

Supplemental Discussion of the RNA -dependent and -independent A3 PPI Network

In our high confidence network, A3A and A3C both represent unconnected networks, with A3C only sharing a single interactor, PHAX, with A3G. Both A3A and A3C are single domain proteins that are expressed mainly in immune cells, though A3C has lower tissue specificity than A3A. A3A is localized to the nucleoplasm of cells and functionally has been linked to a number of cancers (106), restricts LINE-1 elements (96), restricts HPV infection (107), and has been shown to modify foreign and non-genomic DNA (108). We identify 9 high confidence A3A interactions under both conditions. These include TAXB1/TAX1BP1, a protein that regulates TNF-induced apoptosis and NF-kappa-b signaling (109, 110). TAXB1 also interacts with the HTLV-1 oncoprotein Tax1 and it has been suggested that A3A can restrict HTLV-1 infection and edit the HTLV-1 genome (111, 112). A3C localizes to intermediate filaments and in some cases can be nucleolar (113). While A3A has a small RNA independent subnetwork, A3C has a total of 26 high confidence interactors, with 11 having a DIS < -0.5 (higher in NT) and 4 having a DIS > 0.5 (higher in +RNase). Among the A3C interactors, SRRT and PHAX are the strongest that bind A3C under both NT and +RNase conditions. Interestingly both SRRT and PHAX interact with the CBC complex which is comprised of NCBP1 and NCBP2, both of which A3C did not have a high confidence interactions with, though it should be noted that under NT conditions, A3C pulls-down NCBP1 but does not pass both SAINT and CompPASS stringency thresholds (**Table S5**). Potentially A3C interacts with the CBC complex via PHAX and SRRT, but the interaction with CBC is mediated by RNA.

A3G is an important antiviral factor involved in restriction of retroviruses, and potently restricts HIV in the absence of Vif. In total, A3G interacts with 43 high confidence interacting proteins, including 17 proteins that interact specifically with A3G. The A3G sub-network includes 14 proteins that have a DIS < -0.5 and are stronger interactors in NT conditions, and 4 proteins that have a DIS > 0.5, and are stronger interactors in +RNase conditions, but many of the core interactors are RNA independent. While A3G is a double domain A3 protein, it shares only a few interactions with A3B, A3D, and A3F (**Figure S7**), the other double domain A3s. Among the shared interactors between all the double domain A3 proteins are a number of proteins involved in RNP complexes including 7SK RNP proteins (LARP7 and MEPCE), and RNA-binding protein RO60 (RO60/TROVE2) and Lupus La protein (LA/SSB), both of which are Y RNA-binding proteins. RO60 binds to misfolded non-coding RNA, pre-5S rRNA, U2 snRNA, Alu retroelements, LINEs, and Y RNAs, and RO60 RNPs have a number of documented functions in the cell including response to environmental stress, transport of RNA binding proteins, quality control of RNA, and apoptosis (103, 114, 115). La binds to the polyU tail of Y RNAs and is thought to stabilize RNA polymerase III products and prevent degradation (116–118). It has been shown that A3G complex formation with HIV nucleocapsid (NC) can be facilitated by Y RNAs (119). Additional Y RNA binding proteins, like MOV10, (120) have been shown to interact with A3G in previous studies (121), and interestingly in our hands, MOV10 is just below our stringent scoring criteria (**Table S5**), indicating it could be a true interactor, and that A3G interacts with Y RNA. In addition to interactions it shares with other A3 baits, A3G specifically binds to the PHAX-cap binding complex (CBC), including interaction with NCBP1 strongly in both NT and RNase+ conditions, and weaker interaction to both NCBP2 and NCBP3. Since NCBP1 forms a CBC with NCBP2, and an alternative CBC with NCBP3, it is likely that A3G binds to NCBP1 rather than the RNAs or NCBP2 or NCBP3 directly. It is interesting to note that A3G only binds to PHAX under NT conditions indicating that this interaction might be mediated by RNA. A3G has been shown to deaminate RNA (122) and here we find that A3G also interacts with a number of RNA editing and RNA modification enzymes. This includes a number of proteins involved in RNA splicing, tRNA methylation enzymes (THUM1, THUM3, WDR6, and TRM1), and pseudouridine synthesis enzymes (PUS3, PUS7, and TRUA). Notably, all of these tRNA and RNA editing enzyme interactions are only identified in NT conditions and are likely mediated by RNA. This indicates a

strong likelihood that A3G interacts with tRNAs and could be involved in tRNA biology in either a deaminase-specific or a deaminase-independent role. A3G also interacts with another host restriction factor, ZCCHV, also known as zinc-finger antiviral protein (ZAP). ZAP functions to sense non-host RNA by binding to CG dinucleotides, which occur in the human genome less frequently than expected (123). That A3G interacts with ZAP may be a mechanism to ensure viral restriction by delivering two anti-viral factors to the viral genome.

In this study we investigate the two haplotypes of A3H that are relevant to cancer (A3H-I) and virus restriction (A3H-II). These two haplotypes differ at 3 polymorphic sites, 105 G/R, 121 K/D, and 178 E/D for A3H-I and A3H-II, respectively (22). A3H-I is rapidly ubiquitinated in cells leading to a short half-life, but still has demonstrated deamination activity and is one of the most common A3H haplotypes in the human population (22, 23) A3H-II is more stable in cells and has been primarily characterized as an anti-retroviral restriction factor (22, 23). In our hands, even under constitutive overexpression by transient transfection in HEK293T cells, we find a similar trend with A3H-I being expressed at lower levels than A3H-II (**Figure S1**). As expected, the A3H-I and A3H-II subnetworks share a subset of interacting proteins, though A3H-I has more total high confidence interactions than A3H-II, with A3H-I having 46 and A3H-II having 17 (**Figure S7, Table S5**), possibly as a result of lower expression of the bait, but potentially due to functional differences as well. Correspondingly, A3H-I has a larger number of unique interactors, including 23 that are A3H-I specific and not shared with A3H-II or any other A3 protein. Interestingly, three components of the RNase P complex (*i.e.* POP1, POP7/RPP20, and RPP25) were identified as A3H-I interactors (**Figure S7**) under NT and +RNase conditions, although the NT interaction was higher confidence. The human RNase P complex is comprised of 10 proteins and a single catalytic RNA (H1) that acts as an essential endonuclease that processes precursor tRNA, specifically by removing the 5' leader sequence. Recently solved cryo-EM structures of human RNase P in complex with or without tRNA^{val} demonstrated a clamp-like structure of 10 proteins surrounding the H1 RNA that can then bind to the tRNA substrate through both protein-RNA and RNA-RNA interactions (124). In this structure, POP1, POP7/RPP20, and RPP25 form the 'finger' component of the human complex that wraps around and stabilizes the C domain of the H1 RNA, with POP1 also providing contacts with the tRNA substrate (124). In addition to POP1, POP7, and RPP25, A3H-I interacts with a number of proteins involved in tRNA processing and methylation, indicating a potential for A3H-I to bind tRNA or be involved in tRNA processes. Specifically, A3H-I interacts with RNZ2/ELAC2, TRM6/TRMT6, TRM61, TRM1L, and a number of shared interactors of A3G and A3H-II including THUM1, THUM3, NSUN2, GRSF1, DIM1, MEPCE, and LARP7. Interestingly, VCIP1, a deubiquitinase, specifically interacts with A3H-I and this may enable A3H-I to gain cellular stability at times when VCIP1 is upregulated, such as during DNA repair (125).

A3B, A3D, and A3F are all double domain A3 proteins that have NTDs that share sequence similarity and have overlapping high confidence interactors. If we look at the individual subnetworks we find that A3B, A3D, and A3F interact with 31, 31, and 34 high confidence proteins, respectively, with 9 being pulled-down by all three and 23 found in at least two (**Figure S7**). A3B expression has been correlated in a number of cancers, particularly breast cancer, and has been implicated in prognosis, patient survival, drug resistance, and mutagenesis (77, 83, 126–128). A number of the A3B interacting proteins have also been linked to cancer in both pro-oncogenic and tumor-suppressor roles, indicating a potential deaminase-independent role of A3B through PPIs. Scaffold attachment factor B1 (SAFB1) is an RNA binding protein that has been implicated in breast cancer through binding and regulation of estrogen receptor (ER)-alpha (129, 130). Interestingly, it has been shown that A3B expression is inversely correlated to the clinical benefit provided by tamoxifen, which is an ER modulator used to treat metastatic ER+ breast tumors (127). RNA-binding protein 6 (RBM6) is a tumor suppressor that is often down-regulated or mutated in cancers (131). RBM6 has been shown to regulate DNA double-strand breaks and its protein expression is significantly lost in metastatic breast tumors compared to primary tumors

(132). Glyoxalase 2 (GLO2) catalyzes the glutathione-dependent metabolism of methylglyoxal, a pro-apoptotic product of glycolysis. Given many cancers have increased rates of glycolysis, upregulation of the enzymes that turn over toxic byproducts can be steps towards cancer progression and cell survival (133, 134). In addition to urological cancers, overexpression has also been linked to prostate cancer (133, 134). Aberrant expression of the A3B-specific interactor PLOD1 has been linked to tumorigenesis for a variety of cancers (135–138). Of the 9 proteins that come down with A3B, A3D, and A3F, all but one stably co-purify under NT and +RNAse conditions. Five of these are components of the Prefoldin Complex (PFD1-6), with all but PFD2 stably co-purifying with all three A3s, and PFD2 stably interacts with A3B and A3F but not A3D. The other four proteins include HNRL2, ZN326, UGGG1 and HUWE1. HNRL2 and ZN326 are both RNA binding proteins with ZN326 being a component of the DBIRD complex which is also able to bind DNA (139). UGGG1 plays a role in regulating the recognition, refolding, or degradation of misfolded glycoproteins, while HUWE1 is an E3 ubiquitin ligase that has important roles in a number of human diseases and cancers (140–143). Interestingly, HUWE1 interacts with HIV-1 integrase (IN) and virion infectivity factor (Vif), and is known to restrict HIV-1 infection (144, 145).

In general there was common interactions with ribonucleoproteins for A3B, A3F, A3G, A3D, A3H-I, and A3H-II, e.g, GEMI5, RALY, ROA3, ROA2, HNRPL, DHX30, RL27, RENT1, and RO60, however the whether the interaction occurred in NT, +RNAse, or both varied. GEMI5 interacted with A3F, A3G, and A3H-I only in the NT condition. RALY interacted with A3D, A3F, and A3G only in the +RNAse condition. ROA3 interacted with A3B and A3F in both conditions and HRNPL interacted with A3B and A3D under both conditions. Overall, even though the type of interaction varied, for each PPI there was similarity between the type of interaction for all baits across the dataset. A3D uniquely interacted with proteins involved in the biological process of cellular response to stress (TEBP, FKBP5, BAG2, DNJB1, DNJC7), with only STIP1 having an overlapping interaction with A3F and A3B. These interactions occurred in NT and +RNAse conditions.