	10	2	0	30	40	50	60	70	80	90	100
						<u>. .</u>					
	TM1				TM2				тмз		
BePD193	-CVSTAVTTLL	CTGTLEGT	TSCSEATTE			кт. тsк	K LT EH ETMOET		NCTLTTCSAV	AT.VNA TOST	TT. HC
EcYiiP	-IAATAMASLLI	LIKIFAWW	YTGSVSILA	ALVDSLVDIG	ASLTNLLVV	RYSLO	PADDNHSFGH	GRAESLAALA	OSMFISGSAL	FLFLTGIOHI	LISP
ReCH3072	LSLGV	IGLKMVAWI	VTGSVALLS		AFIAFFVI	RY AQK	PADHDHPFGHI	HKAEYLSAVT	EGVLIVVAAL	LIVNEAIGYI	LAAP
ReNepA	SFVILAITA	ALQLVVVF	YS GS VAL LA	DTIENVGDAA	TAIPLWIAF	SLVRR	AATKTFNYGL	GRV EDYAGLI	IVLIILFSAL	VAGYEAIDRI	LNP
HsZnT5	IFYFICLNLLFI	FVELFYGV	LTNSLGLISI	DGF <mark>HMLFDC</mark> S	ALVMGLFAA	LMSRW	KAT RI FSY GY	GRI EILSGFI	NGLFLIVIAF	FVFMESVARI	LI DP
ScZRC1	LITLDTVFF	ILE IT IGY	MSHSLAL IA	SEHMLNDII	SLLVALWAV	DVAKN	GPDAKY TY GWI	KRA EILOALI	NAVELIALCE	SIMIEALQRI	LI EP
AtMTP1	-CIAVVLCLVF	ISVEVVGGI	KANSLAII	AAHLLSDVA	P FRISLFSL	WAAGW	EATPROTYGE	FRIEILGAIV	SIQL <mark>I</mark> WLLT <mark>G</mark>	ILV <mark>YE</mark> AIIRI	EV TE
WmDmeF	LVMWITLAM	IVE IA AGL	VFNSMAL LA	GWHMSSHAL	AIGLSAFAY	AAARR	SQDGRFSFGT	WKI EVLAAFA	SAIFLLGVAG	LMVFGSVERI	FTP
ReCepA	IWLVIGLTAVM		LY GS MAL VA	DGWHMSTHAS	ALLISALAY	LE'ARK	ARN PRFTFGT	GRL GDLAGFA	SALILALIAL		SNP
WINC ZOD	LKIALALTGTFI		MIRSLALIS			ALARR	PADKKRTFGI		NALLLFGVAL	I ILI EAILRI	
ECAICB	LL INFOVINGE		TO GO LALINA		ALLEALEAV	QE SKK	F FI IKHIF GWI			LIVWEALERI	KIP
				TM4		_	M5		(TM6	
										TIMO	
RePD193	-GRP LNP		-IVYAVVTL	GVT AT MGV IG	DRANRKI	KS DFVALDT	KAWIMSAG-L	r ig llvafai	GYF IQGSS F II	DPVALALICI	IVII
EcYiiP	-TPM TDE		-VIVTIVAL	I CT II L VS FQ	RWVVRRT	QS QAVRADM	L <mark>H</mark> YQS <mark>D</mark> VM-MI	NGA ILLALGL	SWYGWHRA	DALFALGIGI	ſŸIL
ReCH3072		VLG-	-LAINLAAG	VINAVWARLI	IRAGRKH	RSAALAADG	QHIMSDVV-T:	SAG VLVGLLL	AL-ATGYAIF	DPVLAILVA	ENIL
Renepa		AAV -				SSAAL LADG					
ScZBC1		$P = \mathbf{R} \mathbf{I} \mathbf{M} - \mathbf{R} \mathbf{I} \mathbf{M} - \mathbf{R} \mathbf{I} \mathbf{M} - \mathbf{R} \mathbf{I} \mathbf{M} \mathbf{M} \mathbf{R}$		I SNWGLEL	HDESKKDCH	RS INMIGUE	LHVIGDAL-G		TWKTEYSWYS	PIUSLITT	
A +MTP1	-SEVNG-	FTM-	-FIVAAFGU		GHHGK RK	RN TNLOGAY	THVIGDST-0	SVGVMTGGAT		DUICTUAES	<u>л л</u> .
WmDmeF	-QPIHY-		-MAITAIGL	IVNLACALII	GGAHHGHRH	HDINLRSAY	LHVVADAA-T	SVLAIVALAG	GWW-LGWSWL	PVMGLVGA	VLVG
ReCepA	-VPIGF-		-IAVAVIGL	AVNLASAWLL	AGG GHHS	GD NN I RAAY	LHVVADAL-T	SVL AI AAL TL	GSL-YGWLWL	DPIMGIVGGI	IVIA
WmCzcD	- PQI ES-		-FVVAVLGL	IINLISMRMI	SSGQS	SS LNVKGAY	LEVWSDLL-G:	SVG VIAGAII	IRF-TGWAWV	SAIAVLIGI	WVL
EcZitB	- R P V E G-		-MAIAVAGL	LANILSFWLL	HHGSE-E	KNLNVRAAA	L <mark>EVIGD</mark> IL – <mark>G</mark>	SVG AIIAALI	IIW-TGWTPA	DPILSILVSI	LVL
	TMA		14		84	01			110		
	ТМо				51	52			пг		55
RePD193	PTPACTVRRATA	DT LIVTPL		VAKTERHGET		-GRGROTEL	HEVVSENT.K-		ETGLATONEG	PSRRI.TI	
EcYiiP	YSALRMGYEAV	SLLDRALP	DEEROEIID	IVISWPGVSG	ADIRTROS	-GPTRFIOI	HLEMEDSLP-	- LV OAHMVAD	OVEOAILREF		HODPCSV-
ReCH3072	YQGWKVISQSIS	GLMDQAVE	POEEEAIKO	AIATAAGSIG	VHDLKTRRA	-GTVTFIDF	HMVVPGTMS-	- VR QAHDI CD	RIEDAIRAVH	E GAKIAI	HVEPEGE-
ReNepA	GIVWQSGRAVVI	RSLDGV-E	PWITDEIRH	A AE HVRGIDE	VVDVKARWL	-GHKLFTDV	VIAVDRSKN-	-VS EANAIAS	ALRRELQGHL	PSLGNATI	EQ FD
HsZnT5	LSWPLIKDAC	VLLLRLPP	EYE-KELHI	ALEKIEGLIS	YRDPHFWRH	SASIVAGTI	HIQVTSDVL-	E Q RIV Q	QVTGILKD	AGVNNLTI	IQVE
ScZRC1	SSALPLSRRASE	ILLQATPS	TISADQIQRI	EILAVPGVIA	VHDFHVWNL	TESIYIASI	HVQ IDCAP	D KFMSSAK	LIRK-IFHQH	GIHSATV	QPEFVSG-
AtMTP1	GTTINMIRNIL	VUMESTIPR	EIDATKLEK	JLL EME EV VA	VHELHIWAI	TVGKVLLAC	HVN IRPEAD-	-ADMVLN	KVI DY IRREY	NISHVTI	121E
wmDmeF ReCenA	KWAIGLMRQSGI		HP VVEEVRE	VLAQGEDGTR TETEEDE		CPCHHAD TV	SLVTHDASL-	- II PQ	KVRHALSIHD	ELVHVSV	/E IN
WmCzcD	PRIWILLKSSIN	VLLEGVPD	DVDLAEVEK	DILAT PGVKS	FHDINI I WAT.	TSCKASLTV	VVND TAVN-	PEMEVT.P	EIKOMLADKE	D IT	OFELAPO
EcZitB	RSAWRLLKDSVN	ELLEGAPV	SLDIAELKR	RMCRIPEVRN	VHHVHVMMV	G-EKPVMTL	HVOVI PPHD-	-HD AL LD	O IOHYLMDHY-		IOMEYOPC-
02									R R	~	-x x- v

Figure S1. Structural and functional key residues conserved in the CDF family (PF01545). Transmembrane domains (TM1-6), residues of metal binding sites A, B, and C (red, blue and green dots) respectively, as well as the interlocked (Lys⁷⁷-Asp²⁰⁷)₂ salt-bridges (black dots) were located by comparison with the reported structure of YiiP⁵. The consensus sequence region reported for the CDF protein family⁹ is black underlined (in TM2-3). The amino acid positions involved in protein function/activity as well as substrate selectivity (yellow highlighted) were located by comparison with previous works ^{4,23,25,26,36,38}. The CDF conserved regions were obtained from the Pfam HMM-domain and aligned with ClustalO. This 11-sequences alignment is a sample of the alignment of 318 CDF sequences used to infer the phylogeny.



Figure S2. Comparison of conserved residues carrying potential N- ligand donors present in CDF proteins of clade III (Ni²⁺/Co²⁺), Zn²⁺-transporting CDF (clades III, V, XI, VII,VIII) and 21 sole Zn-CDF proteins contained in the Table S1. The putative metal binding site A among the different groups is indicated by red arrows. The conserved and group specific Asn and Arg residues (black arrows) in group III proteins were present and functionally analyzed in NepA. Numbering corresponds to *R. etli* NepA, *E. coli* YiiP, *A. thaliana* AtMTP1, *S.cerevisiae* ZRC1 and *H. sapiens* ZnT5 proteins from III, V, VII, X and XI clades, respectively. For the 21 sole Zn CDF proteins the numbering of *E. coli* ZitB was used. Logos were built using Consurf (see methods).



Figure S3. Group-specific His-rich stretches present in Co²⁺, Zn²⁺ and Ni²⁺ transporters grouped in clades XII (A), VII (B) and III (C), respectively. In HvMTP1³⁸ or AtMTP1⁴⁰ proteins of group VII middle His-rich tracts are required for proper protein function and substrate selectivity. The role of equivalent regions in group III and XII proteins is unknown. The number of His-residues over the complete stretch-length is shown in parenthesis. Characterized CDF proteins are shown in black bold. The red line indicates the position of the first residue (TM2) of the putative metal binding site A.