *Science* **2014**, *343*, 506−511.
(4) Caldera, M.; Buphamalai, P.; Müller, F.; Menche, J. Interactome-based approaches to human disease. *Curr. Opin. Syst. Biol.* **2017**, *3*, 88−94.
(5) Miryala, S. K.; Anbarasu, A.; Ramaiah, S. Discerning molecular interactions: A comprehensive review on biomolecular interaction databases and network analysis tools. *Gene* **2018**, *642*, 84−94.
(6) Safari-Alighiarloo, N.; Taghizadeh, M.; Rezaei-Tavirani, M.; Goliaei, B.; Peyvandi, A. A. Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterol. Hepatol.* **2014**, *7*, 17−31.
(7) Sahni, H.; Ross, S.; Barbarulo, A.; Solanki, A.; Lau, C.-I.; Furmanski, A.; Saldaña, J. I.; Ono, M.; Hubank, M.; Barenco, M.; Crompton, T. A genome wide transcriptional model of the complex response to pre-TCR signalling during thymocyte differentiation. *Oncotarget* **2015**, *6*, 28646−28660.
(8) Venkatesan, K.; Rual, J.-F.; Vazquez, A.; Stelzl, U.; Lemmens, I.; Hirozane-Kishikawa, T.; Hao, T.; Zenkner, M.; Xin, X.; Goh, K.-I.; Yildirim, M. A.; Simonis, N.; Heinzmann, K.; Gebreab, F.; Sahalie, J. M.; Cevik, S.; Simon, C.; de Smet, A.-S.; Dann, E.; Smolyar, A.; Vinayagam, A.; Yu, H.; Szeto, D.; Borick, H.; Dricot, A.; Klitgord, N.; Murray, R. R.; Lin, C.; Lalowski, M.; Timm, J.; Rau, K.; Boone, C.; Braun, P.; Cusick, M. E.; Roth, F. P.; Hill, D. E.; Tavernier, J.; Wanker, E. E.; Barabási, A.-L.; Vidal, M. An empirical framework for binary interactome mapping. *Nat. Methods* **2009**, *6*, 83−90.
(9) Stumpf, M. P. H.; Thorne, T.; de Silva, E.; Stewart, R.; An, H. J.; Lappe, M.; Wiuf, C. Estimating the size of the human interactome. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6959−6964.
(10) Luck, K.; Sheynkman, G. M.; Zhang, I.; Vidal, M. Proteome-Scale Human Interactomics. *Trends Biochem. Sci.* **2017**, *42*, 342−354.
(11) Luck, K.; Kim, D.-K.; Lambourne, L.; Spirohn, K.; Begg, B. E.; Bian, W.; Brignall, R.; Cafarelli, T.; Campos-Laborie, F. J.; Charloteaux, B.; Choi, D.; Coté, A. G.; Daley, M.; Deimling, S.; Desbuleux, A.; Dricot, A.; Gebbia, M.; Hardy, M. F.; Kishore, N.; Knapp, J. J.; Kovács, I. A.; Lemmens, I.; Mee, M. W.; Mellor, J. C.; Pollis, C.; Pons, C.; Richardson, A. D.; Schlabach, S.; Teeking, B.;

Yadav, A.; Babor, M.; Balcha, D.; Basha, O.; Bowman-Colin, C.; Chin, S.-F.; Choi, S. G.; Colabella, C.; Coppin, G.; D'Amata, C.; De Ridder, D.; De Rouck, S.; Duran-Frigola, M.; Ennajdaoui, H.; Goebels, F.; Goehring, L.; Gopal, A.; Haddad, G.; Hatchi, E.; Helmy, M.; Jacob, Y.; Kassa, Y.; Landini, S.; Li, R.; van Lieshout, N.; MacWilliams, A.; Markey, D.; Paulson, J. N.; Rangarajan, S.; Rasla, J.; Rayhan, A.; Rolland, T.; San-Miguel, A.; Shen, Y.; Sheykhkarimli, D.; Sheynkman, G. M.; Simonovsky, E.; Taşan, M.; Tejeda, A.; Tropepe, V.; Twizere, J.-C.; Wang, Y.; Weatheritt, R. J.; Weile, J.; Xia, Y.; Yang, X.; Yeger-Lotem, E.; Zhong, Q.; Aloy, P.; Bader, G. D.; De Las Rivas, J.; Gaudet, S.; Hao, T.; Rak, J.; Tavernier, J.; Hill, D. E.; Vidal, M.; Roth, F. P.; Calderwood, M. A. A reference map of the human binary protein interactome. *Nature* **2020**, *580*, 402−408.

(12) von Mering, C.; Krause, R.; Snel, B.; Cornell, M.; Oliver, S. G.; Fields, S.; Bork, P. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **2002**, *417*, 399−403.

(13) Levy, E. D.; Landry, C. R.; Michnick, S. W. How perfect can protein interactomes be? *Sci. Signal.* **2009**, *2*, 11.

(14) Parrish, J. R.; Gulyas, K. D.; Finley, R. L., Jr. Yeast two-hybrid contributions to interactome mapping. *Curr. Opin. Biotechnol.* **2006**, *17*, 387−393.

(15) Yu, H.; Braun, P.; Yildirim, M. A.; Lemmens, I.; Venkatesan, K.; Sahalie, J.; Hirozane-Kishikawa, T.; Gebreab, F.; Li, N.; Simonis, N.; Hao, T.; Rual, J.-F.; Dricot, A.; Vazquez, A.; Murray, R. R.; Simon, C.; Tardivo, L.; Tam, S.; Svrzikapa, N.; Fan, C.; de Smet, A.-S.; Motyl, A.; Hudson, M. E.; Park, J.; Xin, X.; Cusick, M. E.; Moore, T.; Boone, C.; Snyder, M.; Roth, F. P.; Barabasi, A.-L.; Tavernier, J.; Hill, D. E.; Vidal, M. High-quality binary protein interaction map of the yeast interactome network. *Science* **2008**, *322*, 104−110.

(16) Schwartz, A. S.; Yu, J.; Gardenour, K. R.; Finley, R. L., Jr.; Ideker, T. Cost-effective strategies for completing the interactome. *Nat. Methods* **2009**, *6*, 55−61.

(17) Caufield, J. H.; Sakhawalkar, N.; Uetz, P. A comparison and optimization of yeast two-hybrid systems. *Methods* **2012**, *58*, 317−324.

(18) Rajagopala, S. V.; Sikorski, P.; Kumar, A.; Mosca, R.; Vlasblom, J.; Arnold, R.; Franca-Koh, J.; Pakala, S. B.; Phanse, S.; Ceol, A.; Häuser, R.; Siszler, G.; Wuchty, S.; Emili, A.; Babu, M.; Aloy, P.; Pieper, R.; Uetz, P. The binary protein-protein interaction landscape of Escherichia coli. *Nat. Biotechnol.* **2014**, *32*, 285−290.

(19) Choi, S. G.; Olivet, J.; Cassonnet, P.; Vidalain, P.-O.; Luck, K.; Lambourne, L.; Spirohn, K.; Lemmens, I.; Dos Santos, M.; Demeret, C.; Jones, L.; Rangarajan, S.; Bian, W.; Coutant, E. P.; Janin, Y. L.; van der Werf, S.; Trepte, P.; Wanker, E. E.; De Las Rivas, J.; Tavernier, J.; Twizere, J.-C.; Hao, T.; Hill, D. E.; Vidal, M.; Calderwood, M. A.; Jacob, Y. Maximizing binary interactome mapping with a minimal number of assays. *Nat. Commun.* **2019**, *10*, 3907.

(20) Xu, Y.; Piston, D. W.; Johnson, C. H. A bioluminescence resonance energy transfer (BRET) system: application to interacting circadian clock proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 151−156.

(21) Bacart, J.; Corbel, C.; Jockers, R.; Bach, S.; Couturier, C. The BRET technology and its application to screening assays. *Biotechnol. J.* **2008**, *3*, 311−324.

(22) Mo, X.-L.; Fu, H. BRET: NanoLuc-Based Bioluminescence Resonance Energy Transfer Platform to Monitor Protein-Protein Interactions in Live Cells. *Methods Mol. Biol.* **2016**, *1439*, 263−271.

(23) Mo, X.-L.; Luo, Y.; Ivanov, A. A.; Su, R.; Havel, J. J.; Li, Z.; Khuri, F. R.; Du, Y.; Fu, H. Enabling systematic interrogation of protein-protein interactions in live cells with a versatile ultra-high-throughput biosensor platform. *J. Mol. Cell Biol.* **2016**, *8*, 271−281.

(24) Hamdan, F. F.; Audet, M.; Garneau, P.; Pelletier, J.; Bouvier, M. High-throughput screening of G protein-coupled receptor antagonists using a bioluminescence resonance energy transfer 1-based beta-arrestin2 recruitment assay. *J. Biomol. Screen* **2005**, *10*, 463−475.

(25) Vrecl, M.; Jorgensen, R.; Pogačnik, A.; Heding, A. Development of a BRET2 screening assay using beta-arrestin 2 mutants. *J. Biomol. Screen* **2004**, *9*, 322−333.

(26) Couturier, C.; Deprez, B. Setting Up a Bioluminescence Resonance Energy Transfer High throughput Screening Assay to Search for Protein/Protein Interaction Inhibitors in Mammalian Cells. *Front. Endocrinol.* **2012**, *3*, 100.

(27) Robinson, K. H.; Yang, J. R.; Zhang, J. FRET and BRET-based biosensors in live cell compound screens. *Methods Mol. Biol.* **2014**, *1071*, 217−225.

(28) Ong, L. L.; Vasta, J. D.; Monereau, L.; Locke, G.; Ribeiro, H.; Pattoli, M. A.; Skala, S.; Burke, J. R.; Watterson, S. H.; Tino, J. A.; Meisenheimer, P. L.; Arey, B.; Lippy, J.; Zhang, L.; Robers, M. B.; Tebben, A.; Chaudhry, C. A High-Throughput BRET Cellular Target Engagement Assay Links Biochemical to Cellular Activity for Bruton's Tyrosine Kinase. *SLAS Discovery* **2020**, *25*, 176−185.

(29) Jin, H. Y.; Tudor, Y.; Choi, K.; Shao, Z.; Sparling, B. A.; McGivern, J. G.; Symons, A. High-Throughput Implementation of the NanoBRET Target Engagement Intracellular Kinase Assay to Reveal Differential Compound Engagement by SIK2/3 Isoforms. *SLAS Discovery* **2020**, *25*, 215−222.

(30) Baker, A.; Carrier, D. J.; Schaedler, T.; Waterham, H. R.; van Roermund, C. W.; Theodoulou, F. L. Peroxisomal ABC transporters: functions and mechanism. *Biochem. Soc. Trans.* **2015**, *43*, 959−965.

(31) Mosser, J.; Douar, A.-M.; Sarde, C.-O.; Kioschis, P.; Feil, R.; Moser, H.; Poustka, A.-M.; Mandel, J.-L.; Aubourg, P. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. *Nature* **1993**, *361*, 726−730.

(32) Engelen, M.; Kemp, S.; de Visser, M.; van Geel, B. M.; Wanders, R. J.; Aubourg, P.; Poll-The, B. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. *Orphanet J. Rare Dis.* **2012**, *7*, 51.

(33) Kemp, S.; Berger, J.; Aubourg, P. X-linked adrenoleukodystrophy: clinical, metabolic, genetic and pathophysiological aspects. *Biochim. Biophys. Acta* **2012**, *1822*, 1465−1474.

(34) Schlüter, A.; Sandoval, J.; Fourcade, S.; Díaz-Lagares, A.; Ruiz, M.; Casaccia, P.; Esteller, M.; Pujol, A. Epigenomic signature of adrenoleukodystrophy predicts compromised oligodendrocyte differentiation. *Brain Pathol.* **2018**, *28*, 902−919.

(35) Gärtner, J.; Braun, A.; Holzinger, A.; Roerig, P.; Lenard, H. G.; Roscher, A. A. Clinical and genetic aspects of X-linked adrenoleukodystrophy. *Neuropediatrics* **1998**, *29*, 3−13.

(36) Hillebrand, M.; Gersting, S. W.; Lotz-Havla, A. S.; Schäfer, A.; Rosewich, H.; Valerius, O.; Muntau, A. C.; Gärtner, J. Identification of a new fatty acid synthesis-transport machinery at the peroxisomal membrane. *J. Biol. Chem.* **2012**, *287*, 210−221.

(37) Chang, C.-L.; Weigel, A. V.; Ioannou, M. S.; Pasolli, H. A.; Xu, C. S.; Peale, D. R.; Shtengel, G.; Freeman, M.; Hess, H. F.; Blackstone, C.; Lippincott-Schwartz, J. Spastin tethers lipid droplets to peroxisomes and directs fatty acid trafficking through ESCRT-III. *J. Cell Biol.* **2019**, *218*, 2583−2599.

(38) Cohen, S.; Valm, A. M.; Lippincott-Schwartz, J. Interacting organelles. *Curr. Opin. Cell Biol.* **2018**, *53*, 84−91.

(39) Nagai, T.; Ibata, K.; Park, E. S.; Kubota, M.; Mikoshiba, K.; Miyawaki, A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **2002**, *20*, 87−90.

(40) Nyfeler, B.; Michnick, S. W.; Hauri, H. Capturing protein interactions in the secretory pathway of living cells. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (18), 6350−6355.

(41) Ashburner, M.; Ball, C. A.; Blake, J. A.; Botstein, D.; Butler, H.; Cherry, J. M.; Davis, A. P.; Dolinski, K.; Dwight, S. S.; Eppig, J. T.; Harris, M. A.; Hill, D. P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J. C.; Richardson, J. E.; Ringwald, M.; Rubin, G. M.; Sherlock, G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25−29.

(42) Gersting, S. W.; Lotz-Havla, A. S.; Muntau, A. C. Bioluminescence resonance energy transfer: an emerging tool for the detection of protein-protein interaction in living cells. *Methods Mol. Biol.* **2012**, *815*, 253−263.

(43) Mercier, J.-F.; Salahpour, A.; Angers, S.; Breit, A.; Bouvier, M. Quantitative assessment of beta 1- and beta 2-adrenergic receptor

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

homo- and heterodimerization by bioluminescence resonance energy transfer. *J. Biol. Chem.* **2002**, *277*, 44925−44931.
(44) Hamdan, F. F.; Percherancier, Y.; Breton, B.; Bouvier, M. Monitoring protein-protein interactions in living cells by bioluminescence resonance energy transfer (BRET). *Curr. Protoc. Neurosci.* **2006**, *34*, 5−23.
(45) Oner, S. S.; Maher, E. M.; Breton, B.; Bouvier, M.; Blumer, J. B. Receptor-regulated interaction of activator of G-protein signaling-4 and Galphai. *J. Biol. Chem.* **2010**, *285*, 20588−20594.
(46) Kemp, S.; Theodoulou, F. L.; Wanders, R. J. Mammalian peroxisomal ABC transporters: from endogenous substrates to pathology and clinical significance. *Br. J. Pharmacol.* **2011**, *164*, 1753−1766.
(47) Nakagawa, C.; Inahata, K.; Nishimura, S.; Sugimoto, K. Improvement of a Venus-based bimolecular fluorescence complementation assay to visualize bFos-bJun interaction in living cells. *Biosci. Biotechnol. Biochem.* **2011**, *75*, 1399−1401.
(48) Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, A.; Kampf, C.; Sjöstedt, E.; Asplund, A.; Olsson, I.; Edlund, K.; Lundberg, E.; Navani, S.; Szigyarto, C. A.-K.; Odeberg, J.; Djureinovic, D.; Takanen, J. O.; Hober, S.; Alm, T.; Edqvist, P.-H.; Berling, H.; Tegel, H.; Mulder, J.; Rockberg, J.; Nilsson, P.; Schwenk, J. M.; Hamsten, M.; von Feilitzen, K.; Forsberg, M.; Persson, L.; Johansson, F.; Zwahlen, M.; von Heijne, G.; Nielsen, J.; Pontén, F. Proteomics. Tissue-based map of the human proteome. *Science* **2015**, *347*, 1260419.
(49) Pfleger, K. D. G.; Eidne, K. A. Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). *Nat. Methods* **2006**, *3*, 165−174.
(50) Villemure, J. F.; Adam, L.; Bevan, N. J.; Gearing, K.; Chénier, S.; Bouvier, M. Subcellular distribution of GABA(B) receptor homo- and hetero-dimers. *Biochem. J.* **2005**, *388*, 47−55.
(51) Coulon, V.; Audet, M.; Homburger, V.; Bockaert, J.; Fagni, L.; Bouvier, M.; Perroy, J. Subcellular imaging of dynamic protein interactions by bioluminescence resonance energy transfer. *Biophys. J.* **2008**, *94*, 1001−1009.
(52) James, J. R.; Oliveira, M. I.; Carmo, A. M.; Iaboni, A.; Davis, S. J. A rigorous experimental framework for detecting protein oligomerization using bioluminescence resonance energy transfer. *Nat. Methods* **2006**, *3*, 1001−1006.
(53) Cusick, M. E.; Yu, H.; Smolyar, A.; Venkatesan, K.; Carvunis, A.-R.; Simonis, N.; Rual, J.-F.; Borick, H.; Braun, P.; Dreze, M.; Vandenhaute, J.; Galli, M.; Yazaki, J.; Hill, D. E.; Ecker, J. R.; Roth, F. P.; Vidal, M. Literature-curated protein interaction datasets. *Nat. Methods* **2009**, *6*, 39−46.
(54) Srivastava, A.; Mazzocco, G.; Kel, A.; Wyrwicz, L. S.; Plewczynski, D. Detecting reliable non interacting proteins (NIPs) significantly enhancing the computational prediction of protein-protein interactions using machine learning methods. *Mol. Biosyst.* **2016**, *12*, 778−785.
(55) Park, Y.; Marcotte, E. M. Revisiting the negative example sampling problem for predicting protein-protein interactions. *Bioinformatics* **2011**, *27*, 3024−3028.
(56) Braun, P.; Tasan, M.; Dreze, M.; Barrios-Rodiles, M.; Lemmens, I.; Yu, H.; Sahalie, J. M.; Murray, R. R.; Roncari, L.; de Smet, A.-S.; Venkatesan, K.; Rual, J.-F.; Vandenhaute, J.; Cusick, M. E.; Pawson, T.; Hill, D. E.; Tavernier, J.; Wrana, J. L.; Roth, F. P.; Vidal, M. An experimentally derived confidence score for binary protein-protein interactions. *Nat. Methods* **2009**, *6*, 91−97.
(57) Rual, J.-F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Dricot, A.; Li, N.; Berriz, G. F.; Gibbons, F. D.; Dreze, M.; Ayivi-Guedehoussou, N.; Klitgord, N.; Simon, C.; Boxem, M.; Milstein, S.; Rosenberg, J.; Goldberg, D. S.; Zhang, L. V.; Wong, S. L.; Franklin, G.; Li, S.; Albala, J. S.; Lim, J.; Fraughton, C.; Llamosas, E.; Cevik, S.; Bex, C.; Lamesch, P.; Sikorski, R. S.; Vandenhaute, J.; Zoghbi, H. Y.; Smolyar, A.; Bosak, S.; Sequerra, R.; Doucette-Stamm, L.; Cusick, M. E.; Hill, D. E.; Roth, F. P.; Vidal, M. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **2005**, *437*, 1173−1178.
(58) Stelzl, U.; Worm, U.; Lalowski, M.; Haenig, C.; Brembeck, F. H.; Goehler, H.; Stroedicke, M.; Zenkner, M.; Schoenherr, A.; Koeppen, S.; Timm, J.; Mintzlaff, S.; Abraham, C.; Bock, N.; Kietzmann, S.; Goedde, A.; Toksöz, E.; Droege, A.; Krobitsch, S.; Korn, B.; Birchmeier, W.; Lehrach, H.; Wanker, E. E. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **2005**, *122*, 957−968.
(59) Lemmens, I.; Lievens, S.; Tavernier, J. Strategies towards high-quality binary protein interactome maps. *J. Proteomics* **2010**, *73*, 1415−1420.
(60) El Khamlichi, C.; Reverchon-Assadi, F.; Hervouet-Coste, N.; Blot, L.; Reiter, E.; Morisset-Lopez, S. Bioluminescence Resonance Energy Transfer as a Method to Study Protein-Protein Interactions: Application to G Protein Coupled Receptor Biology. *Molecules* **2019**, *24*, 537.
(61) Schlüter, A.; Real-Chicharro, A.; Gabaldón, T.; Sánchez-Jiménez, F.; Pujol, A. PeroxisomeDB 2.0: an integrative view of the global peroxisomal metabolome. *Nucleic Acids Res.* **2010**, *38*, D800.
(62) Yifrach, E.; Fischer, S.; Oeljeklaus, S.; Schuldiner, M.; Zalckvar, E.; Warscheid, B. Defining the Mammalian Peroxisomal Proteome. *Subcell. Biochem.* **2018**, *89*, 47−66.
(63) Osumi, T.; Imamura, A.; Tsukamoto, T.; Fujiwara, C.; Hashiguchi, N.; Shimozawa, N.; Suzuki, Y.; Kondo, N. Temperature sensitivity in peroxisome assembly processes characterizes milder forms of peroxisome biogenesis disorders. *Cell Biochem. Biophys.* **2000**, *32*, 165−170.
(64) Cheng, F.; Desai, R. J.; Handy, D. E.; Wang, R.; Schneeweiss, S.; Barabási, A.-L.; Loscalzo, J. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **2018**, *9*, 2691.
(65) Kemp, S.; Pujol, A.; Waterham, H. R.; van Geel, B. r. M.; Boehm, C. D.; Raymond, G. V.; Cutting, G. R.; Wanders, R. J. A.; Moser, H. W. ABCD1 mutations and the X-linked adrenoleukodystrophy mutation database: role in diagnosis and clinical correlations. *Hum. Mutat.* **2001**, *18*, 499−515.
(66) Ruiz, M.; Jové, M.; Schlüter, A.; Casasnovas, C.; Villarroya, F.; Guilera, C.; Ortega, F. J.; Naudí, A.; Pamplona, R.; Gimeno, R.; Fourcade, S.; Portero-Otín, M.; Pujol, A. Altered glycolipid and glycerophospholipid signaling drive inflammatory cascades in adrenomyeloneuropathy. *Hum. Mol. Genet.* **2015**, *24*, 6861.
(67) Schlüter, A.; Espinosa, L.; Fourcade, S.; Galino, J.; López, E.; Ilieva, E.; Morató, L.; Asheuer, M.; Cook, T.; McLaren, A.; Reid, J.; Kelly, F.; Bates, S.; Aubourg, P.; Galea, E.; Pujol, A. Functional genomic analysis unravels a metabolic-inflammatory interplay in adrenoleukodystrophy. *Hum. Mol. Genet.* **2012**, *21*, 1062−1077.
(68) Bottelbergs, A.; Verheijden, S.; Van Veldhoven, P. P.; Just, W.; Devos, R.; Baes, M. Peroxisome deficiency but not the defect in ether lipid synthesis causes activation of the innate immune system and axonal loss in the central nervous system. *J. Neuroinflammation* **2012**, *9*, 61.
(69) Weinhofer, I.; Zierfuss, B.; Hametner, S.; Wagner, M.; Popitsch, N.; Machacek, C.; Bartolini, B.; Zlabinger, G.; Ohradanova-Repic, A.; Stockinger, H.; Köhler, W.; Höftberger, R.; Regelsberger, G.; Forss-Petter, S.; Lassmann, H.; Berger, J. Impaired plasticity of macrophages in X-linked adrenoleukodystrophy. *Brain* **2018**, *141*, 2329−2342.
(70) Nold-Petry, C. A.; Lo, C. Y.; Rudloff, I.; Elgass, K. D.; Li, S.; Gantier, M. P.; Lotz-Havla, A. S.; Gersting, S. W.; Cho, S. X.; Lao, J. C.; Ellisdon, A. M.; Rotter, B.; Azam, T.; Mangan, N. E.; Rossello, F. J.; Whisstock, J. C.; Bufler, P.; Garlanda, C.; Mantovani, A.; Dinarello, C. A.; Nold, M. F. IL-37 requires the receptors IL-18Ralpha and IL-1R8 (SIGIRR) to carry out its multifaceted anti-inflammatory program upon innate signal transduction. *Nat. Immunol.* **2015**, *16*, 354−365.
(71) Sampaio, N. G.; Kocan, M.; Schofield, L.; Pfleger, K. D. G.; Eriksson, E. M. Investigation of interactions between TLR2, MyD88 and TIRAP by bioluminescence resonance energy transfer is hampered by artefacts of protein overexpression. *PLoS One* **2018**, *13*, No. e0202408.

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

(72) Stellberger, T.; Häuser, R.; Baiker, A.; Pothineni, V. R.; Haas, J.; Uetz, P. Improving the yeast two-hybrid system with permutated fusions proteins: the Varicella Zoster Virus interactome. *Proteome Sci.* **2010**, *8*, 8.
(73) Trabuco, L. G.; Betts, M. J.; Russell, R. B. Negative protein-protein interaction datasets derived from large-scale two-hybrid experiments. *Methods* **2012**, *58*, 343−348.
(74) Rolland, T.; Taşan, M.; Charloteaux, B.; Pevzner, S. J.; Zhong, Q.; Sahni, N.; Yi, S.; Lemmens, I.; Fontanillo, C.; Mosca, R.; Kamburov, A.; Ghiassian, S. D.; Yang, X.; Ghamsari, L.; Balcha, D.; Begg, B. E.; Braun, P.; Brehme, M.; Broly, M. P.; Carvunis, A.-R.; Convery-Zupan, D.; Corominas, R.; Coulombe-Huntington, J.; Dann, E.; Dreze, M.; Dricot, A.; Fan, C.; Franzosa, E.; Gebreab, F.; Gutierrez, B. J.; Hardy, M. F.; Jin, M.; Kang, S.; Kiros, R.; Lin, G. N.; Luck, K.; MacWilliams, A.; Menche, J.; Murray, R. R.; Palagi, A.; Poulin, M. M.; Rambout, X.; Rasla, J.; Reichert, P.; Romero, V.; Ruyssinck, E.; Sahalie, J. M.; Scholz, A.; Shah, A. A.; Sharma, A.; Shen, Y.; Spirohn, K.; Tam, S.; Tejeda, A. O.; Trigg, S. A.; Twizere, J.-C.; Vega, K.; Walsh, J.; Cusick, M. E.; Xia, Y.; Barabási, A.-L.; Iakoucheva, L. M.; Aloy, P.; De Las Rivas, J.; Tavernier, J.; Calderwood, M. A.; Hill, D. E.; Hao, T.; Roth, F. P.; Vidal, M. A proteome-scale map of the human interactome network. *Cell* **2014**, *159*, 1212−1226.
(75) Koh, G. C. K. W.; Porras, P.; Aranda, B.; Hermjakob, H.; Orchard, S. E. Analyzing protein-protein interaction networks. *J. Proteome Res.* **2012**, *11*, 2014−2031.
(76) Ideker, T.; Krogan, N. J. Differential network biology. *Mol. Syst. Biol.* **2012**, *8*, 565.
(77) Greene, C. S.; Krishnan, A.; Wong, A. K.; Ricciotti, E.; Zelaya, R. A.; Himmelstein, D. S.; Zhang, R.; Hartmann, B. M.; Zaslavsky, E.; Sealfon, S. C.; Chasman, D. I.; FitzGerald, G. A.; Dolinski, K.; Grosser, T.; Troyanskaya, O. G. Understanding multicellular function and disease with human tissue-specific networks. *Nat. Genet.* **2015**, *47*, 569−576.
(78) Vinayagam, A.; Gibson, T. E.; Lee, H.-J.; Yilmazel, B.; Roesel, C.; Hu, Y.; Kwon, Y.; Sharma, A.; Liu, Y.-Y.; Perrimon, N.; Barabási, A.-L. Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 4976−4981.
(79) Domling, A. *Protein-Protein Interactions in Drug Discovery*; Wiley, 2013; Vol. *56*, p 334.
(80) Wang, X.; Wei, X.; Thijssen, B.; Das, J.; Lipkin, S. M.; Yu, H. Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat. Biotechnol.* **2012**, *30*, 159−164.
(81) Neshich, I. A.; Kiyota, E.; Arruda, P. Genome-wide analysis of lysine catabolism in bacteria reveals new connections with osmotic stress resistance. *ISME J.* **2013**, *7*, 2400−2410.
(82) Sanders, R. J.; Ofman, R.; Dacremont, G.; Wanders, R. J. A.; Kemp, S. Characterization of the human omega-oxidation pathway for omega-hydroxy-very-long-chain fatty acids. *FASEB J.* **2008**, *22*, 2064−2071.
(83) van Roermund, C. W. T.; Ijlst, L.; Wagemans, T.; Wanders, R. J. A.; Waterham, H. R. A role for the human peroxisomal half-transporter ABCD3 in the oxidation of dicarboxylic acids. *Biochim. Biophys. Acta* **2014**, *1841*, 563−568.
(84) Herzog, K.; Pras-Raves, M. L.; Ferdinandusse, S.; Vervaart, M. A. T.; Luyf, A. C. M.; van Kampen, A. H. C.; Waterham, H. R.; Vaz, F. M. Functional characterisation of peroxisomal β-oxidation disorders in fibroblasts using lipidomics. *J. Inherit. Metab. Dis.* **2018**, *41*, 479−487.
(85) Coppa, A.; Guha, S.; Fourcade, S.; Parameswaran, J.; Ruiz, M.; Moser, A. B.; Schlüter, A.; Murphy, M. P.; Lizcano, J. M.; Miranda-Vizuete, A.; Dalfó, E.; Pujol, A. The peroxisomal fatty acid transporter ABCD1/PMP-4 is required in the C. elegans hypodermis for axonal maintenance: A worm model for adrenoleukodystrophy. *Free Radic. Biol. Med.* **2020**, *152*, 797−809.
(86) Exner, T.; Romero-Brey, I.; Yifrach, E.; Rivera-Monroy, J.; Schrul, B.; Zouboulis, C. C.; Stremmel, W.; Honsho, M.; Bartenschlager, R.; Zalckvar, E.; Poppelreuther, M.; Füllekrug, J. An alternative membrane topology permits lipid droplet localization of peroxisomal fatty acyl-CoA reductase 1. *J. Cell Sci.* **2019**, *132*, jcs223016.
(87) Kong, J.; Ji, Y.; Jeon, Y. G.; Han, J. S.; Han, K. H.; Lee, J. H.; Lee, G.; Jang, H.; Choe, S. S.; Baes, M.; Kim, J. B. Spatiotemporal contact between peroxisomes and lipid droplets regulates fasting-induced lipolysis via PEX5. *Nat. Commun.* **2020**, *11*, 578.

https://doi.org/10.1021/acs.jproteome.1c00330
J. Proteome Res. XXXX, XXX, XXX−XXX

8. Literaturverzeichnis

Aguirre-Plans, J. *et al.* (2018) “Proximal pathway enrichment analysis for targeting comorbid diseases via network endopharmacology,” *Pharmaceuticals*, 11(3), pp. 1–18. doi:10.3390/ph11030061.

Alanis-Lobato, G., Andrade-Navarro, M.A. and Schaefer, M.H. (2017) “HIPPIE v2.0: Enhancing meaningfulness and reliability of protein-protein interaction networks,” *Nucleic Acids Research*, 45(D1), pp. D408–D414. doi:10.1093/nar/gkw985.

Amberger, J.S. *et al.* (2015) “OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an Online catalog of human genes and genetic disorders,” *Nucleic Acids Research*, 43(D1), pp. D789–D798. doi:10.1093/nar/gku1205.

Antonicka, H. *et al.* (2020) “A High-Density Human Mitochondrial Proximity Interaction Network,” *Cell Metabolism*, 32(3), pp. 479-497.e9. doi:10.1016/j.cmet.2020.07.017.

Argmann, C.A. *et al.* (2016) “Perspective A Next Generation Multiscale View of Inborn Errors of Metabolism,” *Cell Metabolism*, 23(1), pp. 13–26. doi:10.1016/j.cmet.2015.11.012.

Ayoub, M. and Pfleger, K.D.G. (2010) “Recent advances in bioluminescence resonance energy transfer technologies to study GPCR heteromerization.,” *Current opinion in pharmacology*, 10(1), pp. 44–52. doi:10.1016/j.coph.2009.09.012.

Bacart, J. *et al.* (2008) “The BRET technology and its application to screening assays.,” *Biotechnology journal*, 3(3), pp. 311–24. doi:10.1002/biot.200700222.

Barabasi, A.-L. (2007) “Network Medicine — From Obesity to the ‘ Diseasome ,’” *The New England Journal of Medicine*, 357(4), pp. 404–407. doi:10.1056/NEJMe078114.

Barábasi, A.-L. and Albert, R. (1999) “Emergence of Scaling in Random Networks,” *Science*, 286(October), pp. 509–512. doi:10.1097/WNF.0B013E31803D357F.

Barabasi, A.-L. and Bonabeau, E. (2003) “Scale-Free Networks,” *Scientific American*, (May), pp. 50–59.

Barabási, A.-L., Gulbahce, N. and Loscalzo, J. (2011) “Network medicine: a network-based approach to human disease,” *Nature Review Genetics*, 12, pp. 56–68. doi:10.1038/nrg2918.

Barábasi, A.L. and Oltvai, Z.N. (2004) “Network Biology: Understanding the Cell’s Functional Organization,” *Nature Review*, 5(February), pp. 101–113. doi:10.1038/nrg1272.

Barbero-Camps, E. *et al.* (2014) “Endoplasmic reticulum stress mediates amyloid β neurotoxicity via mitochondrial cholesterol trafficking,” *American Journal of Pathology*, 184(7), pp. 2066–2081. doi:10.1016/j.ajpath.2014.03.014.

Barrios-Rodiles, M. *et al.* (2017) “Lumier: A discovery tool for mammalian protein interaction networks,” *Methods in Molecular Biology*, 1550, pp. 137–148. doi:10.1007/978-1-4939-6747-6_11.

Basha, O. *et al.* (2015) “MyProteinNet: Build up-to-date protein interaction networks for organisms, tissues and user-defined contexts,” *Nucleic Acids Research*, 43(W1), pp. W258–W263. doi:10.1093/nar/gkv515.

Basha, O. *et al.* (2017) “The TissueNet v.2 database: A quantitative view of protein-protein interactions across human tissues,” *Nucleic Acids Research*, 45(D1), pp. D427–D431. doi:10.1093/nar/gkw1088.

Blasche, S. and Koegl, M. (2013) “Analysis of protein-protein interactions using LUMIER assays,” *Methods in Molecular Biology*, 1064, pp. 17–27. doi:10.1007/978-1-62703-601-6_2.

Boute, N., Jockers, R. and Issad, T. (2002) “The use of resonance energy transfer in high-throughput screening: BRET versus FRET,” *Trends Pharmacol. Sci*, 23, pp. 351–354.

Braun, P. *et al.* (2009) “An experimentally derived confidence score for binary protein-protein interactions,” *Nature Methods*, 6(1), pp. 91–97. doi:10.1038/NMETH.1281.

Breuer, K. *et al.* (2013) “InnateDB: Systems biology of innate immunity and beyond - Recent updates and continuing curation,” *Nucleic Acids Research*, 41(D1), pp. 1228–1233. doi:10.1093/nar/gks1147.

Cafarelli, T. *et al.* (2017) “Mapping, modeling, and characterization of protein–protein interactions on a proteomic scale,” *Current Opinion in Structural Biology*, 44, pp. 201–210. doi:10.1016/j.sbi.2017.05.003.

Caldera, M. *et al.* (2017) “Interactome-based approaches to human disease,” *Current Opinion in Systems Biology*, 3, pp. 88–94. doi:10.1016/j.coisb.2017.04.015.

Carter, H., Hofree, M. and Ideker, T. (2013) “Genotype to phenotype via network analysis,” *Current Opinion in Genetics and Development*, 23(6), pp. 611–621. doi:10.1016/j.gde.2013.10.003.

Cheng, F. *et al.* (2018) “Network-based approach to prediction and population-based validation of in silico drug repurposing,” *Nature Communications*, 9(1), pp. 1–12. doi:10.1038/s41467-018-05116-5.

Choi, S.G. *et al.* (2019) “Maximizing binary interactome mapping with a minimal number of assays,” *Nature Communications*, 10(1). doi:10.1038/s41467-019-11809-2.

Collins, F.S. *et al.* (2004) “Finishing the euchromatic sequence of the human genome,” *Nature*, 431(7011), pp. 931–945. doi:10.1038/nature03001.

Conte, F. *et al.* (2020) “A paradigm shift in medicine: A comprehensive review of network-based approaches,” *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1863(6), p. 194416. doi:10.1016/j.bbagrm.2019.194416.

Couturier, C. and Deprez, B. (2012) “Setting Up a Bioluminescence Resonance Energy Transfer High throughput Screening Assay to Search for Protein/Protein Interaction Inhibitors in Mammalian Cells.,” *Frontiers in endocrinology*, 3(September), p. 100. doi:10.3389/fendo.2012.00100.

Das, J. and Yu, H. (2012) “HINT: High-quality protein interactomes and their applications in understanding human disease,” *BMC Systems Biology*, 6. doi:10.1186/1752-0509-6-92.

Edwards, S.L. *et al.* (2013) “Beyond GWASs: Illuminating the dark road from association to function,” *American Journal of Human Genetics*, 93(5), pp. 779–797. doi:10.1016/j.ajhg.2013.10.012.

Ellis, J.D. *et al.* (2012) “Tissue-Specific Alternative Splicing Remodels Protein-Protein Interaction Networks,” *Molecular Cell*, 46(6), pp. 884–892. doi:10.1016/j.molcel.2012.05.037.

Ewing, R.M. *et al.* (2007) “Large-scale mapping of human protein-protein interactions by mass spectrometry.,” *Mol Syst Biol*, 3(89), p. 89. doi:10.1038/msb4100134.

Frixel, S. *et al.* (2016) “Homooligomerization of ABCA3 and its functional significance,” *International Journal of Molecular Medicine*, 38, pp. 558–566. doi:10.3892/ijmm.2016.2650.

Gersting, S.W., Lotz-Havla, A.S. and Muntau, A.C. (2012) “Bioluminescence Resonance Energy Transfer: An Emerging Tool for the Detection of Protein–Protein Interaction in Living Cells,” in Kaufmann, M. and Klinger, C. (eds) *Methods in Molecular Biology*. Springer New York (Methods in Molecular Biology), pp. 253–263. doi:10.1007/978-1-61779-424-7.

Ghiassian, S.D., Menche, J. and Barabási, A.-L. (2015) “A DIseAse MOdule Detection (DIAMOnD) Algorithm Derived from a Systematic Analysis of Connectivity Patterns of Disease Proteins in the Human Interactome,” *PLOS Computational Biology*, 11(4), p. e1004120. doi:10.1371/journal.pcbi.1004120.

Gordon, D.E. *et al.* (2020) “A SARS-CoV-2 protein interaction map reveals targets for drug repurposing,” *Nature*, 583(July). doi:10.1038/s41586-020-2286-9.

Gstaiger, M. and Aebersold, R. (2013) “Genotype-phenotype relationships in light of a modular protein interaction landscape,” *Molecular BioSystems*, 9(6), pp. 1064–1067. doi:10.1039/c3mb25583b.

Guder, P. *et al.* (2018) “Isoform-specific domain organization determines conformation and function of the peroxisomal biogenesis factor PEX26,” *Biochimica et Biophysica Acta - Molecular Cell Research*, 1866, pp. 518–531. doi:10.1016/j.bbamcr.2018.10.013.

Gupta, G.D. *et al.* (2015) “A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface,” *Cell*, 163(6), pp. 1484–1499. doi:10.1016/j.cell.2015.10.065.

Hamdan, F.F. *et al.* (2005) “High-Throughput Screening of G Protein-Coupled Receptor Antagonists Using a Bioluminsescnece Resonance Energy Transfer 1-Based beta-Arrestin2 Recruitment Assay,” *The Society for Biomolecular Screening*, 10.

Hamdi, A. and Colas, P. (2012) “Yeast two-hybrid methods and their applications in drug discovery.,” *Trends in pharmacological sciences*, 33(2), pp. 109–18. doi:10.1016/j.tips.2011.10.008.

Hartwell, L.H. *et al.* (1999) “From molecular to modular cell biology,” *Nature*, 402(December), pp. 47–52.

Havugimana, P.C. *et al.* (2013) “A Census of Human Soluble Protein Complexes,” *Cell*, 150(5), pp. 1068–1081. doi:10.1016/j.cell.2012.08.011.A.

Hein, M.Y. *et al.* (2015) “A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances,” *Cell*, 163, pp. 712–723. doi:10.1016/j.cell.2015.09.053.

Hillebrand, M. *et al.* (2012) “Identification of a new fatty acid synthesis-transport machinery at the peroxisomal membrane,” *Journal of Biological Chemistry*, 287(1), pp. 210–221. doi:10.1074/jbc.M111.272732.

Huttlin, E.L. *et al.* (2015) “The BioPlex Network: A Systematic Exploration of the Human Interactome,” *Cell*, 162(2), pp. 425–440. doi:10.1016/j.cell.2015.06.043.

Huttlin, E.L. *et al.* (2021) “Dual proteome-scale networks reveal cell-specific remodeling of the human interactome,” *Cell*, 184(11), pp. 3022-3040.e28. doi:10.1016/j.cell.2021.04.011.

Kitsak, M. *et al.* (2016) “Tissue Specificity of Human Disease Module,” *Nature Scientific Reports*, (October), pp. 1–12. doi:10.1038/srep35241.

Kobayashi, H. *et al.* (2019) “Bioluminescence resonance energy transfer–based imaging of protein–protein interactions in living cells,” *Nature Protocols*, 14(4), pp. 1084–1107. doi:10.1038/s41596-019-0129-7.

Kocan, M. *et al.* (2008) “Demonstration of improvements to the bioluminescence resonance energy transfer (BRET) technology for the monitoring of G protein-coupled receptors in live cells.,” *Journal of biomolecular screening : the official journal of the Society for Biomolecular Screening*, 13(9), pp. 888–98. doi:10.1177/1087057108324032.

Koh, G.C.K.W. *et al.* (2012) “Analyzing protein-protein interaction networks.,” *Journal of proteome research*, 11(4), pp. 2014–31. doi:10.1021/pr201211w.

Kotlyar, M. *et al.* (2019) “IID 2018 update: Context-specific physical protein-protein interactions in human, model organisms and domesticated species,” *Nucleic Acids Research*, 47(D1), pp. D581–D589. doi:10.1093/nar/gky1037.

Law, V. *et al.* (2014) “DrugBank 4.0: Shedding new light on drug metabolism,” *Nucleic Acids Research*, 42(D1), pp. 1091–1097. doi:10.1093/nar/gkt1068.

Lemmens, I., Lievens, S. and Tavernier, J. (2010) “Strategies towards high-quality binary protein interactome maps.,” *Journal of proteomics*, pp. 1–6. doi:10.1016/j.jprot.2010.02.001.

Lempiäinen, H. *et al.* (2018) “Network analysis of coronary artery disease risk genes elucidates disease mechanisms and druggable targets,” *Scientific Reports*, 8(3434). doi:10.1038/s41598-018-20721-6.

Li, J. *et al.* (2014) “Integrated systems analysis reveals a molecular network underlying autism spectrum disorders,” *Molecular Systems Biology*, 10(12), pp. 774–774. doi:10.15252/msb.20145487.

Li, P., Wang, L. and Di, L.J. (2019) “Applications of protein fragment complementation assays for analyzing biomolecular interactions and biochemical networks in living cells,” *Journal of Proteome Research*, 18(8), pp. 2987–2998. doi:10.1021/acs.jproteome.9b00154.

Licata, L. *et al.* (2012) “MINT, the molecular interaction database: 2012 Update,” *Nucleic Acids Research*, 40(D1), pp. 857–861. doi:10.1093/nar/gkr930.

Lievens, S. *et al.* (2009) “Array MAPPIT : High-Throughput Interactome Analysis in Mammalian Cells research articles,” *Journal of Proteome Research*, pp. 877–886.

Lievens, S. *et al.* (2016) “Proteome-scale binary interactomics in human cells,” *Molecular and Cellular Proteomics*, 15(12), pp. 3624–3639. doi:10.1074/mcp.M116.061994.

Loscalzo, J. and Barábasi, A.L. (2011) “Systems biology and the future of medicine,” *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 3(6), pp. 619–627. doi:10.1002/wsbm.144.

Luck, K. *et al.* (2017) “Proteome-Scale Human Interactomics,” *Trends in Biochemical Sciences*, 42(5), pp. 342–354. doi:10.1016/j.tibs.2017.02.006.

Luck, K. *et al.* (2020) “A reference map of the human binary protein interactome,” *Nature*, 580(7803), pp. 402–408. doi:10.1038/s41586-020-2188-x.

McDowall, M.D., Scott, M.S. and Barton, G.J. (2009) “PIPs: Human protein-protein interaction prediction database,” *Nucleic Acids Research*, 37(SUPPL. 1), pp. 651–656. doi:10.1093/nar/gkn870.

Menche, J. *et al.* (2015) “Uncovering disease-disease relationships through the human interactome,” *Science*, 347(6224), p. 1257601. doi:10.1126/science.1257601.

Mercer, S.W., Wang, J. and Burke, R. (2017) “In vivo modeling of the pathogenic effect of copper transporter mutations that cause Menkes and Wilson diseases, motor neuropathy, and susceptibility to Alzheimer’s disease,” *Journal of Biological Chemistry*, 292(10), pp. 4113–4122. doi:10.1074/jbc.M116.756163.

Miller, K.E. *et al.* (2016) “BiFC analysis: advances and recent applications for genome-wide interaction studies,” *Journal of Molecular Biology*, 427(11), pp. 2039–2055. doi:10.1016/j.jmb.2015.03.005.Bimolecular.

Mo, X.-L. *et al.* (2015) “Enabling systematic interrogation of protein-protein interactions in live cells with a versatile ultra-high-throughput biosensor platform.,” *Journal of molecular cell biology*, 0, pp. 1–42. doi:10.1093/jmcb/mjv064.

Moutaoufik, M.T. *et al.* (2019) “Rewiring of the Human Mitochondrial Interactome during Neuronal Reprogramming Reveals Regulators of the Respirasome and Neurogenesis,” *iScience*, 19, pp. 1114–1132. doi:10.1016/j.isci.2019.08.057.

Nold-Petry, C.A. *et al.* (2015) “IL-37 requires the receptors IL-18Rα and IL-1R8 (SIGIRR) to carry out its multifaceted anti-inflammatory program upon innate signal transduction,” *Nature Immunology*, 16(4), pp. 354–365. doi:10.1038/ni.3103.

Orchard, S. *et al.* (2014) “The MIntAct project - IntAct as a common curation platform for 11 molecular interaction databases,” *Nucleic Acids Research*, 42(D1), pp. 358–363. doi:10.1093/nar/gkt1115.

Oti, M. and Brunner, H.G. (2007) “The modular nature of genetic diseases,” *Clinical Genetics*, 71(1), pp. 1–11. doi:10.1111/j.1399-0004.2006.00708.x.

Oughtred, R. *et al.* (2021) “The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions,” *Protein Science*, 30(1), pp. 187–200. doi:10.1002/pro.3978.

Pankow, S. *et al.* (2015) “ΔF508 CFTR interactome remodelling promotes rescue of cystic fibrosis,” *Nature*, 528(7583), pp. 510–516. doi:10.1038/nature15729.

Petschnigg, J. *et al.* (2014) “The mammalian-membrane two-hybrid assay (MaMTH) for probing membrane-protein interactions in human cells,” *Nature Methods*, 11(5), pp. 585–592. doi:10.1038/nmeth.2895.

Pfleger, K.D. and Eidne, K.A. (2006) “Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET),” *Nat. Methods*, 3, pp. 165–174.

Piñero, J. *et al.* (2016) “Uncovering disease mechanisms through network biology in the era of Next Generation Sequencing,” *Scientific Reports*, 6(24570), pp. 1–12. doi:10.1038/srep24570.

Prasad, K.T.S. *et al.* (2009) “Human Protein Reference Database - 2009 update,” *Nucleic Acids Research*, 37(SUPPL. 1), pp. 767–772. doi:10.1093/nar/gkn892.

Robinson, K.H., Yang, J.R. and Zhang, J. (2014) “FRET and BRET-Based Biosensors in Live Cell Compound Screens,” *Methods in Molecular Biology*, 1071, pp. 217–225. doi:10.1007/978-1-62703-622-1.

Rolland, T. *et al.* (2014) “A Proteome-Scale Map of the Human Interactome Network,” *Cell*, 159, pp. 1212–1226. doi:10.1016/j.cell.2014.10.050.

Rual, J., Venkatesan, K. and Hao, T. (2005) “Towards a proteome-scale map of the human protein – protein interaction network,” *October*, 437(October), pp. 1173–1178. doi:10.1038/nature04209.

Sahni, N. *et al.* (2013) “Edgotype: a fundamental link between genotype and phenotype.,” *Current opinion in genetics & development*, 23(6), pp. 649–57. doi:10.1016/j.gde.2013.11.002.

Sahni, N. *et al.* (2015) “Widespread Macromolecular Interaction Perturbations in Human Genetic Disorders,” *Cell*, 161(3), pp. 647–660. doi:10.1016/j.cell.2015.04.013.

Sharma, A. *et al.* (2015) “A disease module in the interactome explains disease heterogeneity, drug response and captures novel pathways and genes in asthma,” *Human Molecular Genetics*, 24(11), pp. 3005–3020. doi:10.1093/hmg/ddv001.

Sharma, A. *et al.* (2018) “Integration of Molecular Interactome and Targeted Interaction Analysis to Identify a COPD Disease Network Module,” *Scientific reports*, 8(August), p. 408229. doi:10.1101/408229.

Silverman, E.K. *et al.* (2020) *Molecular networks in Network Medicine: Development and applications*, *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. doi:10.1002/wsbm.1489.

Snider, J. *et al.* (2010) “Detecting interactions with membrane proteins using a membrane two-hybrid assay in yeast,” *Nature Protocols*, 5(7), pp. 1281–1293. doi:10.1038/nprot.2010.83.

Sonawane, A.R. *et al.* (2019) “Network Medicine in the age of biomedical big data,” *Frontiers in Genetics*, 10(April), pp. 1–16. doi:10.3389/fgene.2019.00294.

Stelzl, U. *et al.* (2005) “A human protein-protein interaction network: a resource for annotating the proteome,” *Cell*, 122, pp. 957–968.

Stumpf, M.P.H. *et al.* (2008) “Estimating the size of the human interactome MATHEMATICS,” *PNAS*, 105(19), pp. 6959–6964. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2383957/?tool=pubmed.

Szklarczyk, D. *et al.* (2021) “The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets,” *Nucleic Acids Research*, 49(D1), pp. D605–D612. doi:10.1093/nar/gkaa1074.

Tarassov, K. *et al.* (2008) “An in Vivo Map of the Yeast Protein Interactome,” *Science*, 320(June), pp. 1465–1470.

Venkatesan, K. *et al.* (2009) “An empirical framework for binary interactome mapping,” *Nat. Methods*, 6, pp. 83–90.

Vidal, M. (2016) “How much of the human protein interactome remains to be mapped?,” *Science Signaling*, 9(427), pp. eg7–eg7. doi:10.1126/scisignal.aaf6030.

Vidal, M., Cusick, M.E. and Barabási, A.-L. (2011) “Interactome networks and human disease.,” *Cell*, 144(6), pp. 986–98. doi:10.1016/j.cell.2011.02.016.

Vidal, M. and Fields, S. (2014) “The yeast two-hybrid assay: still finding connections after 25 years,” *Nature Methods*, 11(12), pp. 1203–1206. doi:10.1038/nmeth.3182.

Vizoso Pinto, M.G. *et al.* (2011) “Varicella zoster virus ORF25 gene product: An essential hub protein linking encapsidation proteins and the nuclear egress complex,” *Journal of Proteome Research*, 10(12), pp. 5374–5382. doi:10.1021/pr200628s.

Vyncke, L. *et al.* (2019) “Straightforward protein-protein interaction interface mapping via random mutagenesis and mammalian protein protein interaction trap (MAPPIT),” *International Journal of Molecular Sciences*, 20(9). doi:10.3390/ijms20092058.

Wang, R.S. and Loscalzo, J. (2018) “Network-Based Disease Module Discovery by a Novel Seed Connector Algorithm with Pathobiological Implications,” *Journal of Molecular Biology*, 430(18), pp. 2939–2950. doi:10.1016/j.jmb.2018.05.016.

Xu, Y., Piston, D.W. and Johnson, C.H. (1999) “A bioluminescence resonance energy transfer (BRET) system: application to interacting circadian clock proteins.,” *Proceedings of the National Academy of Sciences of the United States of America*, 96(1), pp. 151–6. doi:10.1073/pnas.96.1.151.

Yadav, A., Vidal, M. and Luck, K. (2020) “Precision medicine — networks to the rescue,” *Current Opinion in Biotechnology*, 63, pp. 177–189. doi:10.1016/j.copbio.2020.02.005.

Yang, X. *et al.* (2016) “Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing,” *Cell*, 164(4), pp. 805–817. doi:10.1016/j.cell.2016.01.029.

Zhang, X., Kuivenhoven, J.A. and Groen, A.K. (2015) “Forward individualized medicine from personal genomes to interactomes,” *Frontiers in Physiology*, 6(DEC), pp. 1–11. doi:10.3389/fphys.2015.00364.

Zhong, Q. *et al.* (2009) “Edgetic perturbation models of human inherited disorders.,” *Molecular systems biology*, 5(321), p. 321. doi:10.1038/msb.2009.80.

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich auf dieser langen Reise stets unterstützt haben.

Zuallererst möchte ich meiner Doktormutter Frau Prof. Dr. med. Ania C. Muntau für die Möglichkeit danken, diese Doktorarbeit durchzuführen. Ich danke auch für Ihre Unterstützung bei meinem Wunsch des Zweitstudiums der Humanmedizin. Ohne diese Unterstützung wäre diese Arbeit nicht zustande gekommen. Neben Ihrem Engagement für die Forschung und den Bemühungen die wissenschaftlichen Ergebnisse stets in die Behandlung von Patienten zu überführen hat mich auch Ihre vorgelebte Begeisterung für die Stoffwechselmedizin stets beeindruckt. Dies führte auch dazu, dass mein Interesse an dieser Fachrichtung der Pädiatrie mehr und mehr wuchs und ich die Ausbildung zum Stoffwechselmediziner zuletzt mit Freude begonnen habe.

Ein besonderer Dank gilt auch Herrn Prof. Dr. med. Søren W. Gersting, der mir während meiner Doktorarbeit allzeit hilfreich zur Seite stand, mich immer wieder angespornt hat und mich durch seine positive Art die Dinge zu denken nie aufgeben hat lassen. Für Probleme fand ich immer ein offenes Ohr bei ihm, Lösungen wurden gemeinsam erörtert und auch die ein oder andere abendliche Laborstunde geriet so zu einer kurzweiligen und dennoch erfolgreichen Angelegenheit.

Natürlich möchte ich mich auch bei allen zahlreichen wissenschaftlichen Weggefährten der letzten Jahre für interessante Gespräche und hilfreiche Ratschläge bedanken. Ein Dankeschön gilt dabei insbesondere Julian Klingbeil, Marta Danecka, Madalena Barroso, Amelie Lotz-Havla, Philipp Guder und Anna Waldhuber.

Diese Arbeit wäre auch nie ohne die vielfältige Unterstützung meiner Eltern Petra Woidy-Kellner und Ludwig Woidy, sowie meiner Geschwister Patricia, Viktoria, Lukas, Lisa und Romina möglich gewesen. Bedanken möchte ich mich auch bei meinen Schwiegereltern Judith und Jürgen Schebesta.

Zu guter Letzt gilt ein besonderer Dank meiner Frau Isabelle und meinem Sohn Emil. Ohne ihre bedingungslose Unterstützung, dauerhafte Geduld und Akzeptanz für meine Arbeit wäre diese Promotion, unser gemeinsamer Umzug nach Hamburg sowie der Beginn meiner klinischen Tätigkeit nicht möglich gewesen.