Published in final edited form as:
J Proteome Res. 2012 December 7; 11(12): 6111–6123. doi:10.1021/pr300760p.

Proteomic Profile Identifies Dysregulated Pathways in Cornelia de Lange Syndrome Cells With Distinct Mutations in *SMC1A* and *SMC3* Genes

Anna Gimigliano[†,‖], **Linda Mannini**[‡,‖], **Laura Bianchi**[†], **Michele Puglia**[†], **Matthew A. Deardorff**[⊥], **Stefania Menga**[‡], **Ian D Krantz**[⊥], **Antonio Musio**[‡,*], and **Luca Bini**[†]

[†]Functional Proteomics Laboratory, Department of Biotechnologies, University of Siena, Siena, Italy

[‡]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Pisa, Italy

[⊥]Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

Abstract

Mutations in cohesin genes have been identified in Cornelia de Lange syndrome (CdLS), but its etiopathogenetic mechanisms are still poorly understood. To define biochemical pathways that are affected in CdLS we analyzed the proteomic profile of CdLS cell lines carrying mutations in the core cohesin genes, *SMC1A* and *SMC3*. Dysregulated protein expression was found in CdLS probands compared to controls. The proteomics analysis was able to discriminate between probands harboring mutations in the different domains of the SMC proteins. In particular, proteins involved in the response to oxidative stress were specifically down-regulated in hinge mutated probands. In addition, the finding that CdLS cell lines show an increase in global oxidative stress argues that it could contribute to some CdLS phenotypic features such as premature physiological aging and genome instability. Finally, the *c-MYC* gene represents a convergent hub lying at the center of dysregulated pathways, and is down-regulated in CdLS. This study allowed us to highlight, for the first time, specific biochemical pathways that are affected in CdLS, providing plausible causal evidence for some of the phenotypic features seen in CdLS.

Keywords

Cohesin; Cornelia de Lange syndrome; SMC1A; SMC3; 2D-DIGE; proteomic profile; dysregulated protein expression; c-Myc

INTRODUCTION

Cornelia de Lange Syndrome (CdLS, OMIM 122470, 300590 and 610759) is a rare multisystem disorder characterized by distinctive craniofacial dysmorphia, upper limb malformations, hirsutism, microcephaly, cardiac defects, gastroesophageal dysfunction, growth retardation and neurodevelopment impairment ranging from moderate to severe,

[*]To whom correspondence may be addressed, Tel: +39 0503152776, Fax: +39 0503153973, antonio.musio@irgb.cnr.it.
[‖]The authors wish it to be known that in their opinion, the first two authors should be regarded as joint First Authors

Supporting information
Supplementary figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org

CONFLICT OF INTEREST
The authors declare no competing financial interest.

with a wide range of variability.[1] Its prevalence is estimated to be 1 in 10,000 births, with most cases being sporadic. CdLS results from mutations in the highly evolutionarily conserved components of the cohesin pathway that mediate cohesion between replicated sister chromatids in dividing cells, in order to ensure proper chromosome segregation. Human core cohesin is composed of four members, SMC1A, SMC3, RAD21 and STAG1/2. Cohesin interacts with many proteins that regulate its activity.[2] About 60% of CdLS probands have been found to have heterozygous mutations in the *NIPBL* gene, which encodes a cohesin regulatory protein homologous to the fungal Scc2-type and the *Drosophila* Nipped-B.[3, 4] Most of the *NIPBL* mutations are point mutations, small insertions and deletions in coding regions or splice junctions. They are expected to produce absent or truncated proteins, so that haploinsufficiency is the presumed pathogenetic mechanism.[5] NIPBL is involved in loading cohesin onto chromatin, and a role in unloading has also been suggested. About 5% of CdLS cases are caused by mutations in *SMC1A*, and one proband was found to have a mutation in *SMC3*.[6-8] All these mutations are missense or small in-frame deletions. To date, no truncating mutations have been identified, suggesting that gross rearrangements, as occur in *NIPBL*, are negatively selected.[9] Very recently, mutations in the *HDAC8* gene, vertebrate SMC3 deacetylase, have been identified in CdLS probands,[10] while mutations in *RAD21* have been associated with a human cohesinopathy.[11]

SMC1A and *SMC3* mutations are associated with milder CdLS phenotype, and *SMC*-mutated probands often demonstrate a normal birth weight and head circumference, mild facial dysmorphism, mild intellectual disability and no limb defects.

The lack of evident defects in chromatid cohesion in CdLS probands[12] argues that the developmental perturbations seen in CdLS do not originate from a deficiency in cohesin's canonical role. Non-canonical roles for cohesin in genome integrity maintenance and in gene expression regulation have been proposed. Cohesin down-regulation leads to G2/M checkpoint defects and genome instability.[13-15] Mutations in *NIPBL, SMC1A* and *SMC3* genes have been identified in colorectal cancer [16] while the dysregulation of core cohesin genes and its regulator factors occurs in many cancers.[17, 18] Finally, CdLS cell lines show spontaneous genome instability and a reduced capacity to tolerate DNA damage following genotoxic treatments.[12, 19] The first study linking cohesin to transcription regulation was the finding that *Nipped-B* mutations promote the activation of the Drosophila *cut* and *Ultrabithorax* homeobox genes by distant transcriptional enhancers.[20] Furthermore, cohesin facilitates the expression of genes in the ecdysone steroid hormone signaling pathway[21, 22] as well as pluripotency genes in embryonic stem cells[23] and the expression of *myc* genes, which promote cell proliferation.[24-28] The mechanism by which cohesin regulates gene expression is not well-characterized, but evidence suggests that cohesin facilitates long-distance DNA looping over several kilobases.[29-32] This notion is further supported by the finding that cohesin co-localizes with CCCTC-binding factor (CTCF) binding sites. CTCF is a well-known zinc-finger protein involved in gene regulation,[33-35] and in particular, is required for transcriptional insulation of promoters from distant enhancers.[36] Cohesin down-regulation reduces long-range interactions between CTCF sites in several genes; however, these reduced interactions are accompanied by modest changes in gene expression.[29-31, 37] These results are consistent with gene expression data obtained in CdLS cell lines and in *Nipbl*$^{+/-}$ mutant mice where cohesin mutation leads to modest but significant transcriptional dysregulation and overall down-regulation of many genes.[26, 27] These studies revealed that gene transcription is abnormal in CdLS; however, the biochemical pathways that are perturbed in CdLS are as yet unknown. To address this, we used seven CdLS cell lines derived from probands carrying mutations in core cohesin genes, *SMC1A* and *SMC3*, and we applied a proteomic approach that allowed us to identify, for the first time, specific biochemical pathways that are dysregulated in CdLS.

EXPERIMENTAL SECTION

Cell culture

Lymphoblastoid cell lines were grown in RPMI 1640 medium with 20mM sodium-pyruvate, 10% fetal calf serum, penicillin (100 U/ml), streptomycin (0.1 mg/ml) and 1% L-glutamine. We analyzed seven CdLS lymphoblastoid cell lines derived from five unrelated and two related (PT2 and PT3, mother and son respectively) probands carrying mutations in either the *SMC1A* or *SMC3* gene. All of the cases have been previously reported by our laboratories.[7] Studies were performed after informed consent was obtained from the families, according to the procedures established by the local Ethical Committees. To identify differentially expressed protein between CdLS probands and controls, age-, gender- and ethnicity-matched samples from seven normal controls were used. Samples are listed in Supporting Information Table S1 with detailed description.

SMC mutations mapping

Mutations were mapped onto protein models derived from crystal structure data of the yeast SMC1 head domain dimer[38] and the Thermotoga SMC hinge domain dimer[39] using the Cn3D program.[40]

Protein extraction and fluorescence dye labelling

Samples were denatured with a solution of SDS 2% and 25 mM TRIS and precipitated by the Clean-Up kit (GE Healthcare, Uppsala, Sweden). The whole lysate was resuspended with lysis buffer 7M Urea, 2M Thiourea, 4% (w/v) CHAPS, 25 mM Tris, buffered to pH 8, and protein concentration was estimated by the Bradford's assay. Minimal protein labelling for quantitative 2D Fluorescence Difference Gel Electrophoresis (2D-DIGE) was performed by fluorescent minimal cyanine dyes Cy2, Cy3 and Cy5, according to the manufacturer's instructions (GE Healthcare). An internal standard (60 μg), resulting from pooling equal aliquots of all experimental samples, was labelled with the fluorescent dye Cy2. Aliquots of 60 μg of individual samples, from control and disease probands, were randomly labelled with Cy3 or Cy5 fluorophores, to avoid any dye-specific staining bias.

2D-DIGE electrophoresis

2D-DIGE gels [41, 42] were run with seven CdLS samples and seven control samples that were combined according to age and type of mutation, thus resulting in a total of 14 protein spot maps. The Cy3- and Cy5-labelled samples were mixed together with an aliquot of Cy2-labelled internal standard and an equal volume of DIGE buffer (7M Urea, 2M Thiourea, 4% (w/v) CHAPS, and 2% (w/v) DTE). The volume of combined samples was increased to 450 μl with DIGE rehydration solution (7M Urea, 2M Thiourea, 4% (w/v) CHAPS, and 1% (w/v) DTE) and loaded onto commercial immobilized non-linear pH 3-10 gradient (24-cm length IPG strips, Biorad Laboratories, CA, USA). 2D electrophoresis was performed mainly according to Görg.[43] The isoelectric focusing run was performed using an Ettan IPGphor system (GE Healthcare) at 16 °C, until a total of 80,000 Vh was reached. The second dimension (SDS-PAGE) was run at 15 °C on 9-16% polyacrylamide gradient gels (24 cm × 20 cm × 1 mm) using the Ettan DALTtwelve Unit (GE Healthcare). A 17 W/gel constant power was applied until the Bromophenol blue dye front reached the lower end of the gels. All electrophoresis procedures were performed in the dark.

Image acquisition and spot quantification

The DIGE gel images were acquired by the fluorescence Typhoon 9400 imager (GE Healthcare), to visualize differentially labelled protein spots. All gels were scanned at 100 μm resolution and a total of 21 gel images were obtained by control and disease samples.

Spot map images were merged and analyzed using the DeCyder 2D version7.0 software (GE Healthcare), which enabled differences in the abundance levels of proteins to be found. The quantification of spot intensity was performed by the differential in-gel analysis (DIA) module of the software and Cy3:Cy2, Cy5:Cy2 normalized volume ratios for each protein spot were calculated by using the individual signal of pooled-sample Cy2-labelled as internal standard. The subsequent Biological Variation Analysis (BVA) module allowed samples to be inter-compared along with the experimental design by the univariate analysis across the seven gels. Two thousand one hundred protein spots were constantly detected and matched in at least 85% of spot maps produced. Protein spots that showed a normalized spot volume with an increase or decrease in average ratio ≥ 1.3 and ≤ -1.3 respectively, were accepted as being differentially expressed between the compared extracts.

Statistical analysis of protein expression

Protein expression data were filtered by the one-way analysis of variance (ANOVA), which was performed at a significance level of 0.05, and only protein spots that were found in at least 18 out of 21 experimental spot maps were further investigated. In order to further establish the power of the proteomic analysis performed, the p-value distribution test [44] was used and the results obtained are shown in Supporting Information Figure S2. Pairwise comparisons of control and disease experimental groups by Tukey's multiple comparison procedure were also calculated. In addition to univariate analysis, the extended data analysis (EDA) module of DeCyder software was carried out by the intra-gel and inter-gel statistical analyses, allowing the characterization and the classification of biological samples. The principal component analysis (PCA) was selected as multivariate statistical analysis. [45-49] The relationships between the spot maps were visualized by performing a PCA according to the intensity values of protein spots differentially expressed as detected by DIGE analysis. All samples were distributed in a two-dimensional space, along with the first two principal components, PCA1 and PCA2, which were the largest sources of variation in the experimental data set.

Spot identification by mass spectrometry (MS)

Eight hundred micrograms of sample was loaded onto each MS-preparative 2D gel which were then stained with SYPRO Ruby (BioRad) in accordance with the manufacturer's guidelines. After image acquisition by a Typhoon 9400 laser densitometer (GE Healthcare) and image analysis, spots of interest were automatically cut from the preparative gels by an Ettan Spot Picker (GE Healthcare). Spots of interest were identified by peptide mass fingerprinting using an Ettan MALDI-TOF Pro (GE Healthcare), as previously described.[50] Protein spectra were acquired by the Ettan MALDI evaluation software (GE Healthcare), using as internal standard the 842.509 and 2211.105 m/z peptides arising from trypsin autoproteolysis. Peptide mass fingerprint searching was performed on UniProtKB database (Swiss_Prot 2011_03: 525997 sequences; 20234 *Homo sapiens* sequences; 185874894 residues) by Mascot online available software (Matrix Science, UK, http://www.matrixscience.com). The experimental and theoretical peptide-fingerprinting patterns were set to a Δmass less than 100 ppm, number of accepted missed cleavage sites equal to one, and alkylation of cysteine by carbamidomethylation was assumed as fixed modification, while oxidation of methionine was considered as a possible one. The database search was taxonomically limited to *Homo sapiens*. The criteria used to accept identifications included the extent of sequence coverage, number of matched peptides, and probabilistic score ($p < 0.05$) as listed in Table S2.

Tryptic digests whose PMF searching did not unambiguously satisfy such criteria were further investigated, performing peptide sequencing by tandem mass spectrometry. ESI-ion trap MS/MS was achieved on a LCQ DECA IT mass spectrometer (Thermo Finnigan, San

Jose, CA, USA). After sample acidification with 2 μl of 1% (v/v) TFA solution, tryptic peptide solutions were enriched using the Zip-Tip™ pipette tips (Millipore, Billerica, MA, USA), that were previously equilibrated in 50% (v/v) ACN solution and abundantly washed in 0.1% (v/v) TFA. Peptide elution from the Zip-Tip™ matrix was achieved using a 70% (v/v) methanol and 0.5% (v/v) formic acid solution. Three microliters of concentrated sample was then injected into the spectrometer by nanospray according to the following parameters: Aux gas flow rate: 44 (arb); capillary and spray voltage: 3 V and 0.7 kV, respectively; capillary temperature: 185 °C; and lens voltage: -30 V. A Δmass of 0.5 Da was allowed in selecting target peptides. The helium gas collision energy was properly set in relation to the mass of doubly charged precursor ions (from 25% to 35%), and spectra were acquired by the Excalibur software (Thermo). MS/MS database searching was carried out using TurboSEQUEST algoritm (Thermo) [51] and Mascot MS/MS ion search software (http://www.matrixscience.com) in the UniProtKB database. The following criteria were applied: MS mass accuracy ± 1.2 Da, MS/MS mass accuracy ± 0.6 Da, monoisotopic experimental mass values, trypsin digestion with one allowed missed cleavage, fixed carbamidomethylation of cysteine, and variable oxidation of methionine. Mascot ion score, peptide coverage by “b” and “y” ions, and expected value were used for acceptance of peptide assignment and protein identification, as shown in MS Supplementary Material and listed in Table S2. Reported pI and mass (Da) values were experimentally determined by co-migration with human serum as internal standard.[52]

Network analysis

The whole set of proteins detected as differentially expressed by DIGE analysis and identified by mass spectrometry were submitted to network analysis using the GeneGo’s MetaCore software version 6.9 (http://www.genego.com). MetaCore is an integrated knowledge base and pathway analysis tool based on a proprietary and frequently updated database of human protein-protein, protein-DNA and protein compound interactions, and metabolic and signaling pathways. The MetaCore database is manually annotated and all information is extrapolated by curators from scientific literature. This tool has been successfully applied in many important biological studies.[48, 49, 53-56]

In order to establish a global functional pathway among the experimental data, the *shortest path algorithm* set to *high trust* interactions was selected; a short connectivity map was built among the proteins (*nodes*) of interest according to a precise score method of prioritized network database. The scoring system of the constructed biological and functional networks was considered to be statistically significant with a p (probability of random intersection) value < 0.0001.

Measurement of oxidative stress

Protein carbonyl content was assayed with enzyme-linked immunosorbent assay (ELISA) after derivatization with 2,4-dinitrophenylhydrazine using Protein Carbonyl ELISA Kit (Cell Biolabs) and following the manufacturer’s instructions.

Immunoprecipitation

A volume containing 800 μg of total protein extracts from control and CdLS cell lines was dissolved in 1 ml of incubation buffer. The solution were pre-cleared with 20 μl Dynabeads protein A (Invitrogen) for 1 h. The supernatants were then incubated with 3 μg of SMC1A or SMC3, or RAD21 antibody coupled to the 40 μl Dynabeads protein A. The loaded suspensions were precipitated, washed four times with incubation buffer and then re-suspended in SDS loading buffer. Precipitates were analyzed by SDS-PAGE and western blotting using anti-SMC1A (Bethyl Laboratories), anti-SMC3 (Bethyl Laboratories) and anti-RAD21 (Bethyl Laboratories).

SDS-PAGE and immunoblotting

Protein whole extracts from disease and control probands were solubilized in Laemmli buffer[57, 58] and proteins, 20 μg per lane, were separated by SDS-PAGE gels (10%, 30:0.8 Acrylamide/PDA). Pooled samples were obtained by collecting 2.5 μg aliquots from all the control or disease samples to attain two representative samples. Gels were blotted onto nitrocellulose membranes (GE Healthcare) which were subsequently saturated using a phosphate-buffered solution (PBS) 0.1% Triton X100 and 3% (w/v) milk powder, and then probed with specific antibodies by overnight incubation at 4°C. The following primary antibodies were used: c-Myc transcription factor (9E10 Santa Cruz Biotechnology, Santa Cruz, CA), diluted 1:1000; p-53 (Ab6 Calbiochem, UK), diluted 1:1000; Estrogen Receptor alpha (G-20 Santa Cruz Biotechnology), diluted 1:500, Stathmin (OP18, FL-149, Santa Cruz Biotechnology), diluted to 1:1000; 5'-Nucleotidase (T-25, Santa Cruz Biotechnology), diluted 1:1000; Peroxiredoxin III (4G10, Santa Cruz Biotechnology), diluted 1:5000. To assess equal protein loads, tested membranes were immunostained with anti- β-Tubulin antibody (H-235, Santa Cruz Biotechnology), diluted 1:1000. Primary antibodies were then detected after a 2-h incubation period with suitable peroxidise-conjugated secondary antibodies (anti-mouse (Biorad), diluted1:3000 or anti-rabbit (Sigma), diluted 1:7000). Immunostained bands were detected by chemiluminescence using the ECL-PLUS western blot analysis kit (GE Healthcare), according to the manufacturer's instructions.

The online available software *ImageJ* (NIH, Bethesda, MD, USA) was used to carry out the semiquantitative image analysis of immunoblotting data, expressed by percent of disease/control ratio.

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed as previously described, with minor modifications.[59] Briefly, around 10^7cells were cross-linked with 1% formaldehyde for 10 min in ice and then incubated with glycine at 125 mM. Pellets obtained by centrifugation were resuspended in lysis buffer and sonicated to obtain a sheared chromatin of lengths of between 500 and 200 base pairs (bp). Subsequently samples were immunoprecipitated with an antibody previously coupled to Dynabeads protein G (Invitrogen) (overnight at 4 °C). Antibody against RAD21 (Bethyl Laboratories) was used. The samples were then washed several times with RIPA Buffer and eluted overnight at 65 °C to reverse the crosslinking reaction. The elutes were incubated with proteinase K and DNA was extracted with phenol-chloroform, followed by purification with QIAquick Purification Kit (Qiagen). DNA concentration was determined using a Nanodrop spectrophotometer (NanoDrop Technologies). For each cell line, three independent ChIP assays were carried out. The ChIP products obtained were ready to use for quantitative Real Time PCR experiments.

Quantitative Real Time PCR analysis (qPCR)

ChIP products were analyzed by qPCR using QuantiTecT SYBR Green PCR mix (Qiagen) on the Rotor Gene 3000 (Corbett). Each sample was run in duplicates and repeated at least three times. Specific primers (listed in Supporting Information Table S3), were used to assay the relative enrichment for either the promoter or the first exon of *c-MYC* gene (see Fig. 7A). Each sample precipitated without antibodies was used as a mock. The results are expressed as a fold enrichment relative to control cells.

Statistical analysis

Results were analyzed by Student's *t*-test. *p*-values of < 0.05 were considered statistically significant.

RESULTS

Proteomic profile of *SMC1A*- and *SMC3*-mutated CdLS cell lines identifies differentially expressed proteins

To define the biochemical pathways that are dysregulated in CdLS, we used a quantitative 2D Fluorescence Difference Gel Electrophoresis (2D-DIGE) analysis followed by Mass Spectrometry (MS). We analyzed CdLS probands carrying mutations in the core cohesin genes, *SMC1A* and *SMC3*, which cause a more homogeneous milder phenotype, compared to the broader phenotype spectrum seen in *NIPBL* mutation positive probands, which range from mild to severe. In an effort to minimize the variability of data, we utilized seven lymphoblastoid cell lines derived from CdLS probands harbouring mutations in *SMC1A* and *SMC3* as well as seven age- and gender-matched controls (Supporting Information Table S1). *SMC1A* and *SMC3* mutations were categorized in the "head" and "hinge" groups, according to their location in the protein domain. Only a single proband carrying a mutation in the *SMC3* gene has been identified to date. The identification of additional *SMC3* mutations will allow us to investigate their similarity to *SMC1A* mutations. The SMC3 E488del, SMC1A R496H and R496C mutations map to the N-terminal coiled-coil/hinge junction, whereas V58_R62del and F1122L mutations map to the head domain near the Walker A motif and the signature motif of the ABC family of ATPases, respectively (Supporting Information Figure S1A-C). Total protein cell extracts obtained from both CdLS and control samples were used to run 2D-DIGE, permitting us to obtain high-resolution protein maps of CdLS and control cells (Fig. 1). After the detection of the fluorescent staining by a three-laser densitometer (Thyphoon), dedicated DeCyder software was used to perform an accurate computer analysis. This analysis revealed a significant variation of 71 protein spots (Table 1, Supporting Information Fig. S3) according to an average ratio fold change of ± 1.3 and one-way ANOVA p-value ≤ 0.05 (see Experimental section). These differentially expressed spots were cut from the gels and analyzed by MS, allowing us to unambiguously identify 46 proteins (Table 1). The presence of different spots corresponding to the same protein is related to the occurrence of isoforms, mainly due to post-translational modifications, as suggested by the differences between theoretical and experimental MW and pI. In addition to MS-identified spots, Supporting Information Table S2 reports the SwissProt ID and Mascot's search parameters results, such as the number of peptide masses that identified the single protein, the percent of sequences corresponding to the protein tag query, the highest scoring of peptide matches with a specified protein sequence, the peptide sequence eventually obtained by MS/MS analysis and the fold change of protein expression between CdLS probands and controls. All identified proteins were grouped by their prevalent biological function as described in the UniProtKB database: metabolism, cytoskeleton organization, antioxidant and detoxification, cellular and protein fate and RNA processing. Most of the identified proteins were down-regulated while 12 proteins were up-regulated (Table 1). Sixteen out of forty-six (34.8%) proteins belong to the metabolism pathway including alpha-enolase, fructose-bisphosphate aldolase A and C, and hypoxanthine-guanine phosphoribosyltransferase. Seven (15.3%) identified proteins play roles in cytoskeleton organization, including actin-related protein 3, stathmin and cofilin-1. Fourteen out of forty-six (30.4%) proteins are in the cellular and protein fate pathways including annexin A6, endoplasmin, heat shock 71 KDa. Three proteins (6.5%), namely peroxiredoxin III, Parkinson disease protein 7 and superoxide dismutase [Cu-Zn], belong to the antioxidant and detoxification pathway. Finally, six spots (13%) identified proteins involved in the RNA processing pathway including elongation factor 2, Heterogeneous nuclear ribonucleoprotein H, Cleavage and polyadenylation specificity factor subunit 5. 2D data were validated by Western blot analysis using pooled samples confirming that 5'(3')-deoxyribonucleotidase (Supporting Information, Fig. S4A), stathmin (Supporting

Information Fig. S4B) and peroxiredoxin 3 (Supporting Information Fig. S4C) proteins were down-regulated in CdLS when compared to control cell lines.

To investigate whether protein expression dysregulation depends on altered incorporation of mutated SMC proteins in cohesin complex, we performed co-IP with SMC1A, SMC3 and RAD21 proteins using CdLS protein extracts. RAD21 was detected in IP-SMC1A (and -SMC3) precipitates (Fig. 2A) as well as SMC1A (and SMC3) co-precipitated with IP-RAD21 (Fig. 2C). No RAD21 (Fig. 2B) or SMC1A (and SMC3, Fig. 2D) signal was detected in control western blotting using IgG-coated beads. Though small alterations would not be detectable by co-IP and subsequent Western blotting, this result allows us to exclude that SMC mutations abolish the incorporation into the cohesin complex.

The proteomic profile discriminates between CdLS probands carrying mutations in the different domains of SMC protein

A detailed computer-assisted analysis was performed with the CdLS probands being subdivided into two groups, hinge and head (see Supporting Information Fig. S1). Hinge and transition domains are necessary to form a cohesin structure with DNA-binding capacity, and codon R496 in the *SMC1A* gene has been found to be mutated in multiple CdLS probands.[60] As shown in Table 1, most of the identified proteins were within the first inclusive comparison group all controls vs all CdLS cell lines), while six unique protein spots were found to be dysregulated when comparing the head group to the control group. To investigate whether hinge mutations differently affect protein expression, a further computer analysis was performed taking into account only the probands with the R496C and R496H mutations. Twelve unique protein spots were reported to be mildly dysregulated in R496H or/and R496C cell lines with respect to controls. It is worth noting that proteins involved in the response to oxidative stress were specifically down-regulated in hinge mutated probands. To study global oxidative stress, protein carbonyls were measured by ELISA in control and CdLS cell lines. All CdLS cell lines showed a significantly higher protein carbonyl content than control cells, with the exception of CdL074, which was close to significance ($p = 0.07$) (Fig. 3).

Unsupervised sample clustering by principal component analysis (PCA) interestingly revealed three distinct groups: control subjects, hinge mutated CdLS probands, and CdLS probands mutated in the head domain (Fig. 4). The spot maps of the head group (V58_R62del and F1122L) fit very close to each other at the bottom dial of the ellipse, as well as the hinge group distributed in the upper quadrant of the PCA plot. Globally, all CdLS samples appeared separated from the control spot maps, which in turn were strictly grouped and very close to the central axis of the ellipse, thus demonstrating a low statistical variance. PCA clearly separated control subjects from CdLS probands and differentiated probands carrying mutations in the hinge domain from those with mutations in other domains of the SMC proteins.

Our data indicate that dysregulated proteins are involved in important biochemical pathways, in particular the response to oxidative stress, and the proteomic profile can discriminate CdLS carrying mutations in the hinge-transition domain from those with mutations in other domains.

c-MYC transcription factor is dysregulated in CdLS

To find a comprehensive relationship between the identified proteins, the MetaCore network analysis tool was applied. The whole set of the identified proteins was imported into the MetaCore loading module and successfully mapped by the MetaCore database, using the “shortest paths algorithm”. The protein interactions were filtered, so that unconnected root

nodes and indirect interactions were discarded. This analysis predicts that the dysregulated proteins are related to three very important factors; tumor suppressor p53, estrogen receptor (ESR-1) and c-MYC proto-oncogene, with c-MYC being the most convergent hub among the identified proteins (Fig. 5). Moreover, c-MYC regulates many genes including alpha-enolase, purine nucleoside phosphorylase, phosphoglycerate mutase 1, triosephosphate isomerase, nucleoside diphosphate kinase A, adenine phosphoribosyltransferase and peroxiredoxin III, all belonging to the biochemical pathways we found to be dysregulated in CdLS cell lines. Since it is very difficult to detect transcription factors by 2D, it is not surprising that c-MYC was not represented in our identified proteins. In order to evaluate the effectiveness of the network analysis predictions we performed western blotting with a specific antibody against human c-MYC. Results showed that a general decrease in c-MYC expression occurred in CdLS probands compared to control subjects (Fig. 6, Supporting Information Fig. S5). The down-regulation of *c-MYC* was also confirmed by qPCR (data not shown). To investigate whether this effect is specific for c-Myc or whether other transcription factors are affected, we analyzed p53 and ESR1. As shown in Fig. 6, no significant difference was found in their band intensity among CdLS probands and control samples. This finding suggests that c-MYC dysregulation is a general and specific phenomenon in CdLS and further supports the notion that c-MYC is a very important hub within the identified pathways. *c-MYC* plays a pivotal role in growth, proliferation and protein synthesis. Recently, it has been shown that *c-MYC* expression is strongly down-regulated by RAD21 depletion while CTCF depletion had no effect, suggesting that *c-MYC* expression is independent from CTCF but dependent on cohesin.[24] The finding that cohesin regulates c-MYC expression raised the question of whether a similar regulation exists in human CdLS cell lines. To gain insight into *c-MYC* regulation in CdLS, we used ChIP coupled with qPCR to examine RAD21 binding in *SMC1A*- and *SMC3*-mutated CdLS cell lines. The *c-MYC* expression is regulated by both a constitutive CTCF binding site called MYC insulator element (MINE) and exon 1 (Fig. 7A). In fact, the latter contains regulatory elements as well as a region that controls the elongation of nascent RNA transcripts,[61, 62] an enhancer element[63] which is important in the control of *c-MYC* expression. We found that RAD21 bound the promoter region of *c-MYC* locus in all CdLS cell lines, with the exception of CdL057 which showed a significant decrease ($p < 0.05$) (Fig. 7B). Furthermore, the scan for cohesin binding at the first exon showed a significant decrease ($p < 0.05$) in CdLVH, CdLSS, CdL057, CdL060, and CdL107. CdL203, which shares the same mutation with CdL060, showed a decrease in RAD21 binding very close to significance ($p = 0.07$) (Fig. 7B). These results suggest that *c-MYC* may be a key node in the cellular biochemical network and its down-regulation affects several biochemical pathways in CdLS.

DISCUSSION

CdLS is the most common disorder of a class of multi-organ system developmental syndromes termed the cohesinopathies. CdLS is genetically heterogeneous and also has significant phenotypic variability with some genotype-phenotype correlation both within and between causative genes. About 60% of individuals with CdLS have a mutation in *NIPBL* with the majority being haploinsufficient, whereas about 5% are due to mutations in *SMC1A* and *SMC3*.[9, 64] The phenotypic spectrum seen in *NIPBL*-mutated probands ranges from mild to severe while *SMC* mutations cause a more homogeneous milder phenotype. In fact, probands show very mild facial dysmorphia, often normal birth weight and head circumference, mild to moderate intellectual disability, and absence of major structural limb defects. The etiopathologic mechanisms leading to CdLS are as yet poorly understood. To address this, we used proteomic analysis of lymphoblastoid cell lines deriving from *SMC1A*- and *SMC3*- mutated CdLS probands. We reasoned that the analysis of probands with a more homogeneous phenotype would allow us to minimize data variability by highlighting specific biochemical pathways altered in CdLS. We identified 46 proteins that

are dysregulated in CdLS cell lines. It is worth noting that although it has been shown that there is no linear relationship between the abundance of mRNA and that of the corresponding protein, [65] seven of these proteins, namely SCOT1, SAHH, PPIB, CH60, HNRH1, ROA2, in addition to c-MYC (see below), correspond to genes previously found to be dysregulated in the CdLS cell line.[26] The identified dysregulated pathways are suggestive of critical roles that can be correlated to the CdLS phenotype. Identified dysregulated proteins are involved in specific cellular processes: metabolism, cytoskeleton organization, oxidative stress response, cellular and protein fate and RNA processing. Most of the down-regulated proteins belong to metabolic pathways critical in processes to produce adenosine triphosphate (ATP) or reduce nicotinamide adenine dinucleotide (NADH) or important intermediates that drive most cellular processes. Down-regulation in glycolytic enzymes, such as alpha-enolase, phosphoglycerate kinase 1, aldolase enzyme A or C isoforms, triosephosphate isomerase, could also affect cell functional processes and survival.

As protein carbonyl content is a common marker of protein oxidation, our findings that CdLS cell lines show both a down-regulation of proteins involved in the response to oxidative stress and an increase in global oxidative stress argue that they directly contribute to some of CdLS phenotypic features. Excessive production of reactive oxygen species and/or the reduction of antioxidant defenses can cause DNA damage, loss of membrane potential and reduced ATP synthesis, ultimately leading to cell death. It is worth noting that *SMC*-mutated probands show evidence of a premature aging process, and manifest several physical changes resulting in a more aged appearance compared to their chronological age.[66] In addition, CdLS cell lines show high levels of DNA damage and spontaneous genome instability.[12, 19] It is likely that a reduction in energy and the down-regulation of proteins involved in antioxidant and detoxification pathways could lead to premature physiological aging and genome instability.

Unsupervised sample clustering by PCA of dysregulated proteins was able to differentiate control subjects from CdLS probands, and among the latter, to separate probands carrying mutations in the hinge domain from those with mutations in other domains of the SMC proteins (Fig. 4), suggesting that these two groups have different expression patterns. This result is likely due to the different roles of hinge with respect to other domains. SMC proteins consist of five distinct domains; N- and C-terminal domains, two coiled-coil domains and the hinge domain. This latter is composed of about 150 amino acids and is extremely flexible, allowing homo- and hetero-dimerization. Mutations in the hinge domain could alter the entire structure of the cohesin core by preventing correct protein association, and as a consequence, gene transcription and subsequent protein expression could be affected. However, since the head group includes two different mutations and the hinge group three different mutations (mapped in SMC1A and SMC3) it is likely that different head domain (or hinge domain) mutations have different consequences. Of note, recently it has been suggested that hinge domain is the entry gate for cohesin loading onto DNA, whereas the head domain is the exit gate.[67]

MetaCore analysis allowed us to clarify the key role of *c-MYC*, which was found to be an important convergent hub. *c-MYC* is involved in many biological processes such as cell growth and differentiation, stem cell fate and development processes during early embryonic stages.[68] By immunoblotting analysis, we found that *c-MYC* is down-regulated in all CdLS cell lines, which may explain the deregulation of proteins that are directly activated by c-MYC in CdLS. Genome-wide data have shown an extensive overlap of cohesin and CTCF in mammalian genomes[69] including the promoter region of *c-MYC*.[25, 34] However, recent studies argue that cohesin regulates *c-MYC* expression independently from CTCF.[24, 70] We found that cohesin binding decreased at the first exon of *c-MYC* gene in six out of seven CdLS cell lines. It is well- known that the first exon contains regulatory elements that are

important in the control of *c-MYC* expression and the elongation of nascent RNA transcripts.[61, 62, 71] Our finding suggests that *c-MYC* expression is positively regulated by cohesin, and changes in the promoter structure lead to its deregulation. Altogether, these data suggests that c-MYC deregulation could be a critical early step in the processes leading to the CdLS phenotype and further suggests that *c-MYC* expression by cohesin is evolutionarily conserved.

CONCLUSION

Our analysis allows the first mapping of networks by grouping molecules into functional categories which are specifically dysregulated in CdLS cell lines. Significantly, the modest protein dysregulation we found is in agreement with previous published transcription data. In fact, human, mouse, *Drosophila* and zebrafish studies indicate that hundreds of genes are transcriptionally dysregulated in cohesin-mutated cells or in cohesin genes knocked down using RNAi. Furthermore, the changes in gene expression are modest, typically less than 1.5-fold.[23, 26-28, 72] Altogether, these data raise the intriguing possibility that CdLS developmental deficits result from the collective action of many modest perturbations. The finding that mutated SMC proteins are normally incorporated into the cohesin complex argues that dysregulation likely depends on dominant activity of mutated cohesin. Furthermore, the identified dysregulated pathways can be clustered around a central hub containing *c-MYC*. We therefore suggest that *c-MYC* down-regulation caused by cohesin mutations in *SMC1A* and *SMC3* genes may be an early/primary event in the pathogenesis of CdLS. Finally, the development and use of biomarkers identified by proteomic analysis can be useful as a diagnostic and classification tool and lead to the design of new treatments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from the Region of Tuscany to AM, FIRB “Italian Human ProteomeNet” (BRN07BMCT_013), from the Italian Ministry of University and Scientific Research to LB, NIH/NICHD PO1HD052860 to IDK and by NIH/NICHD K08HD055488 to MAD.

References

1. Kline AD, Krantz ID, Sommer A, Kliewer M, Jackson LG, FitzPatrick DR, Levin AV, Selicorni A. Cornelia de Lange syndrome: clinical review, diagnostic and scoring systems, and anticipatory guidance. Am J Med Genet A. 2007; 143A:1287–1296. [PubMed: 17508425]
2. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. Annu Rev Genet. 2009; 43:525–558. [PubMed: 19886810]
3. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M, Jackson LG. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nat Genet. 2004; 36:631–635. [PubMed: 15146186]
4. Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet. 2004; 36:636–641. [PubMed: 15146185]
5. Gillis LA, McCallum J, Kaur M, DeScipio C, Yaeger D, Mariani A, Kline AD, Li HH, Devoto M, Jackson LG, Krantz ID. NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. Am J Hum Genet. 2004; 75:610–623. [PubMed: 15318302]

6. Borck G, Zarhrate M, Bonnefont JP, Munnich A, Cormier-Daire V, Colleaux L. Incidence and clinical features of X-linked Cornelia de Lange syndrome due to SMC1L1 mutations. Hum Mutat. 2007; 28:205–206. [PubMed: 17221863]
7. Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodriguez C, Arnedo M, Loeys B, Kline AD, Wilson M, Lillquist K, Siu V, Ramos FJ, Musio A, Jackson LS, Dorsett D, Krantz ID. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet. 2007; 80:485–494. [PubMed: 17273969]
8. Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L. X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. Nat Genet. 2006; 38:528–530. [PubMed: 16604071]
9. Mannini L, Liu J, Krantz ID, Musio A. Spectrum and consequences of SMC1A mutations: the unexpected involvement of a core component of cohesin in human disease. Hum Mutat. 2010; 31:5–10. [PubMed: 19842212]
10. Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, Saitoh K, Komata M, Katou Y, Clark D, Cole KE, De Baere E, Decroos C, Di Donato N, Ernst S, Francey LJ, Gyftodimou Y, Hirashima K, Hullings M, Ishikawa Y, Jaulin C, Kaur M, Kiyono T, Lombardi PM, Magnaghi-Jaulin L, Mortier GR, Nozaki N, Petersen MB, Seimiya H, Siu VM, Suzuki Y, Takagaki K, Wilde JJ, Willems PJ, Prigent C, Gillessen-Kaesbach G, Christianson DW, Kaiser FJ, Jackson LG, Hirota T, Krantz ID, Shirahige K. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. Nature. 2012; 489:313–7. [PubMed: 22885700]
11. Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, Monnich M, Yan Y, Xu W, Gil-Rodriguez MC, Clark D, Hakonarson H, Halbach S, Michelis LD, Rampuria A, Rossier E, Spranger S, Van Maldergem L, Lynch SA, Gillessen-Kaesbach G, Ludecke HJ, Ramsay RG, McKay MJ, Krantz ID, Xu H, Horsfield JA, Kaiser FJ. RAD21 mutations cause a human cohesinopathy. Am J Hum Genet. 2012; 90:1014–27. [PubMed: 22633399]
12. Revenkova E, Focarelli ML, Susani L, Paulis M, Bassi MT, Mannini L, Frattini A, Delia D, Krantz I, Vezzoni P, Jessberger R, Musio A. Cornelia de Lange syndrome mutations in SMC1A or SMC3 affect binding to DNA. Hum Mol Genet. 2009; 18:418–427. [PubMed: 18996922]
13. Musio A, Montagna C, Zambroni D, Indino E, Barbieri O, Citti L, Villa A, Ried T, Vezzoni P. Inhibition of BUB1 results in genomic instability and anchorage-independent growth of normal human fibroblasts. Cancer Res. 2003; 63:2855–2863. [PubMed: 12782591]
14. Watrin E, Peters JM. The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells. EMBO J. 2009; 28:2625–2635. [PubMed: 19629043]
15. Musio A, Montagna C, Mariani T, Tilenni M, Focarelli ML, Brait L, Indino E, Benedetti PA, Chessa L, Albertini A, Ried T, Vezzoni P. SMC1 involvement in fragile site expression. Hum Mol Genet. 2005; 14:525–533. [PubMed: 15640246]
16. Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, Velculescu VE, Kinzler KW, Vogelstein B, Lengauer C, Hieter P. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. Proc Natl Acad Sci U S A. 2008; 105:3443–3448. [PubMed: 18299561]
17. Mannini L, Menga S, Musio A. The expanding universe of cohesin functions: a new genome stability caretaker involved in human disease and cancer. Hum Mutat. 2010; 31:623–630. [PubMed: 20513141]
18. Mannini L, Musio A. The dark side of cohesin: The carcinogenic point of view. Mutat Res. 2011; 728:81–87. [PubMed: 22106471]
19. Vrouwe MG, Elghalbzouri-Maghrani E, Meijers M, Schouten P, Godthelp BC, Bhuiyan ZA, Redeker EJ, Mannens MM, Mullenders LH, Pastink A, Darroudi F. Increased DNA damage sensitivity of Cornelia de Lange syndrome cells: evidence for impaired recombinational repair. Hum Mol Genet. 2007; 16:1478–1487. [PubMed: 17468178]
20. Rollins RA, Morcillo P, Dorsett D. Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. Genetics. 1999; 152:577–593. [PubMed: 10353901]

21. Schuldiner O, Berdnik D, Levy JM, Wu JS, Luginbuhl D, Gontang AC, Luo L. piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. Dev Cell. 2008; 14:227–38. [PubMed: 18267091]
22. Pauli A, van Bemmel JG, Oliveira RA, Itoh T, Shirahige K, van Steensel B, Nasmyth K. A direct role for cohesin in gene regulation and ecdysone response in Drosophila salivary glands. Curr Biol. 2010; 20:1787–1798. [PubMed: 20933422]
23. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA. Mediator and cohesin connect gene expression and chromatin architecture. Nature. 2010; 467:430–5. [PubMed: 20720539]
24. Rhodes JM, Bentley FK, Print CG, Dorsett D, Misulovin Z, Dickinson EJ, Crosier KE, Crosier PS, Horsfield JA. Positive regulation of c-Myc by cohesin is direct, and evolutionarily conserved. Dev Biol. 2010; 344:637–649. [PubMed: 20553708]
25. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. Embo J. 2008; 27:654–666. [PubMed: 18219272]
26. Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, Spinner NB, Vega H, Jackson LG, Shirahige K, Krantz ID. Transcriptional dysregulation in NIPBL and cohesin mutant human cells. PLoS Biol. 2009; 7:e1000119. [PubMed: 19468298]
27. Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, Nussenzweig A, Kitzes L, Yokomori K, Hallgrimsson B, Lander AD. Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome. PLoS Genet. 2009; 5:e1000650. [PubMed: 19763162]
28. Schaaf CA, Misulovin Z, Sahota G, Siddiqui AM, Schwartz YB, Kahn TG, Pirrotta V, Gause M, Dorsett D. Regulation of the Drosophila Enhancer of split and invected-engrailed gene complexes by sister chromatid cohesion proteins. PLoS One. 2009; 4:e6202. [PubMed: 19587787]
29. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature. 2009; 460:410–413. [PubMed: 19458616]
30. Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A. Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS Genet. 2009; 5:e1000739. [PubMed: 19956766]
31. Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M. Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. EMBO J. 2009; 28:1234–45. [PubMed: 19322193]
32. Dorsett D. Cohesin: genomic insights into controlling gene transcription and development. Curr Opin Genet Dev. 2011; 21:199–206. [PubMed: 21324671]
33. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M. Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell. 2008; 132:422–433. [PubMed: 18237772]
34. Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A. 2008; 105:8309–8314. [PubMed: 18550811]
35. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM. Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature. 2008; 451:796–801. [PubMed: 18235444]
36. Ohlsson R, Bartkuhn M, Renkawitz R. CTCF shapes chromatin by multiple mechanisms: the impact of 20 years of CTCF research on understanding the workings of chromatin. Chromosoma. 2010; 119:351–360. [PubMed: 20174815]
37. Hou C, Dale R, Dean A. Cell type specificity of chromatin organization mediated by CTCF and cohesin. Proc Natl Acad Sc i U S A. 2010; 107:3651–6.

38. Haering CH, Schoffnegger D, Nishino T, Helmhart W, Nasmyth K, Lowe J. Structure and stability of cohesin's Smc1-kleisin interaction. Mol Cell. 2004; 15:951–964. [PubMed: 15383284]
39. Haering CH, Lowe J, Hochwagen A, Nasmyth K. Molecular architecture of SMC proteins and the yeast cohesin complex. Mol Cell. 2002; 9:773–788. [PubMed: 11983169]
40. Wang Y, Geer LY, Chappey C, Kans JA, Bryant SH. Cn3D: sequence and structure views for Entrez. Trends Biochem Sci. 2000; 25:300–2. [PubMed: 10838572]
41. Rabilloud T, Adessi C, Giraudel A, Lunardi J. Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis. 1997; 18:307–16. [PubMed: 9150907]
42. Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics. 2001; 1:377–96. [PubMed: 11680884]
43. Gorg A, Postel W, Gunther S, Weser J, Strahler JR, Hanash SM, Somerlot L, Kuick R. Approach to stationary two-dimensional pattern: influence of focusing time and immobiline/carrier ampholytes concentrations. Electrophoresis. 1988; 9:37–46. [PubMed: 2466646]
44. Karp NA, McCormick PS, Russell MR, Lilley KS. Experimental and statistical considerations to avoid false conclusions in proteomics studies using differential in-gel electrophoresis. Mol Cell Proteomics. 2007; 6:1354–64. [PubMed: 17513293]
45. Tarroux P. Analysis of protein patterns during differentiation using 2-D electrophoresis and computer multidimensional classification. Electrophoresis. 1983; 4:63–70.
46. Anderson NL, Hofmann JP, Gemmell A, Taylor J. Global approaches to quantitative analysis of gene-expression patterns observed by use of two-dimensional gel electrophoresis. Clinical Chemistrry. 1984; 30:2031–2036.
47. Rabilloud T, Vincens P, Tarroux P. A new tool to study genetic expression using 2-D electrophoresis data: the functional map concept. FEBS letters. 1985; 189:171–178.
48. Bianchi L, Bruzzese F, Leone A, Gagliardi A, Puglia M, Di Gennaro E, Rocco M, Gimigliano A, Pucci B, Armini A, Bini L, Budillon A. Proteomic analysis identifies differentially expressed proteins after HDAC vorinostat and EGFR inhibitor gefitinib treatments in Hep-2 cancer cells. Proteomics. 2011; 11:3725–3742. [PubMed: 21761561]
49. Bianchi L, Gagliardi A, Gioia R, Besio R, Tani C, Landi C, Cipriano M, Gimigliano A, Rossi A, Marini JC, Forlino A, Bini L. Differential response to intracellular stress in the skin from osteogenesis imperfecta Brtl mice with lethal and non lethal phenotype: A proteomic approach. J Proteomics. 201210.1016/j.prot.2012.01.038
50. Bianchi L, Puglia M, Landi C, Matteoni S, Perini D, Armini A, Verani M, Trombetta C, Soldani P, Roncada P, Greppi G, Pallini V, Bini L. Solubilization methods and reference 2-DE map of cow milk fat globules. J Proteomics. 2009; 72:853–64. [PubMed: 19111954]
51. Lundgren DH, Han DK, Eng JK. Protein identification using TurboSEQUEST. Curr Protoc Bioinformatics. 2005; Chapter 13(Unit 13):3. [PubMed: 18428747]
52. Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez JC, Frutiger S, Hochstrasser D. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. Electrophoresis. 1993; 14:1023–1031. [PubMed: 8125050]
53. Marotta LL, Almendro V, Marusyk A, Shipitsin M, Schemme J, Walker SR, Bloushtain-Qimron N, Kim JJ, Choudhury SA, Maruyama R, Wu Z, Gonen M, Mulvey LA, Bessarabova MO, Huh SJ, Silver SJ, Kim SY, Park SY, Lee HE, Anderson KS, Richardson AL, Nikolskaya T, Nikolsky Y, Liu XS, Root DE, Hahn WC, Frank DA, Polyak K. The JAK2/STAT3 signaling pathway is required for growth of CD44CD24 stem cell-like breast cancer cells in human tumors. J Clin Invest. 2011; 121:2723–2735. [PubMed: 21633165]
54. Nikolsky Y, Kirillov E, Zuev R, Rakhmatulin E, Nikolskaya T. Functional analysis of OMICs data and small molecule compounds in an integrated "knowledge-based" platform. Methods Mol Biol. 2009; 563:177–196. [PubMed: 19597786]
55. Nambiar PR, Gupta RR, Misra V. An "Omics" based survey of human colon cancer. Mutat Res. 2010; 693:3–18. [PubMed: 20691711]
56. Vellaichamy A, Dezso Z, JeBailey L, Chinnaiyan AM, Sreekumar A, Nesvizhskii AI, Omenn GS, Bugrim A. "Topological significance" analysis of gene expression and proteomic profiles from

prostate cancer cells reveals key mechanisms of androgen response. PLoS One. 2010; 5:e10936. [PubMed: 20532174]

57. Gallagher SR. One-dimensional SDS gel electrophoresis of proteins. Curr Protoc Cell Biol. 2007; Chapter 6(Unit 6):1. [PubMed: 18228518]
58. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970; 227:680–685. [PubMed: 5432063]
59. Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, McKim K. Effects of sister chromatid cohesion proteins on cut gene expression during wing development in Drosophila. Development. 2005; 132:4743–4753. [PubMed: 16207752]
60. Mannini L, Menga S, Tonelli A, Zanotti S, Bassi MT, Magnani C, Musio A. SMC1A codon 496 mutations affect the cellular response to genotoxic treatments. Am J Med Genet A. 2011; 158A: 224–228.
61. Bentley DL, Groudine M. Novel promoter upstream of the human c-myc gene and regulation of c-myc expression in B-cell lymphomas. Mol Cell Biol. 1986; 6:3481–3489. [PubMed: 3540591]
62. Bentley DL, Groudine M. A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. Nature. 1986; 321:702–706. [PubMed: 3520340]
63. Yang JQ, Remmers EF, Marcu KB. The first exon of the c-myc proto-oncogene contains a novel positive control element. EMBO J. 1986; 5:3553–62. [PubMed: 3030732]
64. Liu J, Krantz ID. Cohesin and human disease. Annu Rev Genomics Hum Genet. 2008; 9:303–320. [PubMed: 18767966]
65. Anderson L, Seilhamer J. A comparison of selected mRNA and protein abundances in human liver. Electrophoresis. 1997; 18:533–7. [PubMed: 9150937]
66. Kline AD, Grados M, Sponseller P, Levy HP, Blagowidow N, Schoedel C, Rampolla J, Clemens DK, Krantz I, Kimball A, Pichard C, Tuchman D. Natural history of aging in Cornelia de Lange syndrome. Am J Med Genet C Semin Med Genet. 2007; 145C:248–60. [PubMed: 17640042]
67. Nasmyth K. Cohesin: a catenase with separate entry and exit gates? Nat Cell Biol. 2011; 13:1170–1177. [PubMed: 21968990]
68. Lin CH, Lin C, Tanaka H, Fero ML, Eisenman RN. Gene regulation and epigenetic remodeling in murine embryonic stem cells by c-Myc. PLoS One. 2009; 4:e7839. [PubMed: 19915707]
69. Wendt KS, Peters JM. How cohesin and CTCF cooperate in regulating gene expression. Chromosome Res. 2009; 17:201–14. [PubMed: 19308701]
70. Gombert WM, Farris SD, Rubio ED, Morey-Rosler KM, Schubach WH, Krumm A. The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain. Mol Cell Biol. 2003; 23:9338–48. [PubMed: 14645543]
71. Hann SR, King MW, Bentley DL, Anderson CW, Eisenman RN. A non-AUG translational initiation in c-myc exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. Cell. 1988; 52:185–95. [PubMed: 3277717]
72. Muto A, Calof AL, Lander AD, Schilling TF. Multifactorial origins of heart and gut defects in nipbl-deficient zebrafish, a model of Cornelia de Lange Syndrome. PLoS Biol. 2011; 9:e1001181. [PubMed: 22039349]

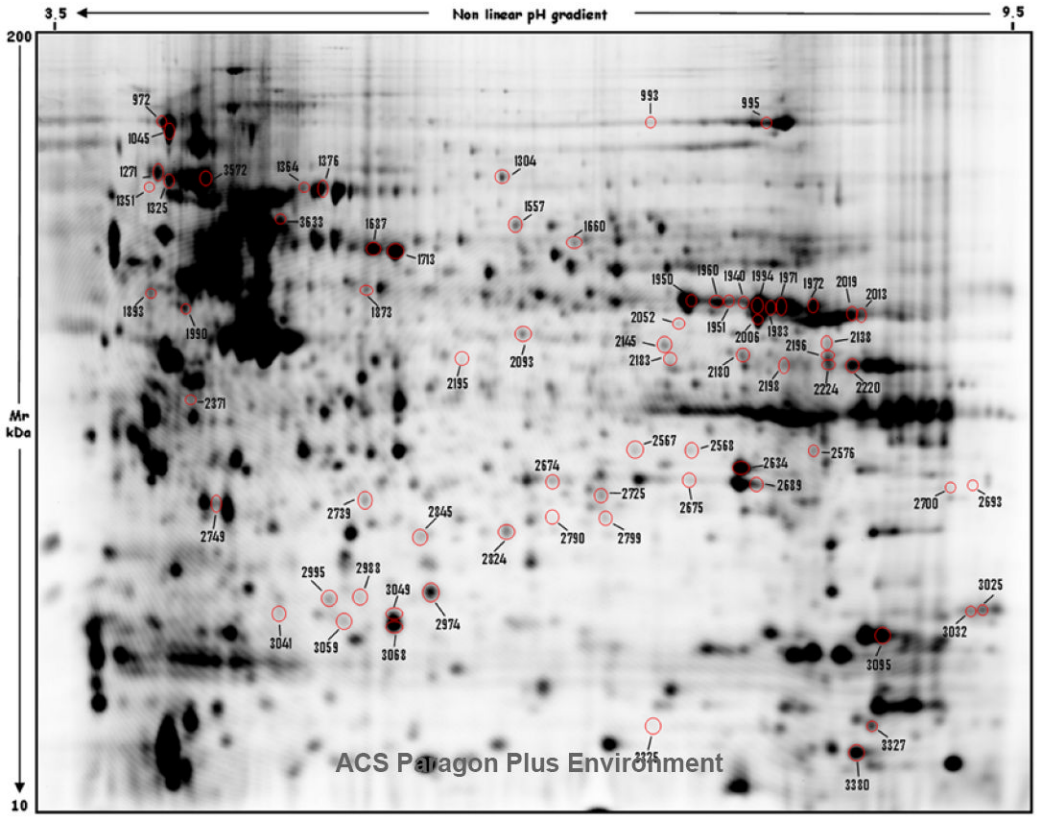

Fig. 1.

Proteomic analysis of CdLS cell lines. Representative DIGE spot map with differently expressed protein spots in CdLS probands compared to control subjects. Seventy-one proteins were found to be dysregulated according to an average ratio fold change of ± 1.3 and $p \leq 0.05$. The identified protein spots were marked by a specific number. Spot numbers match those in Table 1.

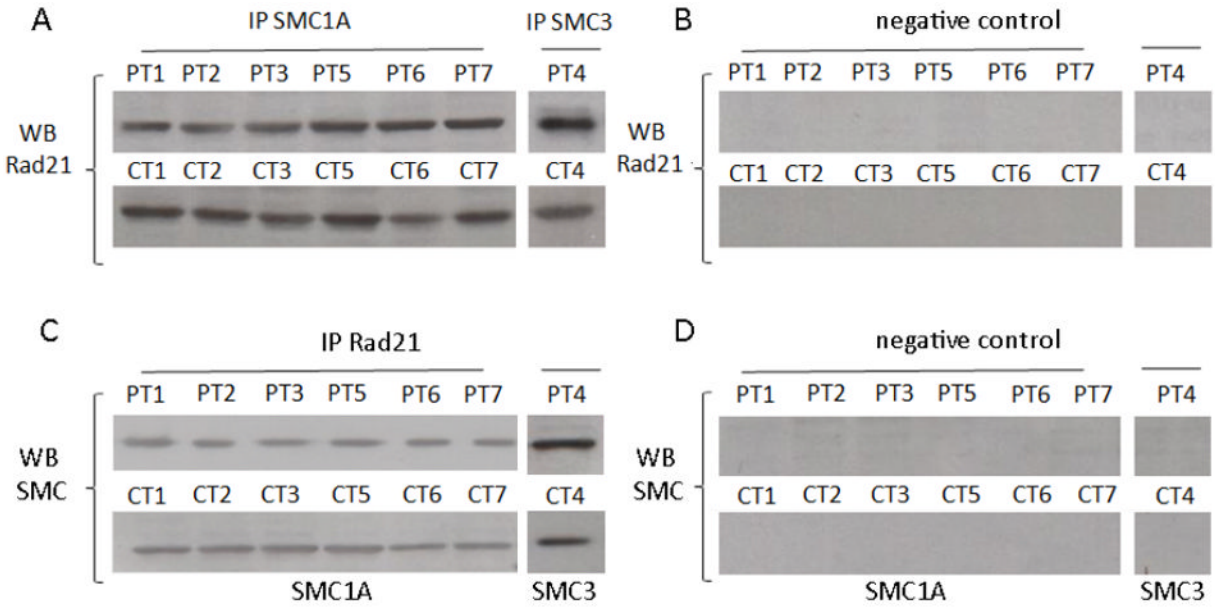

Fig. 2.

Mutated SMC1A and SMC3 co-immunoprecipitate with RAD21 in the CdLS cell lines. (A) SMC1A (and SMC3) was found to be co-precipitated with RAD21, (B) whereas no RAD21 signal was detected in the IPs using IgG-coated beads. (C) In addition, RAD21co-precipitated with SMC1A (and SMC3) and (D) no SMC specific signal was detectable in the IPs using IgG-coated beads.

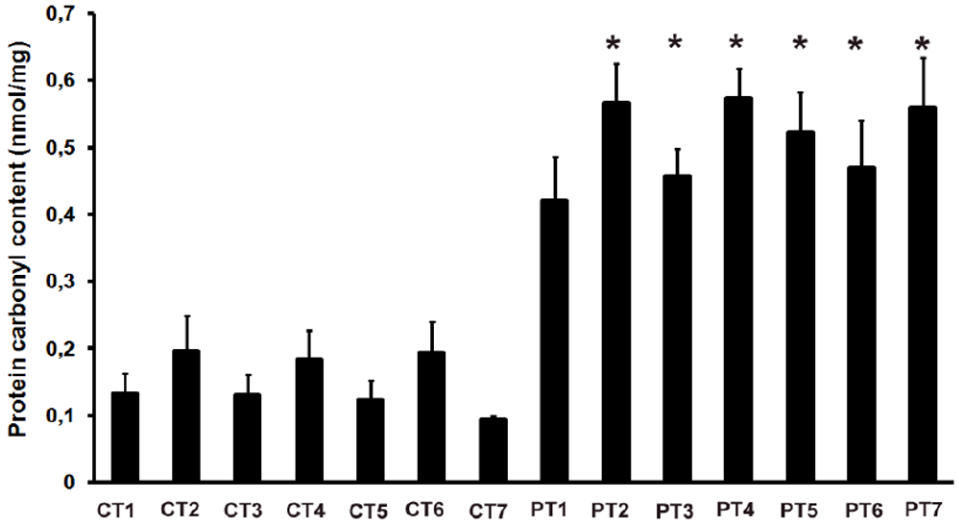

Fig. 3.

Amount of protein carbonyl content in control and CdLS cell lines. All CdLS cell lines showed a significantly higher protein carbonyl content than control cells, with the exception of CdL074 which was close to significance ($p = 0.07$). *$p < 0.05$.

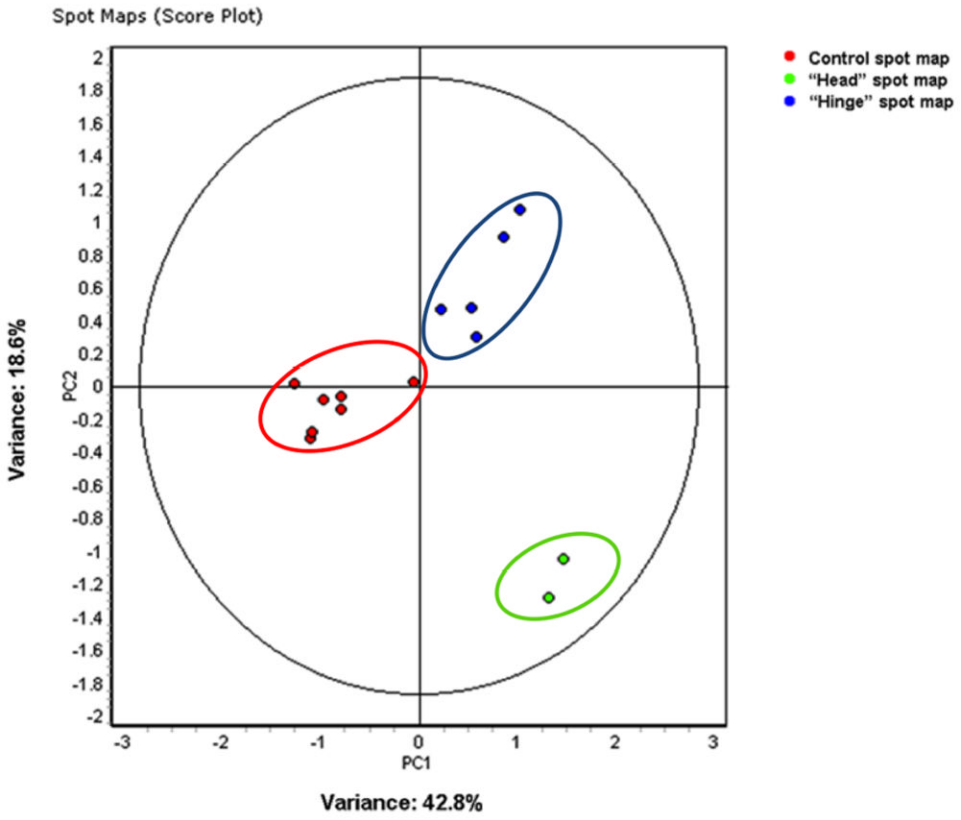

Fig. 4.

Classification of CdLS probands by protein expression. Unsupervised sample clustering by principle component analysis (PCA) of all identified proteins was able to separate the proteins into three different groups: control spot map (red color), head-mutated spot map (green color) and hinge-mutated spot map (blue color).

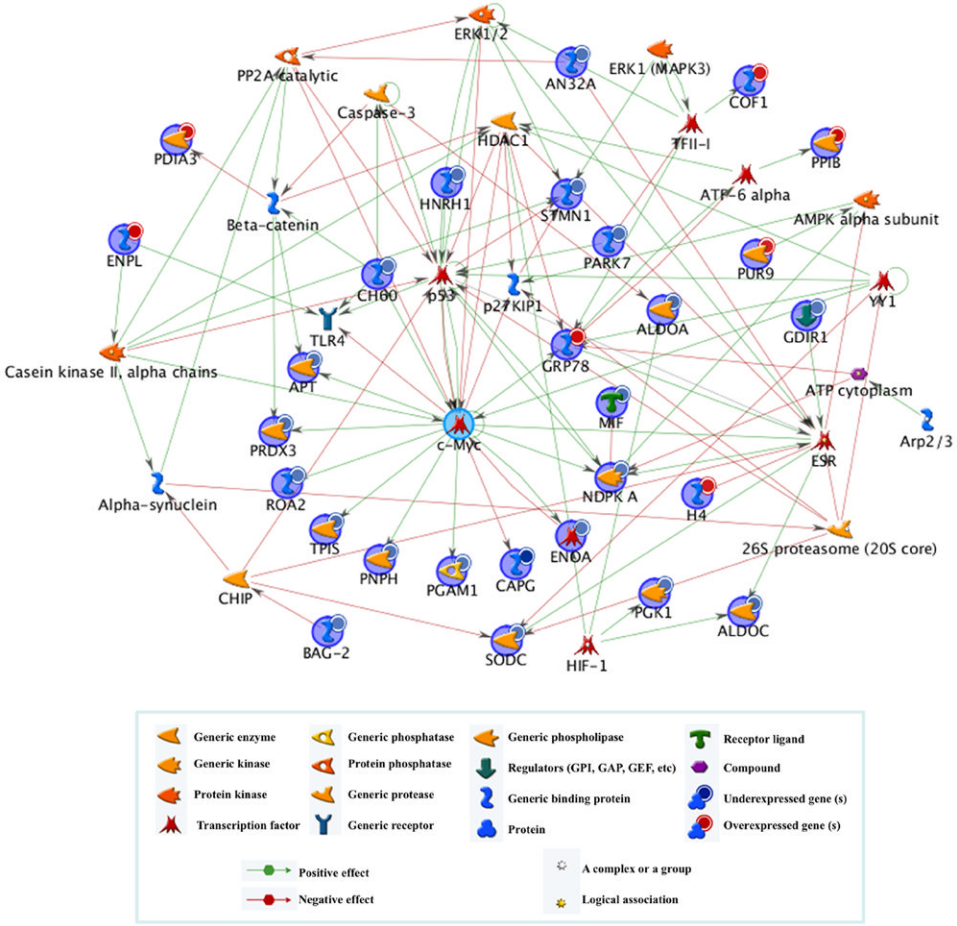

Fig. 5.

c-MYC as the principal “hub” of dysregulated pathways in CdLS cell lines. MetaCore network analysis of 2D results by using the “shortest paths algorithm”. Colored symbols on the edge regard the mechanism, while the effects of interaction between proteins on the network are shown with arrowheads: green for activation and red for inhibition. Dysregulated proteins were related to three transcription factors, namely c-MYC, p53 and estrogen receptor (ESR-1), but c-MYC seems to be the most convergent hub directly or indirectly regulating most of the identified proteins.

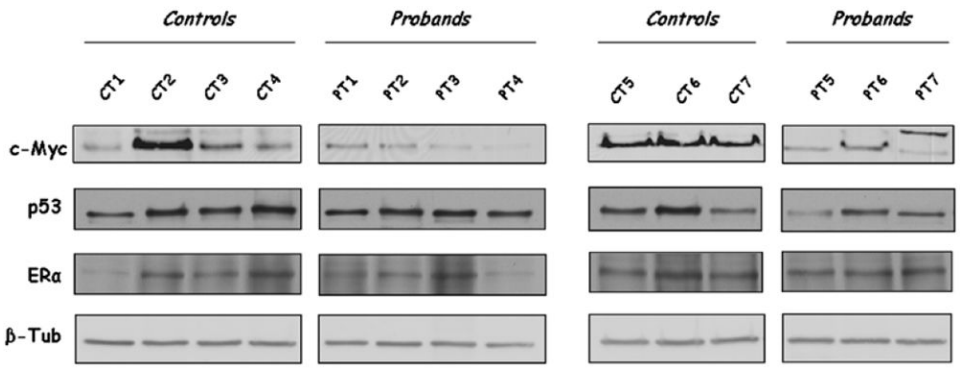

Fig. 6.

c-MYC is down-regulated in CdLS cell lines. Western blotting showed that the transcription factor c-Myc is down-regulated in CdLS cell lines, whereas no difference was found for ESR1 and p53 transcription factors, supporting the notion that c-MYC is an important hub for identified proteins. B-Tubulin (B-TUB) was used as loading control. *$p < 0.05$

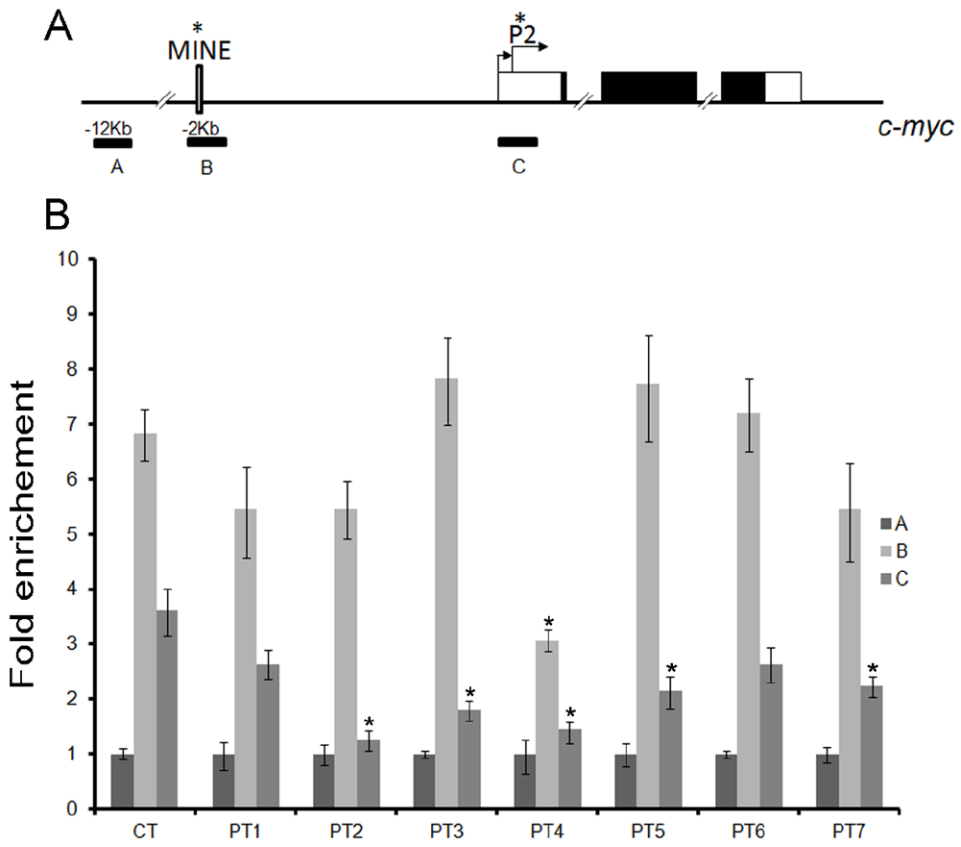

Fig. 7.

Analysis of cohesin binding at the human *c-MYC* locus. (A) Schematic of human *c-MYC* gene. Black solid boxes indicate translated regions. Cohesin binding to MINE and first exon is denoted by an asterisk. Arrows indicate transcriptional start sites of the P1 and P2 promoter. Location of primer pairs A-C are indicated, ChIP assays with antibody against RAD21 were analysed by real-time PCR with primers pairs B and C specific for the promoter region of the *c-MYC* locus. Binding at each site was determined relative to primer A, where no RAD21 binding was predicted. (B) Results showed that RAD21 co-localized to the promoter region of *c-MYC* in all *SMC1A*- and *SMC3*-mutated CdLS cell lines, with the exception of CdL057, whereas bound cohesin was dramatically reduced in exon 1 in CdLVH, CdLSS, CdL057, CdL060, CdL107 CdLS cell lines. CdL203, which shares the same mutation with CdL060, showed a decrease close to significant ($p = 0.07$). Since no difference was found in control cell lines, data was pooled. Results shown are the averages of three independent experiments. The graphs show the average and the standard error of the normalized values and $*p < 0.05$.

Table 1

Differentially expressed protein spots identified by DIGE/Mass Spectrometry analysis.

Master Spot N. *a)*	Spot Entry name *b)*	Description	AC	Fold change *c)*				
				Disease *vs* Control	Head *vs* control	Hinge *vs* control	R496C *vs* control	R496H *vs* control
Metabolism								
1557	PUR9	Bifunctional purine biosynthesis protein	P31939		1.30 *			
1660	SCOT1	Succinyl-CoA:3-ketoacid-coenzyme A transferase 1, mitochondrial	P55809				1.38 *	
1940	ENOA	Alpha-enolase	P06733	-1.37 *		-1.44 **	-1.44 **	
1950	ENOA	Alpha-enolase	P06733				-1.32 *	
1951	ENOA	Alpha-enolase	P06733	-1.30 *		-1.36 *	-1.33 **	
1960	ENOA	Alpha-enolase	P06733	-1.31 *		-1.48 **	-1.30 *	-1.30 *
1971	ENOA	Alpha-enolase	P06733	-1.31 *		-1.37 *	-1.34 *	
1972	ENOA	Alpha-enolase	P06733	-1.33 **	-1.36 **	-1.32 *	-1.44 **	
1983	ENOA	Alpha-enolase	P06733	-1.39 **		-1.44 *	-1.44 *	
1990	ENOG	Gamma-enolase	P09104	-1.52 *		-1.67 *	-1.91 *	
1994	ENOA	Alpha-enolase	P06733	-1.43 **	-1.33 *	-1.48 **	-1.47 **	
2006	ENOA	Alpha-enolase	P06733	-1.39 *	-1.44 *	-1.44 *	-1.42 *	
2013	PGK1	Phosphoglycerate kinase 1	P00558				-1.46 *	
2019	PGK1	Phosphoglycerate kinase 1	P00558	-1.38 *		-1.35 *	-1.55 *	
2093	SAHH	Adenosylhomocysteinase	P23526		1.41 *			
2138	ALDOA	Fructose-bisphosphate aldolase A	P04075		-1.36 **		-1.39 **	
2145	ALDOC	Fructose-bisphosphate aldolase C	P09972	-1.37 **	-1.40 **	-1.35 **	-1.39 **	
2180	ALDOC	Fructose-bisphosphate aldolase C	P09972	-1.55 **	-1.70 *	-1.50 *	-1.83 *	
2183	ALDOC	Fructose-bisphosphate aldolase C	P09972	-1.42 *			-1.70 *	
2196	ALDOA	Fructose-bisphosphate aldolase A	P04075	-1.63 **	-1.61 **	-1.65 **	-1.74 *	
2198	ALDOA	Fructose-bisphosphate aldolase A	P04075	-1.51 *		-1.47 *	-1.78 *	

Master Spot N. *a)*	Spot Entry name *b)*	Description	AC	Fold change *c)*				
				Disease *vs* Control	Head *vs* control	Hinge *vs* control	R496C *vs* control	R496H *vs* control
2220	ALDOA	Fructose-bisphosphate aldolase A	P04075				-1.31 *	
2224	ALDOA	Fructose-bisphosphate aldolase A	P04075	-1.32 **		-1.32 *	-1.48 *	
2567	PNPH	Purine nucleoside phosphorylase	P00491	-1.32 **		-1.33 *	-1.58 **	
2634	PGAM1	Phosphoglycerate mutase 1	P18669				-1.40 *	
2674	TPIS	Triosephosphate isomerase	P60174	-1.43 **	-1.45 *	-1.42 *	-1.54 *	
2675	TPIS	Triosephosphate isomerase	P60174	-1.49 **	-1.65 *	-1.43		
2689	PGAM1	Phosphoglycerate mutase 1	P18669	-1.33 *				
2693	KAD3	GTP:AMP phosphotransferase, mitochondrial	Q9UIJ7	-1.42 *	-1.79 *			
2725	HPRT	Hypoxanthine-guanine phosphoribosyltransferase	P00492				-1.34 *	
2790	NT5C	5’(3’)-deoxyribonucleotidase, cytosolic	Q8TCD5	-1.41 **	-1.98 **		-1.35 *	
2974	NDKA	Nucleoside diphosphate kinase A	P15531					-1.30 *
2995	APT	Adenine phosphoribosyltransferase	P07741	-1.30 **	-1.44 *			
Cytoskeleton organization								
1304	MOES	Moesin	P26038	-1.36 **		-1.44 *	-1.45 **	
1873	ARP3	Actin-related protein 3	P61158	-1.32 **		-1.36 **	-1.32 **	-1.36 **
2195	CAPG	Macrophage-capping protein	P40121	-1.85 *		-2.26 *		-4.28 **
2749	GDIR1	Rho-GDI alpha	P52565	-1.36 **	-1.35 *	-1.30 *	-1.33 *	
3041	ARPC5	Actin-related protein 2/3 complex subunit 5	O15511			-1.39 *	-1.56 *	-1.41 *
3059	STMN1	Stathmin	P16949				-1.58 *	
3068	STMN1	Stathmin	P16949				-1.51 *	
3095	COF1	Cofilin-1	P23528				1.46 **	
Response to oxidative stress								
2824	PRDX3	Peroxiredoxin III	P30048				-1.53 *	
2845	PARK7	Protein DJ-1	Q99497					-1.66 *

Master Spot N. *a)*	Spot Entry name *b)*	Description	AC	Fold change *c)*				
				Disease *vs* Control	Head *vs* control	Hinge *vs* control	R496C *vs* control	R496H *vs* control
3049	SODC	Superoxide dismutase [Cu-Zn]	P00441				-1.38 *	
Cellular fate								
1364	ANXA6	Annexin A6	P08133					1.32 *
1376	ANXA6	Annexin A6	P08133					1.38 *
2371	TXNL1	Thioredoxin-like protein 1	O43396	1.53 *		1.60 *		
2568	AN32A	Acidic leucine-rich nuclear phosphoprotein 32 family member A	P39687	-1.32 **		-1.33 *	-1.58 **	
3380	MIF	Macrophage migration inhibitory factor	P14174	-1.47 **	-1.37 *	-1.51 **	-1.71 **	
Protein fate (folding, modification and degradation)								
972	ENPL	Endoplasmin	P14625	1.50 *		1.52 *		
1045	ENPL	Endoplasmin	P14625		1.85 *			
1271	GRP78	78 kDa glucose-regulated protein	P11021	1.34 *	1.70 *			
1325	HSP7C	Heat shock cognate 71 kDa protein	P11142		1.98 **			
1351	HSP7C	Heat shock cognate 71 kDa protein	P11142	1.34 **		1.48 *	1.65 **	
1687	PDIA3	Protein disulfide-isomerase A3	P30101		1.46 **			
1713	PDIA3	Protein disulfide-isomerase A3	P30101		1.56 **			
2739	PSB4	Proteasome subunit beta type-4	P28070	-2.35 **		-2.27 *		
2988	SSRD	Translocon-associated protein subunit delta	P51571	-1.65 *	-2.37 *			
2799	BAG2	BAG family molecular chaperone regulator 2	O95816	-1.37 **	-1.38 **	-1.33 *		
3025	PPIB	Peptidyl-prolyl cis-trans isomerase B	P23284		1.73 *			
3032	PPIB	Peptidyl-prolyl cis-trans isomerase B	P23284		1.53 *			
3327	FKB1A	Peptidyl-prolyl cis-trans isomerase FKBP1A	P62942			-1.34 *	-1.38 **	
3572	HSP7C	Heat shock cognate 71 kDa protein	P11142		2.31 **			
3633	CH60	60 kDa heat shock protein, mitochondrial	P10809	-1.40 **				
Transcription regulation and RNA processing								

J Proteome Res. Author manuscript; available in PMC 2013 December 07.

Master Spot N. *a)*	Spot Entry name *b)*	Description	AC	Fold change *c)*				
				Disease *vs* Control	Head *vs* control	Hinge *vs* control	R496C *vs* control	R496H *vs* control
993	EF2	Elongation factor 2	P13639	1.37 **	1.42 **	1.35 *		
995	EF2	Elongation factor 2	P13639		1.32 *			
1893	HNRH1	Heterogeneous nuclear ribonucleoprotein H	P31943		-1.32 **	-1.32 *	-1.64 **	
2052	EFTU	Elongation factor Tu, mitochondrial	P49411	-1.43 *				-1.78 *
2576	ROA2	Heterogeneous nuclear ribonucleoproteins A2/B1	P22626		-1.42 **		-1.32 *	
2700	CPSF5	Cleavage and polyadenylation specificity factor subunit 5	O43809	-1.32 **	-1.47 **			
3325	H4	Histone H4	P62805	1.34 **	1.45 *	1.30 *		

a) Master spot numbers showed in Figure 1.

b) Protein identification code according to UniProtKB.

c) Significant average ratios (≤ -1.3 or ≥ 1.3;) among CdLS and control cell lines.

* $p < 0.05$,

** $p < 0.01$