Zusammenfassung:

Herausforderungen in der RNA-Biologie:

* Umfassende, spezifische Identifizierung von RNA-bindenden Proteinen
* Entdeckung von Funktionen von RNA-assoziierten Proteinen

Ziel im Paper:

* Analyse des Konzepts der RNA-Abhängigkeit
* Definieren ein Protein als RNA-abhängig, wenn sein Interaktom von RNA abhängt.

Vorgehensweise des Konzeptes:

* Proteom-weites, unvoreingenommenes und anreicherungsfreies Screening namens R-deep (RNA-abhängige Proteine), basierend auf Dichte Gradienten-Ultrazentrifugation.
* Quantitative Massenspektrometrie:
* 1784 RNA-abhängige Proteine und 537 davon haben keine Verbindung zur RNA
* Nutzung von der quantitativen Natur von R-Deep:

1. Proteine, die nicht RNA-abhängig sind
2. Proteine, die teilweise RNA-abhängig sind
3. Proteine, die vollständig RNA-abhängig sind

* R-Deep identifiziert den Transkriptionsfaktor CTCF bei vollständiger RNA-abhängig
* RNA ist für CTCF-Chromatin-Assoziation erforderlich
* R-Deep: Rekonstruktion von Proteinkomplexen auf Grundlagen der Ko-Segregation

Einleitung:

Transkription:

* Assoziierte RNA-Moleküle mit verschiedenen RNA-bindenden Proteinen (RBPs) um dann Ribonukleoprotein (RNP) Komplexe zu bilden
* RBPs -> Schicksal der RNA-Transkripts beeinflussen
* RNA -> Schicksal der interagierenden Proteine beeinflussen
* RNP -> dynamische Gebilde (Proteinzusammensetzung um die RNA) -> Änderung der RNA-Verarbeitung, -Reifung, -Transport & -Lokalisierung
* Einige Proteine binden konstitutive an RNA, andere dissoziieren oder binden vorübergehend später
* Der dynamische Austausch von RBPs innerhalb von RNPs auf eine Vielzahl von zellulären und umweltbedingten Stimuli abhängig.
* RBPs-Spektrum
* Schlüsselfunktion im RNA-Stoffwechsel und Regulation bei der Genexpression

! -> Defekte in RBP-Funktionen = Krankheit (von neurodegenerativen Störungen bis zu Krebs)

* RBPs beteiligen sich an der Biogenese und Funktionen von kodierenden Boten-RNAs und assoziieren auch mit nicht-kodierenden RNAs (ncRNAs) -> Transkription den Großteil des Genoms.
* Lange nicht.-codierte RNAs (lncRNAs)
* Bildung von Klassen von ncRNAs
* Beteiligung in physiologischen, pathologischen Prozessen und Krebs

! -> Verständnis der Bedeutung von lncRNA -> Neue Perspektive auf Funktionsrepertoire

! -> Schwierigkeit ist die spezifischen molekularen Interaktionen zwischen einzelnen lncRNAs und Interaktionen aufzuklären.

* Herausforderung in RNA-Biologie:
* Entdeckung neuer Mechanismen und des Gesamtspektrums vom Komplex durch die Beeinflussung von (nc)RNAs
* Analyse:
* RBPs von verschieden Säugertiersystmen mit proteomweitem Ansatz

(Pull-down-Assays von polyadenylierten (poly(A)+) RNAs)

* Beweis von Vorhandensein vieler RBPs ohne klassische RNA-bindenden Domäne und auch mit intrinsischen ungeordneten Domänen
* Entdeckung:

Protease Verdauung -> modifizierten Nukleotiden oder org. Phasen- Trennung von proteinvernetzen RNAs wurden entwickelt, um auch Nicht-Poly(A)+-RNA einzubeziehen und die RNA-Bindungsdomänen zu charakterisieren.

! -> Liste der RBP-Kandidaten

1. Überschidung von 8 humanen protemweiten Proteomen nach RBPs
2. 215 Proteine zwischen den Überschneidungen
3. Großteil von Proteinen, die Nicht-Poly(A)+-RNA enthalten

* Komplementäre Strategien:
* Definierung vom Kernsatz von RBPs unabhängig von affinitäts- oder eigenschaftsbasierten Aufreinigungen

ZIEL:

Die Gesamtheit der von RNA betroffenen Proteine zu bestimmen und zu quantifizieren Proteine zu bestimmen und zu quantifizieren und den Weg für die Identifizierung zusätzlicher Wege und Mechanismen, die durch RNA kontrolliert werden.

* Darstellung von einer proteomweite Anreicherungsfreien und quantitativen Screening nach Proteinen, deren Interaktom von RNA abhängt (537 RNA-abhängige Proteine, die nicht mit RNA in Verbindung gebracht wurden z. B HMGN1, CASP7, REEP4 und THYN1)
* Charakterisierung der vollständigen RNA
* RNA-Abhängigkeit des Transkriptionsfaktors CTCF
* RNA wichtiger Rekrutierungs-Faktor von CTCF an Chromatin

Ergebnisse:

* R-Deep: proteomweiter Screen zur Identifizierung von RNA-abhängige Proteinen
* Neue Definierung des Konzeptes der RNA-Abhängigkeit
* Bezeichnung von einem Protein als RNA-abhängig, wenn sein Interaktom von der RNA abhängt, ohne notwendigerweise direkt an RNA zu binden.
* Differenzierung von dem Begriff „RNA-Bindung“: da RNA-abhängige Proteine an Komplexen beteiligt sind, die an RNA gebunden sind, zu denen auch Proteine gehören, die nicht direkt mit RNA interagieren RNA interagieren.
* Konzeptermöglichung: spezifische und quantitative Entdeckung der RNA-Abhängigkeit von Proteinen und Komplexen, auch aus bisher nicht mit RNA in Verbindung gebrachten Stoffwechselwegen.
* Suche nach spezifisch, unvoreingenommen und ohne Anreicherungsstrategien RNA-abhängigen Proteinen (R-DeeP)
* Z. B HeLa S3 Zelllysat auf Dichtegradienten von 5% bis 50% und fraktionierten es nach Ultrazentrifugation in 25 Fraktionen und der Proteingehalt der einzelnen Fraktion wurde mittels Western Blot oder Massenspektrometrie analysiert (Sichtbar: Jedes Protein zeigt eine spezifische Verteilung über den Gradienten auf)
* Eine Verschiebung zwischen den Fraktionen des Kontrollgradienten und des RNase-behandelten Gradienten zeigte die RNA-Abhängigkeit des Proteins (Abbildung 1A)
* Eine Qualitätskontrolle des RNase-Verdaus zeigte, dass RNAs bis zu einer Größe von deutlich unter 100 nt verdaut wurden (Abbildungen S1B und S1C), was zu einem vollständigen Verlust des Nachweises für kleinen und großen RNA-Spezies führte (Abbildung S1D). Darüber hinaus wurden signifikante Verschiebungen für Proteine wie DICER und ARGONAUTE2 beobachtet, von denen bekannt ist, dass sie mit kleinen RNA-Spezies interagiere
* Erzeugung der Kontroll- und RNase-behandelte Saccharose-Dichtegradienten in dreifacher Ausfertigung und quantifizierten die Häufigkeit der Proteine in jeder Fraktion mittels quantitativer Proteomik auf der Grundlage der Massenspektrometrie (-> 4.765 Proteine reproduzierbar identifizieren und quantifizieren)
* Entwickelung einer umfassenden statistische Analyse zur RNA-abhängige Proteine zu identifizieren, basierend auf der Gauß-angepassten Verteilung der einzelnen Proteine im Gradienten (STAR Methods)

Die Verschiebungen wurden charakterisiert durch

(1) die Position der Maxima

in den Kontroll- und RNase-behandelten Gradienten

(2) die Menge der Proteinverschiebung, dargestellt durch die Fläche unter der Gaußschen Anpassungskurve

(3) die Entfernung und Richtung der Verschiebung

(4) die Amplitudendifferenz Differenz bei jedem Maximum zwischen den Kontroll- und den RNase-angepassten Kurven

(5) die statistische Signifikanz der Differenz

* Die Kriterien für eine RNA-abhängige Verschiebung wurden definiert als ein Abstand von mehr als 1 Fraktion und ein signifikanter Unterschied (Falschentdeckungsrate FDR-korrigiert p < 0,05)
* fraktionsweisen Normalisierungsschritt korrelierten die Mengen der einzelnen Proteine pro Fraktion für alle
* Beim Vergleich von Wiederholungspaaren korrelierte die Menge jedes Proteins pro Fraktion für alle Proteine gut, was auf die Reproduzierbarkeit der Methode hinweist (Abbildung 1C; Pearson-Koeffizient R > 0.99)
* Proteine, deren Verteilung signifikante Unterschiede zwischen Kontrollgradienten und RNase-behandelten Gradienten aufwies, wurden in Unterkategorien mit verschiedenen Arten von Verschiebungen eingeteilt: Linksverschiebungen (Verschiebungen in Richtung der Fraktionen mit geringerer Saccharosedichte) als die größte Kategorie, Rechtsverschiebungen (zu höheren Fraktionen) und ausgefällte Proteine (akkumuliert in der letzten Fraktion; Abbildung 1D).
* Anhand dieser gründlichen statistischen Analyse bestimmt man die angepassten Peak-Positionen für jedes Protein und fanden eine hohe Korrelation zwischen Wiederholungen (R > 0,99; Abbildung S1G)
* Kategorisierung:

1. Linksverschiebung (1.931),
2. Rechtsverschiebung (120)
3. ausgefällt (425)
4. die Mehrheit der Proteine keine signifikante Verschiebung zu beobachten
5. (2.857, keine Verschiebung) (Abbildung 1E)
6. Insgesamt 1.784 Proteine, die eine signifikante Verschiebung aufwiesen (Abbildung 1F).

* Vergleich mit 19 früheren Studien und Katalogen von RBPs:

1. 1.247 (69,9 %) RNA-abhängige Proteine, die zuvor mindestens einmal als mindestens einmal als RNA-bindend (Shift und RBP) aufgeführt
2. 537 (30,1%) RNA-abhängige Proteine noch nie mit RNA in Verbindung gebracht worden waren RNA verknüpft waren (Shift und kein RBP)
3. Umgekehrt zeigten 1.402 Proteine, die zuvor als RBP-Kandidaten aufgelistet waren, zeigten keine signifikanten RNA-abhängigen Verschiebungen (keine Verschiebung und RBP)

* Um die spezifischen Eigenschaften der Proteinunterkategorien weiter zu analysieren, untersuchte man ihren isoelektrischen Punkt (pI), den Gehalt an niedrigkomplexen Domänen (LCD), die Aminosäurezusammensetzung und Proteindomänen.
* Shifting-Proteine hatten einen signifikant höheren durchschnittlichen pI als Nicht-Shifting-Proteine, unabhängig von ihrer Identifizierung als RBP-Kandidaten (Abbildung 1G)