Selective accrual and dynamics of proteinaceous compounds during pedogenesis: testing source and sink selection hypotheses

Jinyoung Moon

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Horticulture

Mark A. Williams, Chair

Kang Xia

Brian D. Strahm

Richard F. Helm

Richard E. Veilleux

September 3rd, 2015

Blacksburg, VA

Keywords: Soil organic nitrogen (SON), soil organic matter (SOM) accumulation and formation, soil peptides and proteins, hydrolysable amino acid, hydrolysable amino sugar, mineral-associated organic matters, organo-mineral association, microbial contribution to SOM, soluble amino acid, microbial biomarkers, PLFA, Lake Michigan USA, Haast New Zealand chronosequence, soil ecosystem development, primary ecosystem succession

Selective accrual and dynamics of proteinaceous compounds during pedogenesis: testing source and sink selection hypotheses

Jinyoung Moon

Abstract

The emerging evidence of preferential accumulation and long residence time of proteinaceous compounds in soil are counter to the traditional view that their structure is readily broken down through microbial activity. The shift in thinking of their residence time is, however, heavily influenced by physical and chemical protections in soil, representing an important change for understanding global biogeochemical carbon and nitrogen cycling. We investigated the accumulation patterns of proteinogenic amino acids for a long term (thousands of years) related to their sources and sinks. We found clear patterns of change in the amino acids in a 4000 year-chronosequence adjacent to Lake Michigan, USA (Michigan chronosequence) and they were tightly related to the shifts in their biological sources, namely aboveground vegetative community (r^2=0.66, $p<0.0001$) and belowground microbial community (r^2=0.71, $p<0.0001$). Results also showed great variations of approximately 49% between seasons (summer and winter). Moreover, seasonal dynamic patterns (22% variations) of the amino acids in soil mineral associated fraction were rather counter to the conceptual view that it represents a slow soil organic pool with long residence times. The amino acids enriched in the mineral associated fraction, (e.g., positively charged, aromatic, and sulfur containing amino acids), tended to preferentially accumulate in whole soil pool during the 4000 years of ecosystem development. Their interaction with soil minerals, therefore, may play a

critical role in the long-term sink and selective accumulation of proteinaceous compounds with some degree of the displacement. This was further confirmed by another chronosequence system near Haast River, New Zealand, which is geologically separated and climatically- and ecologically- different from the Michigan chronosequence. Common trends between two chronosequences suggested that either polar interactions or redox reactions may be relatively more important in the mineral interaction of amino acids than non-polar interactions. The consistency of results at two disparate locations in the southern and northern hemispheres is strong evidence that the processes of pedogenesis and ecosystem development are parsimonious and predictable. Our research demonstrated fundamental understanding of behavior of proteinaceous compounds at the molecular species level, and further provided their partitioning mechanisms associated with soil components.

Acknowledgements

I would like to sincerely thank my advisor Dr. Mark A. Williams for the excellent guidance, endless support, valuable advice and consistent patience during my research. I would also like to express my great gratitude to committee members, Dr. Kang Xia, Dr. Brian D. Strahm, Dr. Richard F. Helm, and Dr. Richard E. Veilleux, for their guidance and support during my PhD studies.

Dr. Shankar G. Shanmugam is acknowledged for collecting soil samples from Lake Michigan chronosequence and analyzing PLFA. I would like to thank Dr. Benjamin L. Turner and Dr. Leo M. Condron for providing soil samples and information from Haast chronosequence, New Zealand. Dr. Madhavi L. Kakumanu is thanked for the density gradient fractionation work. I appreciate the HPLC instrumentation advice by Dr. Li Ma and Dr. Chao Shang. I could not present data related to protein work, but I would like to thank Dr. Keith Ray for advising protein purification procedures and MALDI-TOF MS/MS instrumentation.

My colleagues, Richard Rodrigues, Rosana Pineda, Hua Xiao, Kerri Mills in Rhizosphere soil microbial ecology and biogeochemistry lab are thanked for intellectual and mental support and criticism. I also like to acknowledge great help from undergrads, Angi Lantin, Haley Randolph, Tori Nelson, Yoonji Ha, and Audrey Longfellow.

Financial support was obtained from USDA-NIFA.

Table of Contents

List of Figures

Chapter 1. Introduction

Chapter 2. Selective accumulation of amino acids and proteins with minerals and association with plant-microbial communities

Chapter 3. Seasonal dynamics of soil organic nitrogen across a boreal-temperate successional sequence

Chapter 4. Similarity in selecting patterns of protein amino acid during pedogenesis in two disparate chronosequences located in Lake Michigan, USA and Haast River, New Zealand

Appendix A-Chapter 2

Appendix B-Chapter 3

Appendix C-Chapter 4

Appendix D

List of Tables

Appendix A-Chapter 2

Appendix B-Chapter 3

Appendix C-Chapter 4

Attribution

Chapter 2. Selective accumulation of amino acids and proteins with minerals and association with plant-microbial communities

i. Authors: Jinyoung Moon[1], Li Ma[2], Kang Xia[3], Mark A. Williams[1]
ii. Institute:
[1]Soil Microbial Ecology and Biogeochemistry Laboratory, Department of Horticulture, Virginia Polytechnic Institute and State University, 312 Latham Hall, 220 Ag Quad Ln., Blacksburg, VA 24061
[2]Department of Environmental Sciences, University of California, Riverside, CA 92521, USA and USDA-ARS, Soil Physics and Pesticides Research Unit, George E. Brown Jr. Salinity Laboratory, Riverside, CA 92507, USA.
[3]Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, 1880 Pratt Dr., Blacksburg, VA 24061

Chapter 3. Seasonal dynamics of soil organic nitrogen across a boreal-temperate successional sequence

i. Authors: Jinyoung Moon[1], Kang Xia[2], Mark A. Williams[1]
ii. Institute:
[1]Soil Microbial Ecology and Biogeochemistry Laboratory, Department of Horticulture, Virginia Polytechnic Institute and State University, 312 Latham Hall, 220 Ag Quad Ln., Blacksburg, VA 24061
[2]Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, 1880 Pratt Dr., Blacksburg, VA 24061

Chapter 4. Similarity in selecting patterns of protein amino acid during pedogenesis in two disparate chronosequences located in Lake Michigan, USA and Haast River, New Zealand

i. Authors: Jinyoung Moon[1], Kang Xia[2], Benjamin L. Turner[3], Mark A. Williams[1]
ii. Institute:
[1]Soil Microbial Ecology and Biogeochemistry Laboratory, Department of Horticulture, Virginia Polytechnic Institute and State University, 312 Latham Hall, 220 Ag Quad Ln., Blacksburg, VA 24061
[2]Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, 1880 Pratt Dr., Blacksburg, VA 24061
[3]Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancon, Republic of Panama

Chapter 1. Introduction

1.1. New paradigm of soil organic matter (SOM) persistence

Over the last decade, the traditional paradigm of the relationship between the age of carbon (C) molecules of SOM and the recalcitrance of molecular structure to biodegradation has become less accepted from critical reviews and evidence (Amelung *et al.*, 2008, Gleixner, 2013, Grandy & Neff, 2008, Kleber, 2010, Knicker, 2011, Marschner *et al.*, 2008, Schmidt *et al.*, 2011). The initial decomposition rate of plant residues correlates broadly with indices of their bulk chemical composition, such as the content of nitrogen (N) or lignin (often operationally defined as chemically resistant biomolecules to acid hydrolysis) (Melillo *et al.*, 1982). Traditionally, the initial decomposition rates of organic compounds are extrapolated to explain their long-term persistence in soils. In other words, more chemically resistant organic compounds are slowly decomposed during the initial stage of decomposition and are thus predicted to be selectively preserved in soil for a long term. Emerging evidence, however, is less likely to support the traditional concept that the selective preservation of recalcitrant primary biogenic compounds is a major mechanism for long-term SOM stabilization (Gleixner *et al.*, 2002, Hamer & Marschner, 2002, Hamer & Marschner, 2005). Structurally resistant compounds, mainly lignin and its derivatives, for example, are predicted to persist in soils; but their mean turnover times are faster than the bulk of SOM. Molecular fragments of proteins and carbohydrates, which are chemically and biologically labile, are thought to be rapidly metabolized in soil; however, they are observed to have much slower turnover rates than lignin. Taken together, these findings suggest a more complicated picture of recalcitrance and stabilization mechanisms that

allow organic matter to persist in soils. This shift in the paradigm of SOM persistence and turnover enlightens a change in understanding global biogeochemical cycles and presents challenges to developing robust models of global C turnover.

This further highlights the important contribution of so-called biologically labile molecules to SOM formation and their stabilization mechanisms which are not associated with intrinsic structure. Proteins, which are readily cleaved and degraded by various proteases in solution (Milo *et al.*, 2010) but shown to persist in soil for a long term, particularly, are the focus of this dissertation due to their central role in linking soil C and nitrogen (N) cycles and soil fertility (Knicker, 2011). Proteins and their derivatives in soil have shown to be mostly derived from in situ formation through microbial incorporation of plant materials (Kramer & Gleixner, 2006). The decomposition processes of these microbial-derived labile compounds are often found to be retarded through physicochemical protections in the soil matrix (Krull *et al.*, 2003), suggesting that their interactions with solid components in soil, such as minerals and organic aggregates, can be important to their long residence times. Moreover, they are constantly resynthesized in all organisms of the soil food chain so they are continuously present in soil due to their biological importance (Gleixner, 2013). Many important questions related to the turnover and persistence of proteins in soil, however, remain to be resolved. This dissertation is intended to question the relationship of source and sink of proteins with their long-term accrual patterns.

1.2. Source and sink of proteinaceous compounds

The origins of proteins are soil organisms including plants, animals, microbes, and microbial fauna. Proteinaceous compounds (including proteins, peptides, and their

derivatives) are the predominant form of N in these many organisms. Most vascular plants have a relatively low content of proteinaceous compounds per biomass. Between 2% and 15% of the plant mass is assigned to N-containing compounds and mostly to amino acids (Knicker, 2004). About 50% of bacterial and 30% of fungal biomass can be assigned to proteinaceous compounds (Christias *et al.*, 1975, Neidhardt *et al.*, 1990). In addition, bacteria, especially gram-positive groups, contain abundant peptides in cell wall-peptidoglycans. In some fugal cell walls, melanins, dark-colored pigments are observed. Peptides also are commonly used for communication and signaling between organisms; however, the extent of peptide production and turnover for these purposes is not well described (Farrell *et al.*, 2011, Farrokhi *et al.*, 2008). Because of the relatively high content of proteinaceous compounds in microbes, they have the potential to provide substantial amounts of proteinaceous molecules in soil. The concentration of proteinaceous compounds in plants is small, but their biomass inputs are responsible for all the C flow into soil (Kögel-Knabner, 2002). Therefore, they are also likely to play important roles in the fate of proteinaceous materials.

When plant derived proteinaceous compounds enter the soil system, they can be subject to attack from microbial extracellular enzymes (Fig.1.1, Pathway (1)). Through enzymatic activities, the proteinaceous macromolecules break down into lower molecular weight compounds (e.g., peptides and amino acids) and into the even smaller, inorganic N compounds (e.g., ammonium and nitrate). Small enough sizes of proteinaceous compounds are utilized by microbes, microbial fauna, and are taken back up by plant roots (Schimel & Bennett, 2004). Some portions of them, however, can escape from biodegradation. The remaining plant derived proteinaceous compounds

are associated with mineral surfaces and organic aggregates becoming physically and chemically protected in the soil matrix and avoid biological attacks. This can be one way for them to be preserved in soil. However, it is not known how much undecomposed plant-derived proteinaceous compounds contribute to SOM formation.

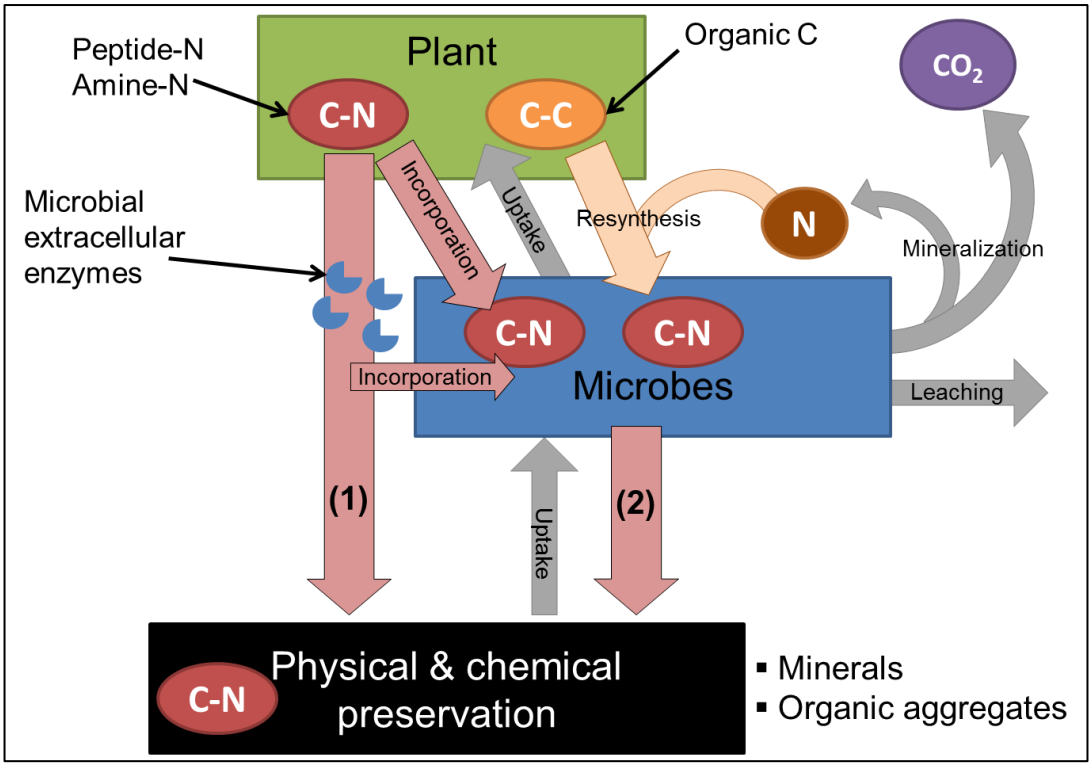


Figure.1. 1. Conceptual model of formation and fate of proteinaceous compounds in soil. Two main pathways for proteinaceous compounds to undergo the preservation processes: **(1)** direct pathway of plant materials and **(2)** microbial mediated pathway including incorporation of proteinaceous compounds into microbial cellular biomass and resynthesis of new molecules of proteinaceous compounds (Gleixner, 2013). "C-N" represents proteinaceous compounds including peptide-N and amine-N. "C-C" represents other C-rich organic compounds derived from plants, including lipids, carbohydrates and lignin. The width and length of arrows does not represent the size or rate of pool fluxes.

Plant derived C is incorporated into cellular biomass through microbial assimilation, with the supply of N from SOM (Fig.1.1, Pathway (2)) (Gleixner, 2013). Microbes recycle C atoms derived from plant material to resynthesize new molecules for cellular needs, (e.g., proteins for structure and function). The cellular proteins eventually are released to the soil as part of cell death and other functional purpose (e.g.,

extracellular enzymes and signaling peptides). They are again taken up for cell growth, but some of these cellular proteins that are not recycled can form the basis for the production of non-living biomass SOM pools. There are several mechanisms used to explain why proteins would tend to remain undecomposed and are not cycled in soil. Their reactivity due to chemical varieties of functional side chains can provide extra stabilization and long persistence in spite of their lability to biological decomposition (Knicker & Hatcher, 1997, Rillig *et al.*, 2007, Wershaw, 1986). Physical stabilization mechanisms, such as the binding to mineral surfaces due to electrostatic force, metal–ligand interactions, atomic bonds, and van der Waals forces, have been suggested to protect organic molecules from decomposition because enzymes cannot access these bound molecules (Kaiser *et al.*, 2002, Kleber *et al.*, 2007, Mikutta *et al.*, 2006, Sollins *et al.*, 2006, Wershaw & Pinckney, 1980).

The mesopore protection hypothesis proposes that organic matter (OM) may be protected by sequestration within mineral mesopores (2–50 nm diameters). Because mineral surfaces are often dominated by the internal surfaces of mesopores (Mayer, 1994), it has been suggested that mineral mesopores may play a major role in the preservation of OM in sediments by protecting OM from degradative attack by bacteria or bacterial extracellular enzymes through physical occlusion within small microbial free mineral pores (Harms & Bosma, 1997, Hulthe *et al.*, 1998, Mayer, 1994). Amino acid monomers and polymers (<1.4 nm diameter) have shown to adsorb strongly onto mesoporous minerals while the size exclusion was found for proteins larger than the mesopores in aqueous suspension experiment (Zimmerman *et al.*, 2004). However, the proteins larger than mesopores were found to undergo different interactions with

mineral surfaces, showing strong adsorption to minerals without changing their original forms (Ding & Henrichs, 2002). Alternatively, but also based on adsorption to mineral surfaces, Sollins *et al.* (2006) suggested that proteinaceous compounds may form a stable inner organic layer around a mineral surface and this inner layer may help less polar organic compounds bind more readily to the surfaces of this mineral-organic layer. Similarly, the model of hemimicellar coatings on the mineral surface have suggested that proteinaceous compounds may play a prominent role in the structure of organo-mineral complexes due to their ability to adsorb irreversibly to mineral surfaces (Kleber *et al.*, 2007).

The biopolymer interaction hypothesis proposes interactions between organic molecules for their stabilization, explaining ubiquitous preservation of proteinaceous compounds regardless of the presence of inorganic minerals and metal-ions. In the so-called encapsulation model, proteinaceous compounds are connected to resistant aliphatic polymers (hydrophobic macromolecules) and surrounded by these polymers, and therefore they are protected from biological degradation (Knicker & Hatcher, 1997, Zang *et al.*, 2000). The mechanisms also include chemical incorporations and reactions of proteinaceous compounds with reducing sugars (Maillard reaction), polyphenols, quinones, and tannins (Espeland & Wetzel, 2001, Fan *et al.*, 2004, Nguyen & Harvey, 2001). Alternatively, Wershaw (1986) proposed a molecular aggregate model based on supramolecular chemistry of biomolecule residues. In this model, proteinaceous compounds play an important role in forming an amphiphilic structure and they are enveloped and stabilized in the core of aggregates through the various interactions with other organic matter constituents. The bonding structure of the molecular aggregates

may be weak interactions such as hydrogen bonding and hydrophobic interactions, rather than covalent bonds.

Another suggested hypothesis is intrinsic stabilization of proteinaceous compounds by modification of their key groups that are recognized by enzymes or conformational restrictions. For example, amyloid aggregates and fibrils efficiently protect proteinaceous compounds from biodegradation in the soil ecosystem (Nelson *et al.*, 2008, Rillig *et al.*, 2007). These filamentous proteins (e.g., hydrophobins and other membrane and cell wall proteins) are more resistant to biodegradation compared to cytoplasmic proteins due to the complexity of bimolecular mixtures and their rigid structural functions (Wessels, 1997). Increasing evidence of the persistence of cell wall constituents in soil has suggested the patchy fragment formation cycle where microbial necromass disintegrates into fragments, especially flat cell wall fragments, attached to mineral surfaces and forms a substantial part of the SOM (Miltner *et al.*, 2012). Although the focus of research in this dissertation is on understanding how proteinaceous molecules contribute to SOM formation, some results will provide clues related to the hypotheses of stabilization that result in OM persistence over relatively long time periods as its sink mechanisms.

Amino sugars, another cell wall related group of molecules, are sometimes used as biomarker for microbial biomass. The assessment of amino sugars provides tracing microbial components in soil. The advantage of four amino sugars (glucosamine, muramic acid, mannosamin, and, galactosamine) provide contributions of microbial groups due to their different origins (Amelung, 2001). Although identifying amino sugars and using their ratios help understand the contribution of microbial groups to SOM

formation, however, simplicity can mislead the interpretation because each amino sugar has different reactivity and turnover rates in soil; thus, care certainly is needed to use these biomarkers (Hobara *et al.*, 2014). Amino sugars are structural constituents of the microbial cell wall and often coexist with amino acids, (e.g., peptidoglycans). Amino sugars, therefore, provide a complementary means of describing SOM formation and are expected to provide the context of amino acid base analysis.

1.3. Objectives and hypotheses

Core hypothesis: *Long-term persistence of soil proteinaceous compounds is affected by (1) source and (2) sink. They selectively accumulate in relationship with their sources through biological cycling and their sink through chemical interaction with minerals in soil.*

The source of proteinaceous compounds largely controls their abundance in soil. Through recycling and resynthesizing processes, in situ formation of proteinaceous compounds occurs continuously in addition to plant material inputs. By changing biological sources, consequently the proteinaceous compounds will change. It is hypothesized that biological sources and their cycling selectively change proteinaceous compounds that remain in soil. The sink mechanisms of proteinaceous compounds are explained by physical and chemical interactions of these compounds with mineral surfaces and organic aggregates. Therefore, it is hypothesized that the proteinaceous compounds selectively associate with mineral particles in soil.

Chapter 2. Selective accumulation of amino acids and proteins with minerals and association with plant-microbial communities

Objective 1: *To determine if the distribution of proteinogenic amino acids in whole soil organic matter (OM) pool and mineral associated OM sub-pool change during 4000 years of ecosystem development and if their distribution are different between the whole soil pool and the mineral associated sub-pool.*

Hypothesis 1-1: *The relative distribution of the amino acids changes during ecosystem development and pedogenesis.*

Hypothesis 1-2: *There is correlation between biological community successions and change of amino acid distribution during ecosystem development (source hypothesis).*

Hypothesis 1-3: *Positively charged amino acids are preferentially associated with primary silicate mineral (sink hypothesis).*

To investigate the long-term dynamics of proteinaceous compounds, a chronosequence approach was used. The gradients of ecosystem development in the chronosequence provided an ideal place to determine the change of proteinaceous compounds related to SOM formation for long term (~4000 years) pedogenesis and aboveground- and belowground-biological community successions under similar climates and soil parent materials.

The main focus is on variations in molecular species of proteinaceous compounds (proteinogenic amino acids) associated with whole and mineral derived soil pools during pedogenesis and ecosystem development. Here, the whole soil OM pool represents the bulk of SOM, which largely consists of organic aggregates of non-living biomass. The mineral associated OM sub-pool is part of whole soil OM and operationally defined by the density gradient fractionation where OM binds to minerals

is relatively heavier than freely existing OM. By comparing the distribution of individual amino acids containing various functional side chains, we can determine their selective accumulation patterns and turnovers related to biochemical processes influenced by biological succession as well as their physicochemical role in mineral associations. Generally, fungal contribution to SOM formation is expected to relatively increase during ecosystem development compared to bacteria; thus we expect to see that Eukarya derived amino acids accumulate throughout long-term development. Primary silicate minerals dominate these study sites and silicate minerals have permanent negative charges on the surfaces. Thus, it is expected, that a greater enrichment of positively charged amino acids in the mineral associated OM sub-pool will be found when compared to the whole soil OM pool.

Chapter 3. Seasonal and pedogenic effects on dynamics of soil organic nitrogen

Objectives 2-1: *To determine seasonal and pedogenic effects on the dynamics of proteinogenic amino acids in the whole soil OM pool and sub-pools associated with mineral particles, soil solution, and microbial biomass (Two factorial design: Season X Age).*

Objectives 2-2: *To determine seasonal and pedogenic effects on the dynamics of amino sugars and phospholipid fatty acids (PLFA) (microbial biomarker) in whole soil pool (Two factorial design: Season X Age).*

Hypothesis: 2-1: *Seasonal variations in biotic and abiotic factors affect the dynamics of proteinogenic amino acids and their association with whole soil, minerals, soil solutions and microbial biomass.*

Hypothesis: 2-2: *Seasonal variations of proteinogenic amino acids in soil solutions will be greater than those in whole and mineral associated pools.*

Hypothesis: 2-3: *Seasonal variations affect the dynamics of amino sugars and PLFA in whole soil pool.*

Amino acids have shown to turnover very rapidly in soil solution. This is thought to be due to the uptake competitions between microbes and plants and sorption to mineral surfaces. This work was done to further understand natural variations that occur as a result of turnover between microbes and the soil matrix during pedogenesis. This investigation was not intended to estimate turnover rates but rather their cycling among the soil constituents between seasons.

Chapter 4. Common trends in accrual of protein amino acids in two soil chronosequences: Lake Michigan, USA and Haast River, New Zealand

Objectives 3: *To compare proteinogenic amino acid distribution in whole soil and mineral associated pools between Michigan and Haast chronosequences.*

Hypothesis 3: *There will be common patterns of proteinogenic amino acids during pedogenesis in two independent ecosystems.*

A comparison of two geographically separate and climatically different chronosequences will test the veracity of the research results across the ecosystems and help better understand variations in the SOM accrual and turnover process.

References

Amelung W (2001) Methods using amino sugars as markers for microbial residues in soil. *Assessment methods for soil carbon*, 233-272.

Amelung W, Brodowski S, Sandhage-Hofmann A, Bol R (2008) Chapter 6 Combining Biomarker with Stable Isotope Analyses for Assessing the Transformation and Turnover of Soil Organic Matter. pp 155-250. San Diego, Elsevier Science & Technology.

Christias C, Couvaraki C, Georgopoulos S, Macris B, Vomvoyanni V (1975) Protein content and amino acid composition of certain fungi evaluated for microbial protein production. *Applied microbiology,* **29**, 250-254.

Ding X, Henrichs SM (2002) Adsorption and desorption of proteins and polyamino acids by clay minerals and marine sediments. *Marine Chemistry,* **77**, 225-237.

Espeland E, Wetzel R (2001) Complexation, stabilization, and UV photolysis of extracellular and surface-bound glucosidase and alkaline phosphatase: implications for biofilm microbiota. *Microbial Ecology,* **42**, 572-585.

Fan TM, Lane A, Chekmenev E, Wittebort R, Higashi R (2004) Synthesis and physico-chemical properties of peptides in soil humic substances. *The Journal of peptide research,* **63**, 253-264.

Farrell M, Hill PW, Farrar J, Bardgett RD, Jones DL (2011) Seasonal variation in soluble soil carbon and nitrogen across a grassland productivity gradient. *Soil Biology and Biochemistry,* **43**, 835-844.

Farrokhi N, Whitelegge JP, Brusslan JA (2008) Plant peptides and peptidomics. *Plant Biotechnology Journal,* **6**, 105-134.

Gleixner G (2013) Soil organic matter dynamics: a biological perspective derived from the use of compound-specific isotopes studies. *Ecological Research,* **28**, 683-695.

Gleixner G, Poirier N, Bol R, Balesdent J (2002) Molecular dynamics of organic matter in a cultivated soil. *Organic Geochemistry,* **33**, 357-366.

Grandy AS, Neff JC (2008) Molecular C dynamics downstream: The biochemical decomposition sequence and its impact on soil organic matter structure and function. *Science of the Total Environment,* **404**, 297-307.

Hamer U, Marschner B (2002) Priming effects of sugars, amino acids, organic acids and catechol on the mineralisation of lignin and peat. *Journal of Plant Nutrition and Soil Science,* **165**, 261.

Hamer U, Marschner B (2005) Priming effects in different soil types induced by fructose, alanine, oxalic acid and catechol additions. *Soil Biology and Biochemistry,* **37**, 445-454.

Harms H, Bosma T (1997) Mass transfer limitation of microbial growth and pollutant degradation. *Journal of industrial microbiology and biotechnology,* **18**, 97-105.

Hobara S, Osono T, Hirose D, Noro K, Hirota M, Benner R (2014) The roles of microorganisms in litter decomposition and soil formation. *Biogeochemistry,* **118**, 471-486.

Hulthe G, Hulth S, Hall PO (1998) Effect of oxygen on degradation rate of refractory and labile organic matter in continental margin sediments. *Geochimica et Cosmochimica Acta,* **62**, 1319-1328.

Kaiser K, Eusterhues K, Rumpel C, Guggenberger G, Kögel-Knabner I (2002) Stabilization of organic matter by soil minerals—investigations of density and particle-size fractions from two acid forest soils. *Journal of Plant Nutrition and Soil Science,* **165**, 451-459.

Kleber M (2010) What is recalcitrant soil organic matter? *Environmental Chemistry,* **7**, 320-332.

Kleber M, Sollins P, Sutton R (2007) A Conceptual Model of Organo-Mineral Interactions in Soils: Self-Assembly of Organic Molecular Fragments into Zonal Structures on Mineral Surfaces. *Biogeochemistry,* **85**, 9-24.

Knicker H (2004) Stabilization of N-compounds in soil and organic-matter-rich sediments—what is the difference? *Marine Chemistry,* **92**, 167-195.

Knicker H (2011) Soil organic N - An under-rated player for C sequestration in soils? *Soil Biology and Biochemistry,* **43**, 1118-1129.

Knicker H, Hatcher PG (1997) Survival of Protein in an Organic-Rich Sediment: Possible Protection by Encapsulation in Organic Matter. *Naturwissenschaften,* **84**, 231-234.

Kögel-Knabner I (2002) The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biology and Biochemistry,* **34**, 139-162.

Kramer C, Gleixner G (2006) Variable use of plant-and soil-derived carbon by microorganisms in agricultural soils. *Soil Biology and Biochemistry,* **38**, 3267-3278.

Krull ES, Baldock JA, Skjemstad JO (2003) Importance of mechanisms and processes of the stabilisation of soil organic matter for modelling carbon turnover. *Functional Plant Biology,* **30**, 207-222.

Marschner B, Brodowski S, Dreves A *et al.* (2008) How relevant is recalcitrance for the stabilization of organic matter in soils? *Journal of Plant Nutrition and Soil Science,* **171**, 91-110.

Mayer LM (1994) Relationships between mineral surfaces and organic carbon concentrations in soils and sediments. *Chemical Geology,* **114**, 347-363.

Melillo JM, Aber JD, Muratore JF (1982) Nitrogen and Lignin Control of Hardwood Leaf Litter Decomposition Dynamics. *Ecology,* **63**, 621-626.

Mikutta R, Kleber M, Torn MS, Jahn R (2006) Stabilization of Soil Organic Matter: Association with Minerals or Chemical Recalcitrance? *Biogeochemistry,* **77**, 25-56.

Milo R, Jorgensen P, Moran U, Weber G, Springer M (2010) BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Research,* **38**, D750-D753.

Miltner A, Bombach P, Schmidt-Brücken B, Kästner M (2012) SOM genesis: microbial biomass as a significant source. *Biogeochemistry,* **111**, 41-55.

Neidhardt FC, Ingraham JL, Schaechter M (1990) *Physiology of the bacterial cell: a molecular approach,* Sunderland, Mass, Sinauer Associates.

Nelson DL, Lehninger AL, Cox MM (2008) *Lehninger principles of biochemistry*, Macmillan.

Nguyen RT, Harvey HR (2001) Preservation of protein in marine systems: Hydrophobic and other noncovalent associations as major stabilizing forces. *Geochimica et Cosmochimica Acta,* **65**, 1467-1480.
Rillig MC, Caldwell BA, Wösten HaB, Sollins P (2007) Role of Proteins in Soil Carbon and Nitrogen Storage: Controls on Persistence. *Biogeochemistry,* **85**, 25-44.
Schimel JP, Bennett J (2004) Nitrogen Mineralization: Challenges of a Changing Paradigm. *Ecology,* **85**, 591-602.
Schmidt MWI, Torn MS, Abiven S *et al.* (2011) Persistence of soil organic matter as an ecosystem property. *Nature,* **478**, 49-56.
Sollins P, Swanston C, Kleber M *et al.* (2006) Organic C and N stabilization in a forest soil: Evidence from sequential density fractionation. *Soil Biology and Biochemistry,* **38**, 3313-3324.
Wershaw R, Pinckney D (1980) Isolation and characterization of clay-humic complexes. *Contaminants and sediments,* **2**, 207-219.
Wershaw RL (1986) A new model for humic materials and their interactions with hydrophobic organic chemicals in soil-water or sediment-water systems. *Journal of Contaminant Hydrology,* **1**, 29-45.
Wessels JG (1997) Hydrophobins: proteins that change the nature of the fungal surface. *Advances in Microbial Physiology,* **38**, 1-45.
Zang X, Van Heemst JDH, Dria KJ, Hatcher PG (2000) Encapsulation of protein in humic acid from a histosol as an explanation for the occurrence of organic nitrogen in soil and sediment. *Organic Geochemistry,* **31**, 679-695.
Zimmerman AR, Goyne KW, Chorover J, Komarneni S, Brantley SL (2004) Mineral mesopore effects on nitrogenous organic matter adsorption. *Organic Geochemistry,* **35**, 355-375.

Chapter 2. Selective accumulation of amino acids and proteins with minerals and association with plant-microbial communities

i. Authors: Jinyoung Moon[1], Li Ma[2], : Kang Xia[3], Mark A. Williams[1]

ii. Institute:
[1]Soil Microbial Ecology and Biogeochemistry Laboratory, Department of Horticulture, Virginia Polytechnic Institute and State University, 312 Latham Hall, 220 Ag Quad Ln., Blacksburg, VA 24061
[2]Department of Environmental Sciences, University of California, Riverside, CA 92521, USA and USDA-ARS, Soil Physics and Pesticides Research Unit, George E. Brown Jr. Salinity Laboratory, Riverside, CA 92507, USA.
[3]Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, 1880 Pratt Dr., Blacksburg, VA 24061

iii. Corresponding Author: Mark A. Williams, Phone: 540-231-2547, FAX 540-231-3083, Email: markwill@vt.edu

iv. Keywords: Lake Michigan Chronosequence, Sand dune, pedogenesis, soil organic matter (SOM), soil organic nitrogen (SON), soil protein, hydrolysable amino acid, organo-mineral associations, HPLC

v. Type of paper: Primary Research Articles

Title: Selective accumulation of amino acids and proteins with minerals and association with plant-microbial communities.

2.1. Abstract

The dynamics and persistence of proteinaceous compounds during pedogenesis are major mechanisms of soil formation and determinants of organic matter (OM) turnover. We investigated the accumulation patterns of proteinogenic amino acids associated with minerals dominated by permanent negative charges (primary silica minerals) and related these to the vegetative and belowground microbial successions during the soil ecosystem development. Positively-charged amino acids (arginine, lysine, histidine) showed clear patterns of accumulation, increasing ~65% during 4010 years of development, while negatively charged amino acids (glutamic acid, artic acid) decreased ~13%. In the mineral fraction, positively charged amino acids were approximately ~431% more-, while negatively charged amino acids were ~38% less-enriched. The belowground bacterial community based on a 16s ribosomal RNA phylogenetic analysis and the aboveground plant community predicted 71% (p<0.0001) and 66% (p<0.0001) of the amino acid dynamics, respectively, during soil ecosystem development. For example, Ala-rich Actinobacteria abundance declined with the year of development, concomitant with the Ala content in soil (r^2=0.82, p=0.0019). His-rich Acidobacteria and His in soil both increased with the year of development (r^2=0.92, p=0.0022). In support of the main hypothesis, the relative distribution of proteinogenic amino acids changed during pedogenesis; with evidence indicating that both microbes and minerals played a role as source and sink of soil organic matter (SOM), respectively. The close relationship in the dynamics of microbial and plant communities and the

process of pedogenesis, especially during early ecosystem development, suggested a tight linkage between biological communities and the formation and accrual of OM. Selective preservation of proteinaceous compounds also suggested that the properties of the soil sink also play an important role in determining how OM accrues in soil during ecosystem development.

2.2. Introduction

Amino acids, peptides, and proteins are the major form of nitrogen (N) in soil organisms and plants; for example, they comprise approximately 50% and 30% of the cellular weight of bacteria and fungi respectively (Christias *et al.*, 1975, Neidhardt *et al.*, 1990). These proteinaceous compounds compose a large fraction of soil organic matter (SOM; ~30%) (Knicker, 2011, Rillig *et al.*, 2007) and are a dominant form of total N (70-90%) in soil (Giagnoni *et al.*, 2010, Knicker & Hatcher, 1997, Miltner *et al.*, 2009, Nannipieri & Eldor, 2009, Schulten & Schnitzer, 1997). The compositions of biotic communities and their proteins, thus, are important determinants of SOM turnover and global biogeochemical cycles.

Cycling of proteinaceous compounds in soil will determine their relative distribution of bioavailable and long-term stabilized pools of N. The breakdown of soil peptides and proteins to amino acids is a primary rate limiting step for N mineralization (Jones & Kielland, 2002). Amino acids, peptides, and proteins thus play an important role in regulating available N for plants and pool sizes of organic matter in soil. The accrual of soil peptides/proteins through the turnover by microorganisms (Hobara *et al.*, 2014) and association with minerals (Mikutta *et al.*, 2006, Peng *et al.*, 2015) contribute to SOM formation and preservation. The role and stabilization of proteinaceous

compounds have been explained in models such as the molecular aggregates model (Wershaw, 1986), onion layering model (Sollins *et al.*, 2006), and encapsulation model (Knicker & Hatcher, 1997). In the models in part emphasizing amphiphilic and amphoteric functional groups, proteinaceous compounds interact with SOM and minerals and are thought to be less mobile and more protected from disassociation and decomposition. However, there is lack of evidence supported by the direct measurement on the process of protein accrual associated with soil minerals.

The emerging evidence of preferential accumulation and long residence time of protein-derived compounds in soil (Cotrufo *et al.*, 2013, Schmidt *et al.*, 2011) is counter to the traditional view that peptide bonds are highly labile to break down through heterotrophic activity (Alexander, 1981, Huguet *et al.*, 2008, Kokinos *et al.*, 1998, Schnitzer, 1985, Sollins *et al.*, 1996, Zonneveld *et al.*, 2010); and not a common form of stable soil organic matter. The foundational principles of ecosystem and soil carbon (C) cycling models were dominated by intrinsic molecular resistance as one of the major controllers of C turnover and storage. Accordingly, the molecular structure and lability of organic material has long been thought to determine long-term decomposition rates. However, recent observations show molecular structure is only part of the story. Protein and sugar compounds are more susceptible to chemical attack and biologically labile than aromatic ring structures, but their mean residence times rather tend to be longer (Schmidt *et al.*, 2011). The shift in thinking of the residence time of protein-derived molecules represents an important change for understanding global biogeochemical C and N cycling.

The mineralogical effects on the distribution of proteinaceous compounds were investigated in Hawaiian rainforest chronosequence (~4.1milion years of development), showing the important role of noncrystalline or short-range ordered minerals in the retention of compounds including acidic amino acids (artic acid and glutamic acid) in soil (Mikutta *et al.*, 2010). The research revealed that the portion of microbial-derived OM largely defines organo-mineral associations (Dümig *et al.*, 2012, Mikutta *et al.*, 2010). Although the mechanisms suggested based on the observation for this site may be generalizable to soils in humid environments where the concentration of noncrystalline minerals are high (Torn *et al.*, 1997), the mineralogical impact on the retention of proteinaceous compounds under other climate regions and/or with other parent materials is still uncertain.

In this study, we investigated the variation in the distribution of proteinaceous compounds by analyzing amino acids — the structure unit of peptides/proteins — along an eolian sand dune chronosequence (~4010 years of development; mineralogy is dominated by quartz) adjacent to Lake Michigan under a temperate-boreal climate. We hypothesized that the relative distribution of proteinogenic amino acids in soil may change; and the changes can be associated with both biotic and abiotic shifts across the chronosequence. The primary objective of this study is to determine if the distributions of amino acid in soil organic matter show patterns across the chronosequence and selective accumulation associated with minerals. To relate with biotic and abiotic factors, we investigated the correlation of amino acid profiling with successional shifts of vegetative and microbial communities and variations of edaphic conditions.

2.3. Materials and methods

2.3.1. Site descriptions and sampling

Soil chronosequences are a key tool for studying chemical, biological, and physical changes that occur in ecosystems as a consequence of pedogenesis. The study site consists of a series of beach-dune ridges bordering Lake Michigan (N 45.72729, W84.94076), and is located in the Wilderness State Park. The chronosequence of sediments have been derived from intermittent deposition of Lake Michigan for ~5000 years. The site is at the interface of temperate and boreal climate region. Temperature and precipitation averaged 6.28°C and 77.2 cm per year, respectively, between 1951 and 1980 at Mackinaw City, 15 km to the east.

The dune ridges have parent material originating from glacial deposits and Paleozoic bedrock underlying the lake basin. The parent material is assumed to be similar across the dune sequence. Fine sands deposited on the lake shore are dominated by quartz and contain other minerals in minor quantities (Lichter, 1995). The youngest soils (<100 y) are mapped as dunes which then develop into Deer Park sands (soil series) and described taxonomically as mixed, frigid, Spodic Udipsamments. The oldest soils (>1475 y) tend to be mapped to the Roscommon series, and are mixed, frigid Mollic Psammaquents.

Five replicates of top soil samples were collected from the incipient A-E horizon (0-15cm, 5-cm dia.) in nine dunes of age 105, 155, 210, 450, 845, 1475, 2385, 3210, and 4010y, using the same method in previous published literature (Williams *et al.*, 2013). Each replicate was separated by 10-m intervals across transects along each

dune's crest. The soil samples were stored in sterile Whirlpak bags, and frozen immediately in coolers with dry ice. Five replicates of sand samples were also collected along the beach to simulate the material that might be the source material that formed the eolian deposits of the dunes. Samples were collected in August 2008. The vegetation and soil properties have been characterized (Lichter, 2000, Williams *et al.*, 2013).

2.3.1.1. Aboveground vegetative succession

The change in plant community structure was greater during early compared to late ecosystem development. Generally speaking, dune-building grass species were replaced by evergreen shrubs and these were then replaced by mixed pine forests. This shift in early-succession to late-succession plant species happened at 450 years of soil and ecosystem development, when the early-succession species began to disappear and the mixed pine forest began to develop. Early succession was thus defined by considerable turnover of plant species. Indeed, plant community composition in the young dunes (105-155 y) was completely different from communities observed at 210 y, which were again taxonomically different from those >450 y of ecosystem development. Once the forest matured, the plant species composition stabilized and there was no major change in the plant community structure during late ecosystem development ($P = 0.59$) (Williams *et al.*, 2013).

2.3.1.2. Belowground bacterial succession

Bacterial communities showed patterns of change across the chronosequence during early ecosystem development (<845 y) but changed little during latter (845-4010 y) ecosystem development. The chronosequence gradients showed a number of

changes in phyla but were generally dominated by the abundance and dynamics of Acidobacteria, Actinobacteria, and Alphaproteobacteria, comprising 71% of all the sampled sequences. Other less abundant phyla (<4%) were Bacteroidetes, Cyanobacteria, Firmicutes, Planctomycetes, Betaproteobacteria, and Gammaproteobacteria. Between early (<450 y) and late (>450 y) ecosystem development, Acidobacteria increased approximately 6-fold from around 4% to w30%. Actinobacterial abundance declined, in contrast, from around 60 to w35% during this same time. The gradient of ecosystem development also was described by changes in low abundance taxa, with Bacteroidetes and Firmicutes, for example declining and Planctomycetes and Gammaproteobacteria increasing 4-fold. Cyanobacterial abundance declined from 5% to less than 0.5% following 210 y of ecosystem development (Williams *et al.*, 2013).

2.3.2. Whole soil hydrolysable amino acid analysis

The hydrolysable amino acids in the whole soil were acid digested, purified, and then analyzed using post-derivatization high performance liquid chromatography (HPLC). Two to five grams (dry weight) of moist soil was hydrolyzed in 10 ml of 6 M HCl with an internal standard (L-norvaline) at 110 °C for 24 h (Amelung & Zhang, 2001). After hydrolysis, the soil hydrolysates were centrifuged at 10,000 *Xg* for 10 min. The aliquot of the 400 µl supernatant was diluted in 55 ml ultra-pure water and cleaned on a preconditioned Dowex 50Wx8 resin (hydrogen form, 50-100 mesh; Alfa Aesar, Cat# B22109) (Küry & Keller, 1991, Norman & BOAS, 1953). The interfering metals were removed by rinsing with 0.1 M oxalic acid (pH 1.6-1.7). Amino acids retained on the resin were eluted with 30 ml 3M NH_4OH, filtered through a 0.22 µm polyvinylidene

fluoride (PVDF) membrane syringe filter, vacuum-dried, reconstituted in 10 µl 0.05 M HCl, and finally derivatized using the AccQ Fluor^{TM} reagent kit (Fluorescent 6-Aminoquinoly-N-Hydroxysuccinimidyl Carbamate derivatizing reagent; Waters Co. Cat# WAT052880) following the standard protocol from Bosch *et al.* (2006) and Hou *et al.* (2009). Chromatographic separations on the HPLC 1260 Infinity system (Agilent Technologies, USA) were carried out on a reversed phase column (Waters X-Terra MS C18, 3.5µm, 2.1X150mm). The mobile phase consisted of A: an aqueous solution containing 140 mM sodium acetate, 17 mM TEA, and 0.1% (g/L, w/v) EDTA-2Na (pH 5.05, adjusted with phosphoric acid solution) and B: ACN/water (60:40, v/v). The gradient conditions were 0 - 17 min 100 - 93% A, 17 - 21 min 93 - 90% A, 21 - 30 min 90 - 70% A, 30 - 35 min 70% A, 35 - 36 min 70 - 0% A, and 0% A for 4 min. The column was thermostated at 50 °C and operated at a flow rate of 0.35 ml/min. The sample injection volume was 5 µL. The analytes detection was carried out using a fluorescence detector (λ_{ex} = 250 nm and λ_{em} = 395 nm) (Bosch *et al.*, 2006, Hou *et al.*, 2009). Hydrolysable amino acids in the samples were qualified and quantified by comparison with amino acid standard solutions at different concentrations. Each amino acid standard solution contained 20 amino acids including alanine (Ala), arginine (Arg), aspartic acid (Asp), asparagine (Asn), cystine (Cys–Cys), glutamic acid (Glu), glutamine (Gln), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tyrosine (Tyr), tryptophane (Trp), and valine (Val). Because of the transformation of Asn to Asp and Gln to Glu and the destruction of Trp during acid hydrolysis, 17 amino acids except Asn, Gln, and Trp were quantified for hydrolysable proteinogenic amino acids.

Non-protein amino acid, ornithine (Orn) was also quantified as an indicator of bacterial contribution in soil.

2.3.3. Soil mineral associated amino acid analysis

Soil mineral associated fraction was isolated by the density gradient fractionation method (Kaiser & Guggenberger, 2007), followed by amino acid analysis in the mineral associated fraction (heavy fraction). Air-dried soils (2.5 g) were fractionated using sodium metatungstate (SMT, $H_2\ Na_6\ O_{40}\ W_{12}$) solutions with a density of 2.4 g/cm^3. The mixture was vigorously agitated on a shaker until the soil was completely dispersed. After the dispersion, the sample was centrifuged and the floating particulate (light fraction) was carefully separated from the heavy fraction. The heavy fraction was thoroughly cleaned with distilled water and completely dried at 60^{o}C in an oven overnight. The dried heavy fraction was weighed and hydrolyzed by using the same procedure with the whole soil hydrolysable amino acid analysis as described. The heavy fraction is referred to as mineral associated OM fraction.

2.3.4. N (1s) K-edge near edge X-ray adsorption fine structure (NEXAFS) analysis

The N (1s) K-edge NEXAFS spectrum for the SMT-isolated mineral portion of each soil sample was collected at room temperature under high vacuum (10^{-8}-10^{-9} Torr) on the soft X-ray beamline U4B at the National Synchrotron Light Sources, Brookhaven National Laboratory. Approximately a 1 mm layer of the SMT-isolated mineral portion of a soil sample was evenly spread on N-free carbon tape mounted on a sample holder before it was loaded into the vacuum spectrum collection chamber. The total electron yield was measured within the photon energy scan range of 390-440 eV. The L_{II} edge of Sc (403.9eV) was used for energy calibration. The Igor Pro data processing software

(Version 5.05A, WaveMetrics, Inc., Lake Oswego, OR) was used for N (1s) K-edge NEXAFS spectra averaging, averaged spectrum background subtraction and normalization. The processed N (1s) K-edge NEXAFS spectrum for each sample was then fitted, using the Solver function of Microsoft EXCEL®, with five Gaussians curves corresponding to N functional groups: Pyridine/Aromatic-N (400.2 eV), Pyridone/Aromatic-N (401.2 eV), Amide/Peptide-N (402.2 eV), Nitro-aromatic-N (405.8 eV), and Mineral fixed-NH_4^+ (407.0 eV) (Lehmann *et al.*, 2009). The abundance of Amide/Peptide-N relative to the total N was calculated based on the curve fitting result.

2.3.5. Statistics

For the multivariate comparison, molecular species of amino acid concentration were transformed by using the general relativization to remove the potentially strong influence of absolute abundance on distribution. Multi-Response Permutation Procedures (MRPP) and Nonmetric multidimensional scaling (NMS) ordination were performed using the PC-ORD software version 6.0 (MjM Software, Gleneden Beach, OR, USA) to compare the effect of soil age on the relative abundance (mol%) of 17 proteinogenic amino acids in whole soil and mineral associated OM hydrolysates. The cutoff of statistical significance in relative abundance data was p=0.01. Univariate comparisons were conducted by using One-way Analysis of Variance (ANOVA) and Student's t-test on the absolute abundance of amino acid, using SAS JMP pro11 (SAS Institute Inc., SAS Campus Drive, Cary, NC, USA). The cutoff of statistical significance in absolute abundance data was p=0.05. SigmaPlot version 11.0 (Systat Software, San José, CA, USA) was used to make graphs.

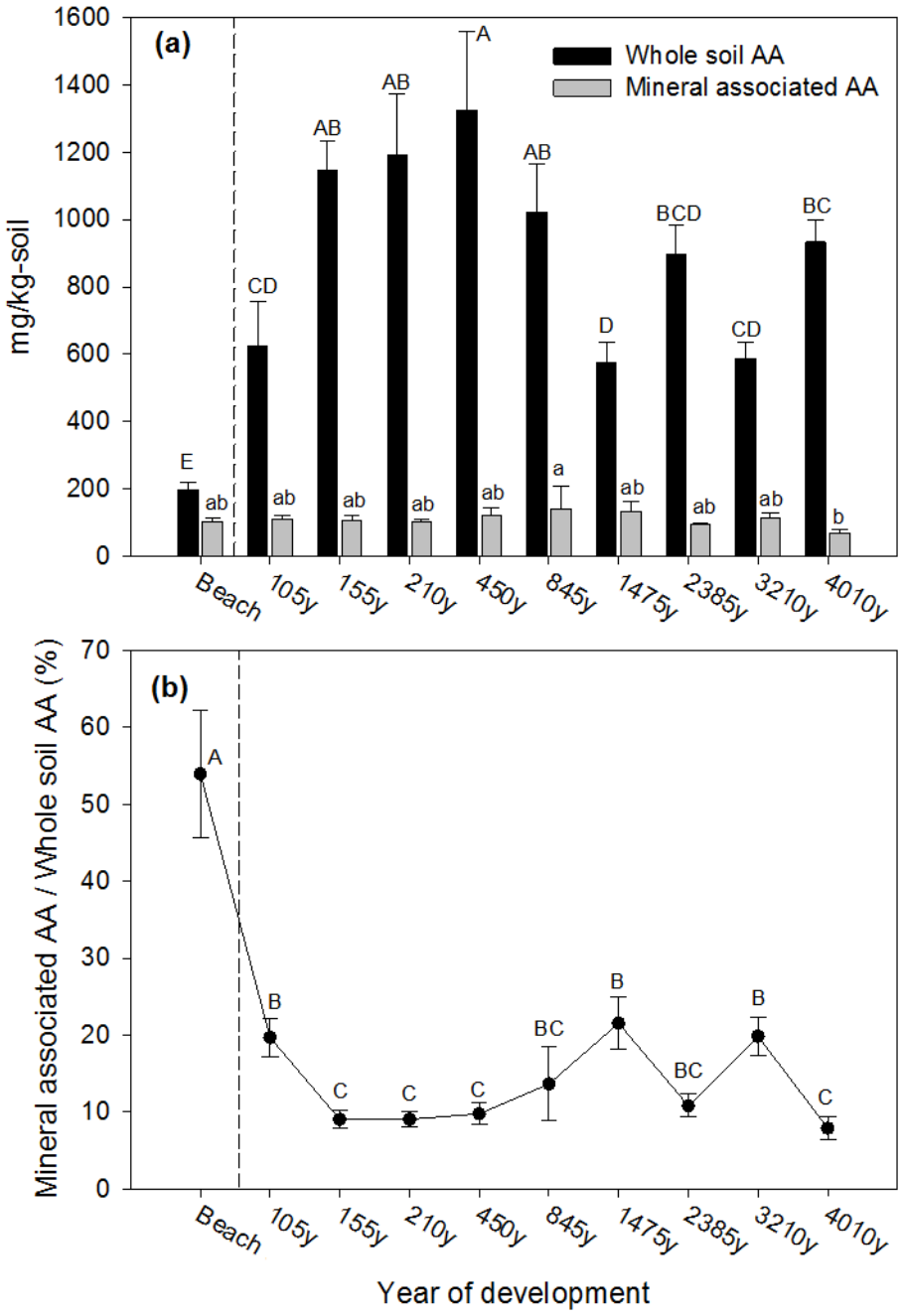


Figure.2. 1. Sum of 17 proteinogenic amino acids in the whole soil pool (whole soil AA) and mineral associated sub-pool (mineral associated AA) in mg/kg-soil (a), and the percentage of the mineral associated amino acid content over amino acid content of whole soil (b) with the age of sites across the Lake Michigan chronosequence. “Beach sand” represents the parent material of sand dunes without the influence of vegetation. Letters denote significant differences and the amino acid contents of two pools were separately tested by Student's t ($P< 0.05$) between the years of development: upper case=whole soil, lower case=mineral association. Error bars represent standard error (n=5).

2.4. Results

2.4.1. Abundance of amino acids

The amino acid abundance in the whole soil pool increased during the early years of soil ecosystem development (Fig.2. 1.a). The average amino acid content in the whole soil pool was 623 mg/kg-soil at 105y. Coinciding with vegetative colonization, the values peaked at 1,325 mg/kg-soil between 450-845y; thereafter, amino acid amounts declined, somewhat, but remained similar or greater than initial 105y pool sizes. Although the change of the amino acid content was dynamic in the whole soil pool, that in the mineral associated sub-pool was relatively consistent across the chronosequence (excluding beach sand) at 109±7 mg/kg-soil (Fig.2. 1.a), and accounting for 13±1% of the whole soil amino acid (Fig.2. 1.b). The total amino acid content of the whole OM pool in the beach sand without vegetation was significantly lower than those from the dunes with vegetation (Fig.2. 1.a). However, the total amino acid content of the mineral associated fraction in the beach sand was similar with chronosequence soils. It is notable that the lake derived beach sediment (sand) had significantly greater percentage of mineral associated amino acid (avg. 54%) compared to chronosequence soils (avg. 13%) (Fig.2.1.b). Overall, the results indicate a dynamic whole soil pool compared to a relatively stable mineral associated amino acid pool during ecosystem development.

2.4.2. Peptide-N in mineral associated fraction

Proteins and peptides were a dominant organic N form on the surface of mineral increasing from 35% at105y to 68% at 4010y (Fig. 2.2). The majority of the amino acids that we have determined in the mineral associated fraction are, thus, expected to be in

the form of peptides and proteins. Overall the contribution of peptides to mineral associated amino acid pool increased with pedogenesis. Peptide-N form among other proteinaceous compounds became relatively more abundant component of SOM that interact with minerals.

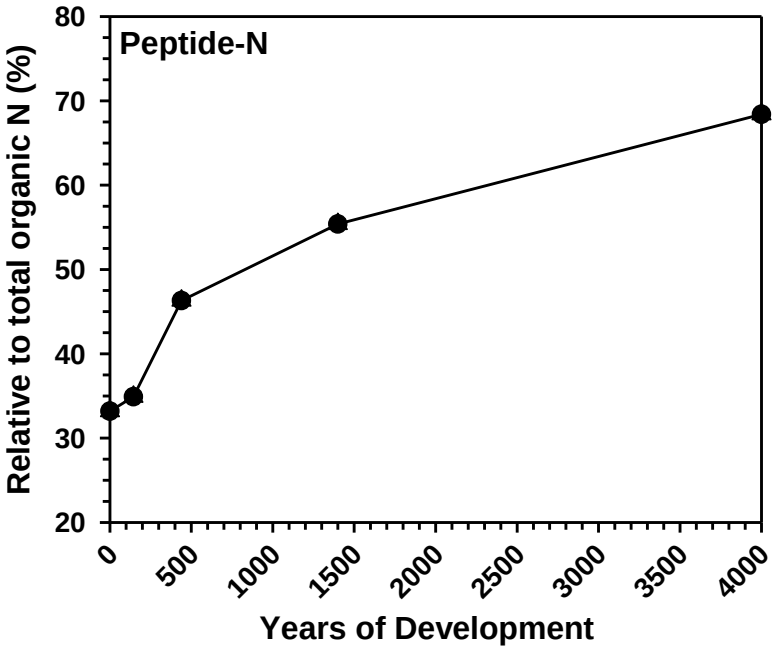


Figure.2. 2. Abundance of peptide-N relative to total N associated with the mineral portion of the Lake Michigan chronosequence soils at various ecosystem development stages (n=1). The relative abundance of amide/peptide-N was obtained using the synchrotron based N (1s) K-edge Near Edge X-ray Absorption Fine Structure (NEXAFS) spectroscopy. Beach sand is shown as “0” year.

2.4.3. Relative distribution of amino acids

Clear patterns of change in the relative distribution of amino acids with ecosystem development were shown in both whole soil and mineral associated pools ($p < 0.0001$ from MRPP for both, Fig.2. 3.a and b). For the whole soil samples, two shifts of relative distribution of amino acids were apparent: (1) from 105y to 450y, which is indicated with the solid blue arrow of “early development” in Fig.2. 3.a and (2) from 450y to 4010y, which is indicated with the dash blue arrow of “late development”. Positively

charged amino acids (His, Arg, Lys) and Pro were positively- and Gly, Ala, and Asp were negatively- correlated with age during early ecosystem development. Ser was positively- and Glu was negatively correlated with age during late ecosystem development. For the mineral associated fraction, the shift of relative distribution of amino acids was strongly associated with axis1, which is shown as blue arrow in Fig.2.3.b. Amino acid distributions at the pedogenically younger sites grouped to the left and gradually changed to the right along with the axis1 in Fig.2.3.b. Those at 4010y were relatively more distinct from the rest of chronosequence soils where Cys was positively correlated with 4010y. The relative abundances of Gly, Ala, Asx, Leu, and Ile were negatively correlated with axis1. The relative distribution of amino acid in the whole OM in beach sand was different from those in dunes with vegetation (Fig. 2.3.a and Appendix_Table A2.1.a; $p = 0.005$ or less from pairwise MRPP). Despite the distinct amino acid profiles of beach sand in the whole OM pool, the relative distribution of beach sand in the mineral associated amino acid was similar to those in younger dunes with vegetation (Fig.2. 3a and Appendix_Table A2 1b; $p = 0.477$ or less from pairwise MRPP). This may indicate relatively slower turnover of proteinaceous compounds associated with mineral surfaces compared those not retained to the mineral surfaces. Both their shifts of amino acid distribution in the whole soil pool and mineral associated sub-pool during ecosystem development were conspicuous; this, therefore, indicated important SOM composition change during pedogenesis.

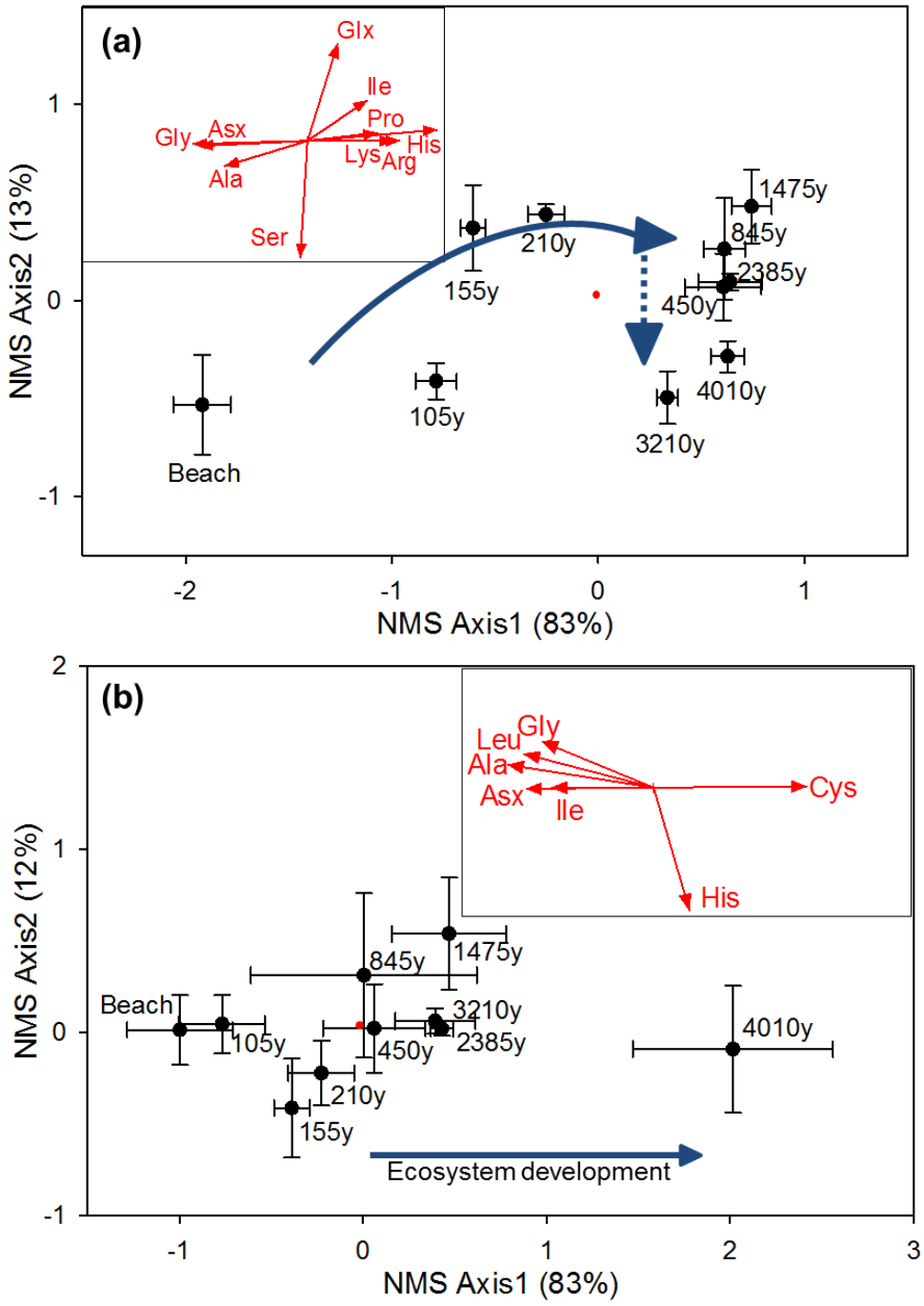


Figure.2. 3. Relationship between the distribution of 17 proteinogenic amino acids and soil ecosystem development plotted by Nonmetric multidimensional scaling (NMS) ordination in the whole soil **(a)**; and in the mineral associated fraction **(b)** in the Lake Michigan sand dune chronosequence. Freshly deposited "beach" sand was also sampled to assess the amino acid distribution of parent material expected to be similar to the source material that formed the eolian deposits of the dune soils. Error bars in **(a)** and **(b)** represent standard error (n=5). Percentages on each axis in each plot denote the amount of variability associated with each axis. Red vectors show the direction and strength of the relationship between individual amino acids and ordination scores with the cutoff of r^2=0.5 for **(a)** and **(b).**The Pearson and Kendall correlations of the vectors are provided in the supplementary document (Appendix_Table A2.3 and 2.4).

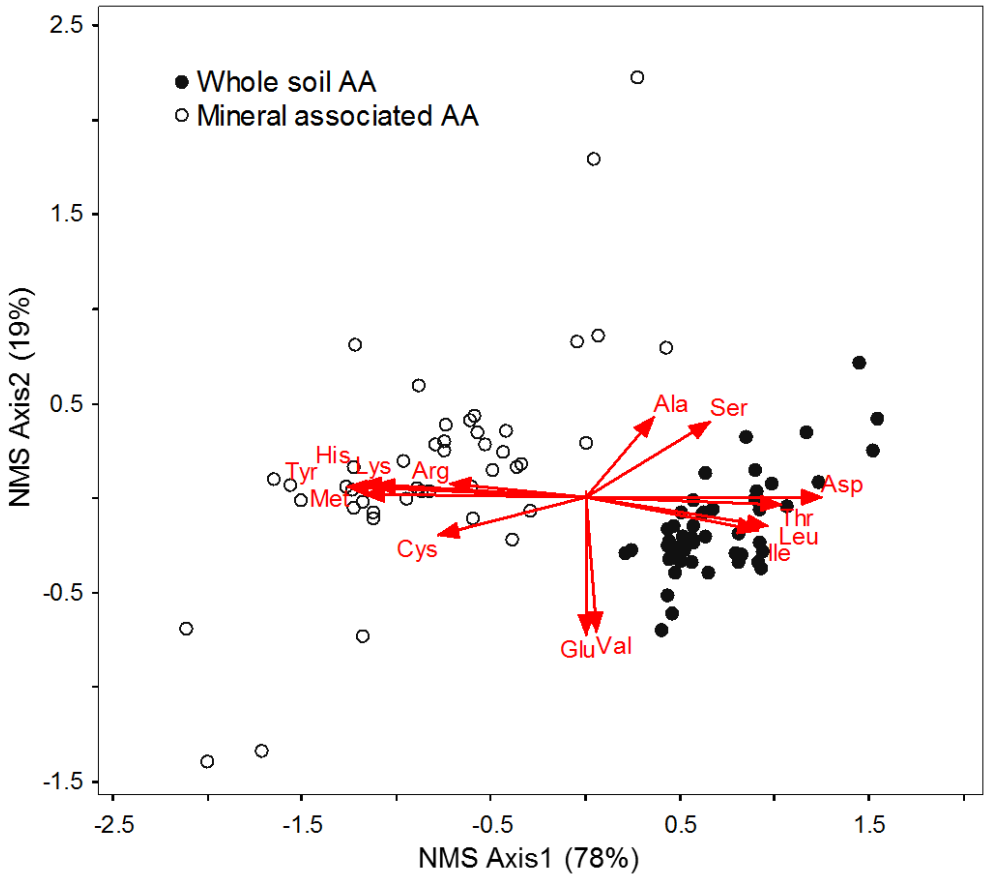


Figure.2. 4. Differences in amino acid distribution between whole soil and mineral associated fraction in the Lake Michigan sand dune chronosequence. Percentages on each axis in each plot denote the amount of variability associated with each axis. Red vectors show the direction and strength of the relationship between individual amino acids and ordination scores with the cutoff of r^2=0.3. The Pearson and Kendall correlations of the vectors are provided in the supplementary document (Appendix_Table A2.5).

2.4.4. Comparison between whole soil pool and mineral associated sub-pool

The dominant amino acids were Gly, Ala, Asx, Glx, Ser, Val, and Thr (Appendix_Fig.A2.1), but the relative distribution of amino acid was different between whole soil pool and mineral associated sub-pool (Fig. 2.4). The relative abundance of amino acids with a carboxyl functional group thus contributing to negative charges on the structure (termed negatively charged amino acids in this paper, including Asp and Glu) , those with the side chain of aliphatic group (Val, Leu, and Ile), and Thr which has

hydroxyl functional group, were relatively depleted in the mineral associated sub-pool than in the whole soil pool. Amino acids with the side chain of an amino functional group contributing to the positive charges on the structure (termed positively charged amino acids in this paper, including Arg, His, and Lys), those of a sulfur functional group (Cys and Met), and Tyr which has both aromatic and hydroxyl functional groups were enriched in the mineral associated sub-pool compared to those from the whole soil pool. The relative abundances of positively charged amino acids enriched in mineral associated fraction compared to those in the whole soil pool; for example, His was enriched ~431% in the mineral associated fraction (Fig .2. 5.a). On the other hand, the proportion of the negatively charged amino acids were depleted in the mineral associated fraction compared to those in the whole soil pool; for example, Asp was ~38% less in the mineral associated fraction than the whole soil pool (Fig.2. 5.a). The percentage of difference was calculated by

using $\%\ \mathrm{Difference} = \frac{(mol\%\ of\ mineral\ associate\mathrm{d\ AA}) - (mol\%\ of\ whole\ soil\ AA)}{(mol\%\ of\ whole\ soil\ AA)} \times 100\%$. The mean relative abundance of the positively charged amino acid group increased ~65% comparing the beach sand with the 4010y soil, while that with the negatively charged amino acid group decreased 13% during the same period of time (Fig .2. 5.b).

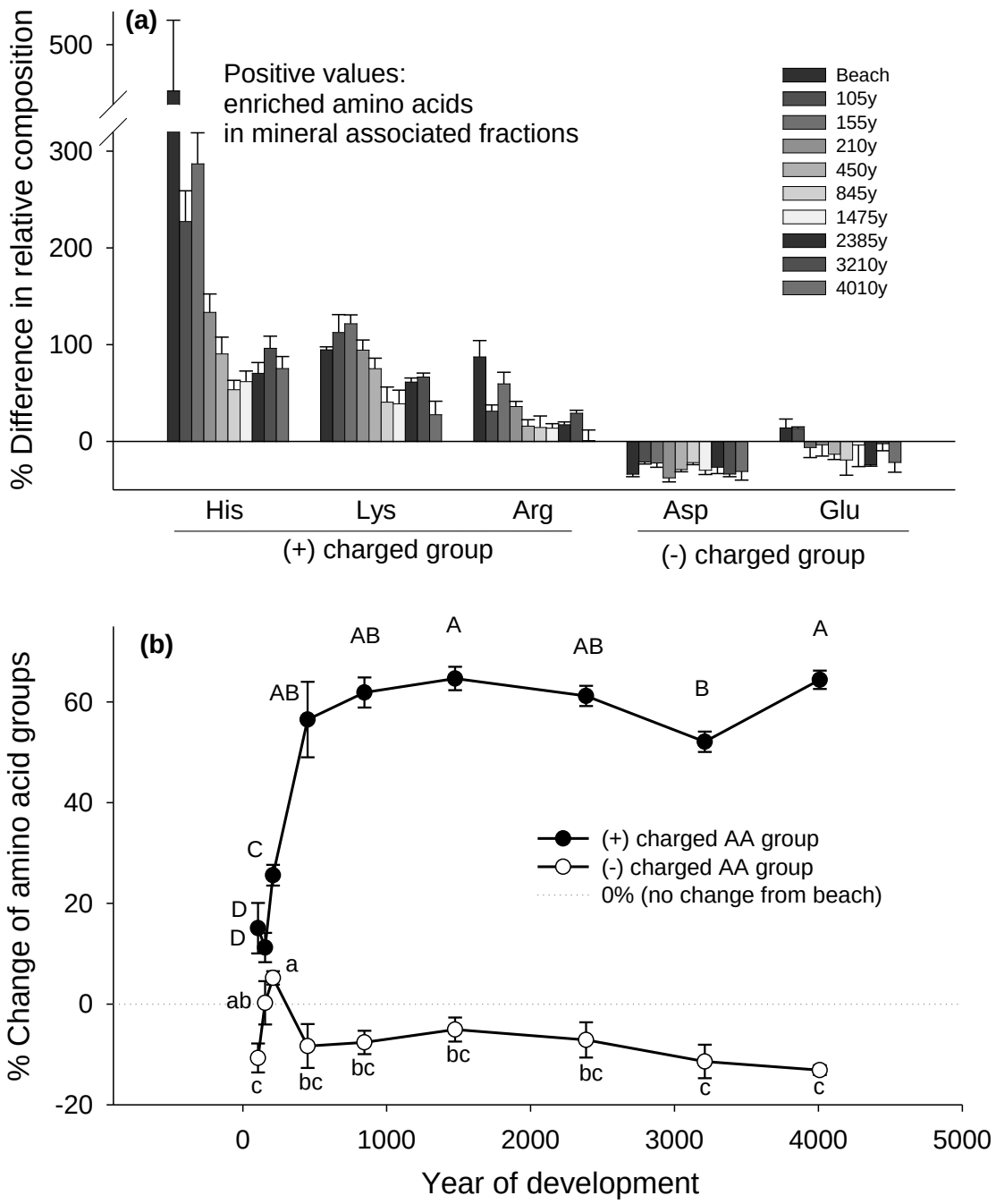


Figure.2. 5 Percentage of difference in relative abundance of charged amino acids between mineral associated sub-pool and whole soil pool **(a)**; and the percentage change of charged amino acid groups **(b)** during soil development across the Lake Michigan sand dune chronosequence. For **(a)**, the calculation was % Difference = ((mol% of mineral associated AA)-(mol% of whole soil AA))/((mol% of whole soil AA))×100%. For **(b)**, the initial abundance of amino acids (Y_0) is at the beach sand and the relative abundance (Y_i) at each year ($_i$). $Y=(Y_i-Y_0)/Y_0$ *100%. Letters denote significant difference and the amino acid contents of two pools were separately tested by Student's t ($P< 0.05$) between the years of development: upper case=whole soil, lower case=mineral association. Error bars represent standard error (n=5).

2.4.5. Relationship between amino acid dynamics and biotic and abiotic changes during pedogenesis

The change of amino acid distribution in the whole soil pool was highly correlated with both aboveground plant (r^2=0.66, p<0.0001) and belowground bacterial communities (r^2=0.71, p<0.0001) during ecosystem development (Fig. 2.5 and Appendix_Fig.A2.2). Dune-building grass species were replaced by evergreen shrubs between 155y and 210y, and these were then replaced by mixed pine forests at around 450y. Once the forest matured, the plant species composition stabilized and there was no major change in the plant community structure during late ecosystem development (Williams et al., 2013). Before and after the aboveground establishment of conifer forest at around 450y, belowground microbial community also showed the shift in composition. For example, Acidobacteria increased approximately 6-fold from around 4% to ~30%, while Actinobacterial abundance declined from around 60 to ~35% during this same time. The amino acid distribution as well as the plant and bacterial community compositions rapidly changed from 105y to 450y, but varied less for the next 3000 years.

Along with the change of biotic communities, the abiotic factors such as pH, cation content, and organic matter content changed. Soil Ca and Mg levels decreased in a log-linear pattern and were concurrent with declining pH (7.6-3.5) as soils aged from younger to older across the chronosequence. Soil organic matter and total soil organic C decreased along the chronosequence from younger to older soils ($r = 0.76$; $P < 0.05$). Soil Na (~149 mg/g) and P (~4 mg/g), in contrast, did not change with soil development (Lichter, 1998). The change of amino acid distribution was correlated with

pH (r^2=0.80), Mg (r^2=0.77), Ca (r^2=0.70), and K (r^2=0.61) content during the pedogenesis process (Appendix_Table A2.7).

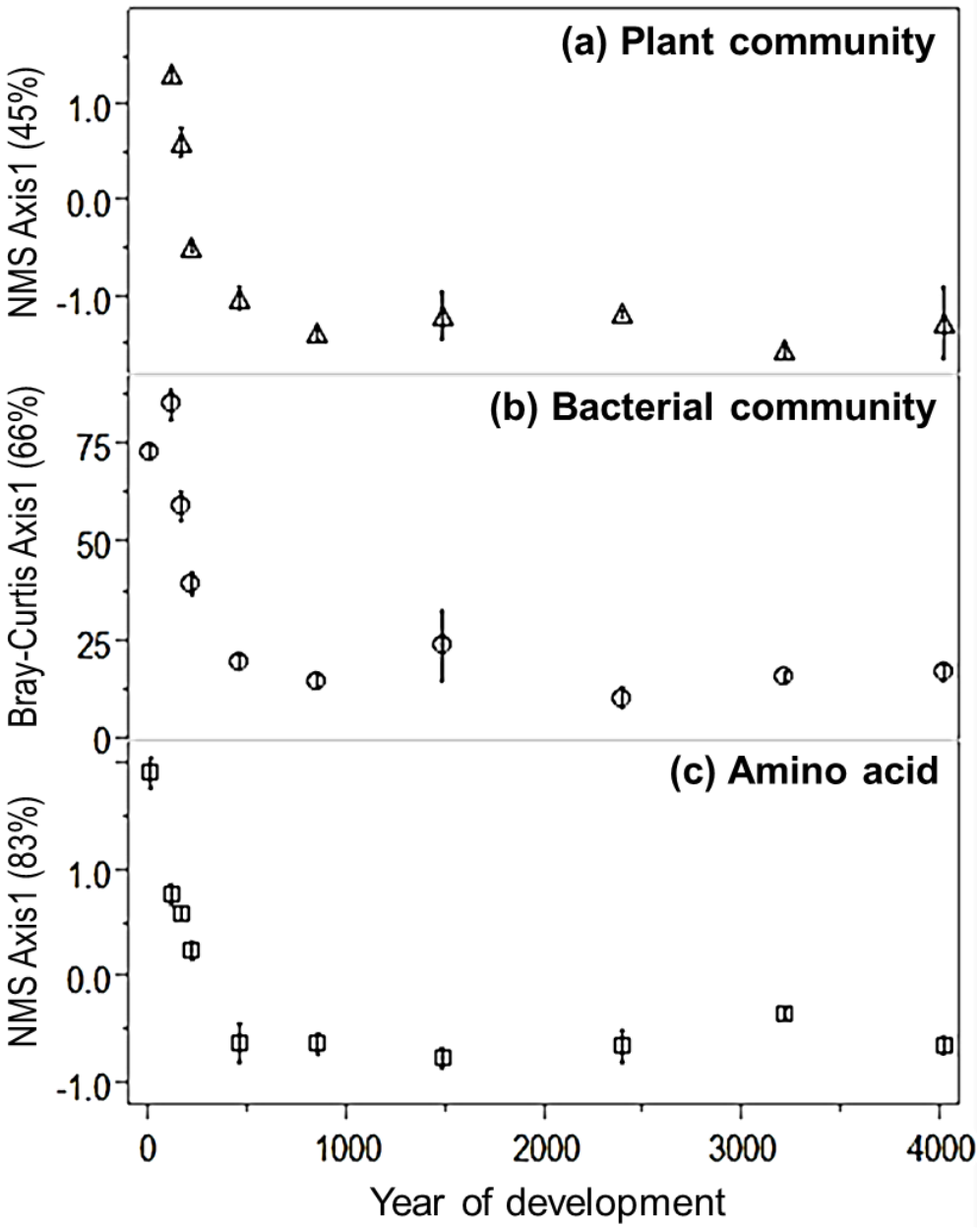


Figure.2. 6. The relationship between year of development and Axis1 from NMS ordination of plant community **(a)**; from Bray-Curtis ordination of bacterial community **(b)**; and NMS ordination of the relative distribution of 17 amino acids from the whole soil pool **(c)** in the Lake Michigan sand dune chronosequence. (a) and (b) were reconstructed based on Williams et al., 2013. Error bars represent standard error (n=5). The regression model graphs are provided in the supplementary document (Appendix_Fig.A2.2).

2.4.6. Water soluble amino acids from soil

Water soluble OM pool contributed to avg. 1%-amino acid in the whole soil OM pool (Appendix_Fig.B3.3.f). The extraction and analysis methods for water soluble amino acid were described in the section of 3.3.4. The abundance of amino acid in the soluble hydrolysate increased during early development and peaked at 845y at about 13 mg/kg-soil. It decreased to about 9 mg/kg-soil during late development after 845y (Appendix_Fig.B3.3.b). The change of amino acid distribution in the soluble hydrolysate was presented with age ($p < 0.0001$ from MRPP, Appendix_Fig.A2.1). Two shifts of relative distribution of amino acids were apparent: (1) from 105y to 1475y, which is indicated with the solid blue arrow of “early development” and (2) from 1475y to 4010y, which is indicated with the dash blue arrow of “late development”. Val, Leu, Ile, and Asx were relatively more abundant at beach sand and early stage of development (105y-210y). Gly, Lys, and Cys were more distributed at intermediate stage of development (450y-1475y). Lastly, Thr was positively correlated with late stage of development (2385y-4010y).

Soluble free amino acids (monomer) were accounted for approximately 0.16±0.01% to amino acids in the whole soil hydrolysate (Appendix_Fig.B3.3.g) and 14.4±0.7% to amino acids in the soluble hydrolysate. Soluble amino acids were predominated by polymer of amino acids (avg. 87.0±0.9%) rather than monomer of amino acids (Appendix_Fig.B3.1). The abundance of the soluble free amino acids was consistent across the chronoseqeunce (ANOVA, $p=0.2022$; Appendix_Fig.B3.3.c). Pedogenesis, however, was a strong driver to change the relative distribution of the soluble free amino acids (MRPP, $p<0.0001$; Fig.3.6). Gln+His (coeluted amino acids

during HPLC separation) were positively- and Phe, Val, Ile, Lys+Leu (coeluted amino acids), and Thr were negatively-correlated with year of development.

2.4.7. Comparison in pedogenic dynamics of amino acid among different OM pools

In comparison among different pools, there was lack of relationship of pattern of change in amino acid abundance of the soluble and mineral associated OM sub-pools with those of the whole soil OM pool during ecosystem development. The proportions of the amino acids in both sub-pools to those in the whole soil pool decreased during early ecosystem development (105-450y) and increased during late ecosystem development (845y<) (Appendix_Fig.B3.3.e and 3.3.f.), while abundance in the whole soil pool increased during early and decreased during late development.. The pedogenic dynamics of amino acid distribution were shown differently among the different OM pools as the amino acid vectors associated pedogenesis in the NMS bi-plots of three pools showed differently (Fig.2.3. and Appendix_Fig.A2.1). Asx, however, showed similar pedogenic trends across the three pools. The relative abundance of Asx decreased with year of development in the mineral associated and soluble OM sub-pools as well as whole soil OM pool. The relative abundance of Ile , in contrast, decreased with age in the mineral associated and soluble OM sub-pools, but increased with age in the whole soil OM pool.

2.4.8. Microbial derived amino acids and amino sugars

Soluble free (monomer) amino acids that were released by microbial lysis were determined and referred as microbial amino acid in this dissertation. The extraction and analysis methods for microbial amino acids were describes in the section 3.3.5. The

abundance and fraction size of microbial amino acid significantly decreased from early to late ecosystem development (Appendix_Fig.B3.3.d and 3.3.h; MRPP p=0.0022 for abundance and p=0.0008 for fraction size). The relative abundance of microbial amino acid changed dramatically from early development (105-155y) to late development (>210y) (MRPP, p<0.0001; Fig.3.6). Glu was strongly correlated with the early stage sites while Met was relatively more abundant at the late stage sites. It was notable that the relative distribution of microbial amino acid in beach sand was distinct from those of chronosequence soils (data not shown).

Ornithine (Orn) is non-protein amino acid and is often used for a bacterial biomarker because it occurs in bacterial peptidoglycan and in Orn-containing lipids (Lehninger, 1979, Ratledge & Wilkinson, 1988). The relative abundance of Orn compared to protein amino acids was consistent across the chronosequence (ANOVA, p=0.2194; Appendix_Fig.B3.6.d).

The abundance and distribution of amino sugars in whole soil hydrolysate changed significantly with the year of development (Two way-ANOVA p=0.0184 and Two way-MANOVA p=0.0002 respectively; Appendix_Fig.B3.2.b and Fig.3.7), mostly derived from change of glucosamine (GlcN) (Appendix_Fig.B3.4). Overall, the trend of change in amino sugar abundance with dune age mimicked the shift of the amino acid abundance in the whole soil hydrolysate described in the section 2.4.1 (r^2=0.48, p <0.0001), which may also reflect the change of total N content during the ecosystem development. The percentages of C as amino acid and amino sugar in the whole soil pool changed over year of development, ranged from 6 to 27% for amino acid-C and from 1 to 5% for amino sugar-C (Appendix_Fig.B3.5). The amino acid and amino sugar

in the whole soil pool generally accumulated during ecosystem development with greater accumulation in amino sugar than amino acid. Due to larger content of amino acid-C but similar level of amino sugar-C content in bacteria compared to fungi (Appendix_Fig.B3.5, Hobara *et al.* (2014)), the ratio of amino sugar-C to amino acid-C can reflect fungi to bacteria contribution to organic C content in soil. The ratio of amino sugar-C to amino acid-C increased during the earlier ecosystem development (Fig.3.14.a), suggesting the increase in fungal contribution to SOM during early ecosystem development (105y-450y).

In addition, the ratio GlcN to galactosamine (GalN) is often used for the indicator of fungal to bacterial contribution (Amelung, 2003, Joergensen & Wichern, 2008).The ratio of GlcN to muramic acid (MurA) is also indicator of fungal to bacterial contribution (Amelung, 2001). Both ratios increased for the first 1000 years of ecosystem development (Appendix_Fig.B3.6.b and c), also suggesting the increase in fungal contribution to SOM during this time period.

2.5. Discussion

Proteinogenic amino acids from whole- and mineral-associated soil organic matter were used as indicators of organic matter formation and change across a 105- to 4010-y dune-soil chronosequence. Distinctive shifts in these soil amino acids across the pedogenic gradient (Fig.2.3) supported the hypothesis that mineral, microbial, and plant communities each contribute to soil organic matter accrual. The types of amino acids found to change during pedogenesis, furthermore, support the individual role that mineral binding plays as a sink, and that organisms provide as organic matter sources. Overall, the close relationship in the dynamics of microbial and plant communities and

the process of pedogenesis, especially during early ecosystem development, suggest a tight linkage between these factors in the formation and accrual of soil organic matter.

2.5.1. Dominant amino acids in soil

Soil organisms and plants contribute to SOM formation through their biomass and physiological- and metabolic- products (Cotrufo *et al.*, 2013). As building blocks of the final products of genomic information, the dominant amino acids of organisms were predicted based on a genomic database (Chen *et al.*, 2013) and were confirmed to resemble those in soil pools (Friedel & Scheller, 2002, Werdin-Pfisterer *et al.*, 2009), For example, Gly, Ala, Asx, Glx, Ser, Val, and Leu, accounting for 70% of the total amino acids, were abundant in eukaryotic and bacterial cells (Chen *et al.*, 2013) as well as whole soil hydrolysis pools (Appendix_Fig.A2.1.a). The abundance of these common amino acids from soil organisms and plant debris might indicate that they are a major source of proteinaceous compounds, supporting the hypothesis that they are contributors to SOM accrual.

Although the amino acids of the whole soil hydrolysable pool share overall common dominant amino acids with their biological sources, there were other notable distinctions in the whole soil hydrolysates and evidence for greater contributions of specific amino types. Gly and Ala, for example, were about 81 and 29% greater in soil organic matter than the theoretical average protein of living organisms (Appendix_Fig.A2.1.e). This might be because these two are the most thermostable amino acids and contribute to greater persistence for millions of years (Wang *et al.*, 2012). Another specific contribution of these amino acids to SOM can be due to the abundance in the peptide interlayer bridges of peptidoglycan. S-layer proteins, for

example, of *Aeromonas hydrophylla* (AN: L37348) are 18% Ala. There is very high diversity of sequence among S-layer proteins, but evidence for some conservation across lineages has been described. These cell wall peptides and proteins play a key role in secretion (e.g. secretome) and signaling. Muropeptides from bacteria and cell wall glycoproteins of fungi are key communication pathways, and so the importance of these extracellular cell-wall attached peptides and proteins may help to explain their disproportionate contribution to SOM. In addition, Gly and Ala are the simplest and smallest amino acids. They are the major product of forms of the intercycling of tricarboxylic acid and amino acid metabolisms (Lodwig *et al.*, 2003, Nelson *et al.*, 2008). Therefore, the soil amino acid distribution may be affected by overall turnover and production from amino acid metabolisms carried out through microbial processes as well as by their structural stability in soil.

2.5.2. Amino acid shifts associated with microbial community change and pedogenesis

Based on our previous study on 16s ribosomal RNA phylogenetic bacterial community analysis (Williams *et al.*, 2013), shifts in the amino acid profile of the hydrolysable pool also show patterns that resemble the proteinogenic amino acid composition of bacterial groups during soil and ecosystem development (Fig.2. 6). The relative abundance of the dominant phylum Actinobacteria decreased dramatically from 60 to ~35% during early ecosystem development, which coincided with the decline in Ala and Gly during this stage of pedogenesis (relationship between Ala and Actinobacteria: r^2=0.82, p=0.0019 and between Gly and Actinobacteria: r^2=0.41, p=0.1670 respectively). Actinobacteria contain high guanine-cytosine (GC) content in

their genomes, which would help explain high abundances of Ala and Gly early and the declining levels by mid and latter periods of pedogenesis (see Appendix_Table A2.2). High GC content also is associated with low Lys and Phe in genomes (Chen et al., 2013), and in agreement with the initially low but increasing levels of hydrolysable soil Phe and Lys during pedogenesis.

Acidobacteria were the second largest bacterial phylum and based on amino acid coding in their genomes they produce relatively high amounts of His (Chen *et al.*, 2013). This again shows agreement with the 120 % increase in His when Acidobacteria become more dominant at the intermediate and later stage of soil development (R^2=0.94, p=0.0022). Similarly, the increases of His in water soluble free (monomer) amino acid pool along a young (105 to 450y) boreal alluvial forest successional sequence (Werdin-Pfisterer *et al.*, 2009) are supportive of our findings. It is likely that the 120% increase of this amino acid is related to shifts in the relative biomass of living Acidobacterial groups as well as the turnover of these organisms over the relatively long term periods of soil development. Since His is also found in greater amounts in eukaryotic organisms, such as plants and fungi (see Appendix_Fig.A2.2.d), their contributions to pedogenesis are likely also important. Overall, the results indicate that biological organisms can have strong and specific influences, related to their phylogeny and genomics, on the occurrence and accrual of organic matter in soil (Miltner *et al.*, 2009, Schmidt *et al.*, 2015).

Changes in vegetation can affect amino acid pools, for example, Glx and Asx, might be indicative of the influence of plant debris inputs to SOM. Glx and Asx are common constituents of plant xylem and phloem (Kielland, 1994) and storage amino

acids in plant tissues (Nordin & Näsholm, 1997). The decrease in the proportion of Glx and Asx with soil ecosystem development (Fig.2. 6.b) might be the result of depolymerization/proteolysis of these storage amino acids and the tendency for lower contributions of these amino acids from members of the Pinaceae (Hobara *et al.*, 2014), which dominate latter ecosystem developmental stages. Rapid turnover or preferential uptake of Gly as a source of plant and mycorrhizal fungi N (Geisseler *et al.*, 2009) might also result in a decrease in the relative proportions of this amino acid during 4010 years of successional change. However, the biochemical basis for differences in uptake rates among amino acids is unclear.

2.5.3. Mineral association and binding of amino acids

The proportion of mineral associated amino acids in the whole soil pool was relatively consistent and was low as avg. 13% (Fig.2. 1). This is in agreement with observations of other studies that have shown relatively lower adsorption of SOM on the mineral surfaces in sandy soils compared to the finer textured soils (Keil & Mayer, 2013, Mikutta *et al.*, 2007). However, the quantitatively consistent contribution of mineral associations to SOM suggests that the mineral may play a role in stabilization of SOM. The rate of weathering of tectosilicates e.g. quartz and feldspar, which are the most stable structure of primary silicate minerals, tends to be slow (McBride, 1994). Relatively slow mineralogical change of the dominant minerals and relatively younger development history of this site may not result in sufficient change in proteinaceous compound contents during pedogenesis. In addition, since rates of adsorption and desorption of proteinaceous compounds may maintain the balance and equilibrium, the variation in quantity of these compounds on the mineral surface has not been fully

described. The relative distribution of amino acids in mineral associated sub-pool over the long term, however, appear to shift in relationship to biogeochemistry related to pedogenesis and biological changes that occur during the ecosystem development (Fig.2.3.b).

The patterns of amino acid distribution change support the concept that mineral binding may play an important role in determining the amount and type of amino acid that accrue in soil organic matter during pedogenesis. Positively charged amino acids were preferentially accumulated on negatively charged exchange sites, while negatively charged amino acids were shown to decrease during pedogenesis. The decline of negatively charged amino acids was likely because they are weakly adsorbed, readily leachable from soil systems or more bioavailable for organisms to uptake. In the mineral associated fraction, furthermore, this result is consistent with enriched positively charged- and depleted negatively charged- amino acids through electrostatic forces. This may consequently reflect the increasing positively charged- and decreasing negatively charged- amino acids during pedogenesis, which has shown to be reasonable for the soil composed with mainly permanent negative surface charge generated from mineral isomorphic substitution.

The selective adsorption of positively charged amino acids to permanently negatively charged mineral surface and their limited release to a soil solution helps explain their preferential accumulation during 4010 years of pedogenesis. Due to the mineral specific mechanisms of stabilization of proteinaceous compounds, soils with different mineral compositions will show different amino acid profiles because of variable mineral surface chemistry. Three hundred to four million years of pedogenesis

in Hawaii, for example, showed preferential accumulation of negatively charged amino acids in the mineral associated fraction. This is likely because of the poorly crystalline and metal-hydroxide minerals that provide positive rather than negative exchange sites for organic matter accrual (Mikutta *et al.*, 2010, Strahm & Harrison, 2008). These results, though showing different binding characteristics, are mechanistically consistent with our study. The two observations tell us that the importance of amino acids that are associated with minerals varies depending on mineral composition of the sites.

During soil development, the pH of soil dropped from 7.6 at 105y to 3.5 at 4010y, which coincided with weathering and loss of total soil Ca and Mg from the ecosystem. The dissolution of minerals and leaching of cations may affect the adsorption strength of positively charged amino acids. Divalent cations such as Ca and Mg cations may tend to adsorb to cation exchange sites stronger than monovalent cations such as Na and K. Leaching of Ca and Mg, in contrast, may create opportunity for the replacement by positively charged amino acid on mineral exchange sites. In addition, multivalent cations such as Ca and Mg are responsible for creating multivalent cation bridging complex between negatively charged mineral surface and organic anions such as negatively charged amino acids (McBride, 1994). The removal of Ca and Mg during weathering process may result in the disrupting and weakening of the bridging complex.

In addition, the chemical interaction between proteinaceous compounds and their surroundings may contribute to the SOM stabilization and selective accumulation patterns. Soil solution surrounding proteins, for example, may cleave the hydrophilic moiety on the outer surface of globular structures and unfold them though non-enzymatic deamination. Carboxyl side chains on the amino acid polymers are the hot

spots for such chemical degradation (Geiger & Clarke, 1987, Jaenicke, 2000). This may explain the relative abundance of Asp and Glu or negatively charged amino acids decreased in whole soil pool with time as a result of their contribution to relatively rapid proteolysis compared to outer surface composed by other hydrophilic amino acids.

2.6. Conclusions

The molecular mechanisms contributing to longer residence times of SOM in soil are fundamental to pedogenesis, soil organic matter accrual, and ecosystem development. There were distinctive shifts in soil amino acids across the pedogenic gradient, which supported the hypotheses that mineral, microbial, and plant communities each contribute to soil organic matter accrual. Biological organisms were shown to have a strong and specific influence, related to their phylogeny, on the occurrence and accrual of organic matter in soil. The patterns of amino acid change also support the concept that mineral binding may play an important role in determining the amount and type of amino acid and protein that accrue in soil organic matter during pedogenesis. Overall, a tight linkage between sink and source factors suggest that there are important non-random mechanisms that contribute to the formation and accrual of soil organic matter. These results provide a valid alternative model of soil organic matter formation and accrual that can develop beside current sink based mechanisms that limit decomposability (e.g. aromatic groups) and source based mechanisms (e.g. structurally complex phenylpropanoid structure of plant lignin) in support of a conceptual model as major drivers of organic matter residence times in soil.

References

Alexander M (1981) Biodegradation of chemicals of environmental concern. *Science,* **211**, 132-138.

Amelung W (2001) Methods using amino sugars as markers for microbial residues in soil. *Assessment methods for soil carbon*, 233-272.

Amelung W (2003) Nitrogen biomarkers and their fate in soil. *Journal of Plant Nutrition and Soil Science,* **166**, 677-686.

Amelung W, Zhang X (2001) Determination of amino acid enantiomers in soils. *Soil Biology and Biochemistry,* **33**, 553-562.

Bosch L, Alegría A, Farré R (2006) Application of the 6-aminoquinolyl-N-hydroxysccinimidyl carbamate (AQC) reagent to the RP-HPLC determination of amino acids in infant foods. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences,* **831**, 176-183.

Chen W, Shao Y, Chen F (2013) Evolution of complete proteomes: guanine-cytosine pressure, phylogeny and environmental influences blend the proteomic architecture. *BMC Evolutionary Biology,* **13**, 219.

Christias C, Couvaraki C, Georgopoulos S, Macris B, Vomvoyanni V (1975) Protein content and amino acid composition of certain fungi evaluated for microbial protein production. *Applied microbiology,* **29**, 250-254.

Cotrufo MF, Wallenstein MD, Boot CM, Denef K, Paul E (2013) The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Global Change Biology,* **19**, 988-995.

Dümig A, Häusler W, Steffens M, Kögel-Knabner I (2012) Clay fractions from a soil chronosequence after glacier retreat reveal the initial evolution of organo–mineral associations. *Geochimica et Cosmochimica Acta,* **85**, 1-18.

Friedel JK, Scheller E (2002) Composition of hydrolysable amino acids in soil organic matter and soil microbial biomass. *Soil Biology & Biochemistry,* **34**, 315-325.

Geiger T, Clarke S (1987) Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *Journal of Biological Chemistry,* **262**, 785-794.

Geisseler D, Horwath WR, Doane TA (2009) Significance of organic nitrogen uptake from plant residues by soil microorganisms as affected by carbon and nitrogen availability. *Soil Biology and Biochemistry,* **41**, 1281-1288.

Giagnoni L, Magherini F, Landi L *et al.* (2010) Extraction of microbial proteome from soil: potential and limitations assessed through a model study. *European Journal of Soil Science,* **62**, 74-81.

Hobara S, Osono T, Hirose D, Noro K, Hirota M, Benner R (2014) The roles of microorganisms in litter decomposition and soil formation. *Biogeochemistry,* **118**, 471-486.

Hou S, He H, Zhang X, Zhang W, Xie H (2009) Determination of soil amino acids by high performance liquid chromatography-electro spray ionization-mass spectrometry derivatized with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate. *Talanta,* **80**, 440-447.

Huguet C, De Lange GJ, Gustafsson Ö, Middelburg JJ, Sinninghe Damsté JS, Schouten S (2008) Selective preservation of soil organic matter in oxidized marine sediments (Madeira Abyssal Plain). *Geochimica et Cosmochimica Acta,* **72**, 6061-6068.

Jaenicke R (2000) Stability and stabilization of globular proteins in solution. *Journal of Biotechnology,* **79**, 193-203.

Joergensen RG, Wichern F (2008) Quantitative assessment of the fungal contribution to microbial tissue in soil. *Soil Biology and Biochemistry,* **40**, 2977-2991.

Jones DL, Kielland K (2002) Soil amino acid turnover dominates the nitrogen flux in permafrost-dominated taiga forest soils. *Soil Biology and Biochemistry,* **34**, 209-219.

Kaiser K, Guggenberger G (2007) Distribution of hydrous aluminium and iron over density fractions depends on organic matter load and ultrasonic dispersion. *Geoderma,* **140**, 140-146.

Keil RG, Mayer LM (2013) Mineral matrices and organic matter. *IN Treatise on Organic Geochemistry,* **12**, pp. 337-359.

Kielland K (1994) Amino Acid Absorption by Arctic Plants: Implications for Plant Nutrition and Nitrogen Cycling. *Ecology,* **75**, 2373-2383.

Knicker H (2011) Soil organic N - An under-rated player for C sequestration in soils? *Soil Biology and Biochemistry,* **43**, 1118-1129.

Knicker H, Hatcher PG (1997) Survival of Protein in an Organic-Rich Sediment: Possible Protection by Encapsulation in Organic Matter. *Naturwissenschaften,* **84**, 231-234.

Kokinos JP, Eglinton TI, Goñi MA, Boon JJ, Martoglio PA, Anderson DM (1998) Characterization of a highly resistant biomacromolecular material in the cell wall of a marine dinoflagellate resting cyst. *Organic Geochemistry,* **28**, 265-288.

Küry D, Keller U (1991) Trimethylsilyl-O-methyloxime derivatives for the measurement of [6, 6-2 H 2]-d-glucose-enriched plasma samples by gas chromatography—mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications,* **572**, 302-306.

Lehmann J, Solomon D, Brandes J *et al.* (2009) Synchrotron-based near-edge X-ray spectroscopy of natural organic matter in soils and sediments. *Biophysico-Chemical Processes Involving Natural Nonliving Organic Matter in Environmental Systems*, 729-781.

Lehninger AL (1979) Biochemistry, 1975. *Worth, New York*, 268-271.

Lichter J (1998) Rates of weathering and chemical depletion in soils across a chronosequence of Lake Michigan sand dunes. *Geoderma,* **85**, 255-282.

Lichter J (2000) Colonization constraints during primary succession on coastal Lake Michigan sand dunes. *Journal of Ecology,* **88**, 825-839.

Lichter JP (1995) Mechanisms of plant succession in coastal Lake Michigan sand dunes. Unpublished 9612986 Ph.D., University of Minnesota, Ann Arbor, 252-252 p. pp.

Lodwig EM, Hosie AH, Bourdès A *et al.* (2003) Amino-acid cycling drives nitrogen fixation in the legume–Rhizobium symbiosis. *Nature,* **422**, 722-726.

Mcbride MB (1994) *Environmental chemistry of soils,* New York, Oxford University Press.

Mikutta R, Kaiser K, Dörr N *et al.* (2010) Mineralogical impact on organic nitrogen across a long-term soil chronosequence (0.3–4100 kyr). *Geochimica et Cosmochimica Acta,* **74**, 2142-2164.
Mikutta R, Kleber M, Torn MS, Jahn R (2006) Stabilization of Soil Organic Matter: Association with Minerals or Chemical Recalcitrance? *Biogeochemistry,* **77**, 25-56.
Mikutta R, Mikutta C, Kalbitz K, Scheel T, Kaiser K, Jahn R (2007) Biodegradation of forest floor organic matter bound to minerals via different binding mechanisms. *Geochimica et Cosmochimica Acta,* **71**, 2569-2590.
Miltner A, Kindler R, Knicker H, Richnow H-H, Kästner M (2009) Fate of microbial biomass-derived amino acids in soil and their contribution to soil organic matter. *Organic Geochemistry,* **40**, 978-985.
Nannipieri P, Eldor P (2009) The chemical and functional characterization of soil N and its biotic components. *Soil Biology and Biochemistry,* **41**, 2357-2369.
Neidhardt FC, Ingraham JL, Schaechter M (1990) *Physiology of the bacterial cell: a molecular approach,* Sunderland, Mass, Sinauer Associates.
Nelson DL, Lehninger AL, Cox MM (2008) *Lehninger principles of biochemistry*, Macmillan.
Nordin A, Näsholm T (1997) Nitrogen Storage Forms in Nine Boreal Understorey Plant Species. *Oecologia,* **110**, 487-492.
Norman F, Boas F (1953) Method of determination of hexosamines in tissues. *J. biol. Chem*, 553-563.
Peng X, Yan X, Zhou H, Zhang Y, Sun H (2015) Assessing the contributions of sesquioxides and soil organic matter to aggregation in an Ultisol under long-term fertilization. *Soil and Tillage Research,* **146**, 89-98.
Ratledge C, Wilkinson S (1988) Fatty acids, related and derived lipids. *Microbial lipids,* **1**, 23-52.
Rillig MC, Caldwell BA, Wösten HaB, Sollins P (2007) Role of Proteins in Soil Carbon and Nitrogen Storage: Controls on Persistence. *Biogeochemistry,* **85**, 25-44.
Schmidt J, Schulz E, Michalzik B, Buscot F, Gutknecht JLM (2015) Carbon input and crop-related changes in microbial biomarker levels strongly affect the turnover and composition of soil organic carbon. *Soil Biology and Biochemistry,* **85**, 39-50.
Schmidt MWI, Torn MS, Abiven S *et al.* (2011) Persistence of soil organic matter as an ecosystem property. *Nature,* **478**, 49-56.
Schnitzer M (1985) Nature of nitrogen in humic substances.
Schulten HR, Schnitzer M (1997) The chemistry of soil organic nitrogen: a review. *Biology and Fertility of Soils,* **26**, 1-15.
Sollins P, Homann P, Caldwell BA (1996) Stabilization and destabilization of soil organic matter: mechanisms and controls. *Geoderma,* **74**, 65-105.
Sollins P, Swanston C, Kleber M *et al.* (2006) Organic C and N stabilization in a forest soil: Evidence from sequential density fractionation. *Soil Biology and Biochemistry,* **38**, 3313-3324.
Strahm BD, Harrison RB (2008) Controls on the sorption, desorption, and mineralization of low-molecular-weight organic acids in variable-charge soils.(FOREST, RANGE & WILDLAND SOILS). *Soil Science Society of America Journal,* **72**, 1653.

Torn MS, Trumbore SE, Chadwick OA, Vitousek PM, Hendricks DM (1997) Mineral control of soil organic carbon storage and turnover. *Nature,* **389**, 170-173.
Wang S-Y, Cappellini E, Zhang H-Y (2012) Why collagens best survived in fossils? Clues from amino acid thermal stability. *Biochemical and Biophysical Research Communications,* **422**, 5-7.
Werdin-Pfisterer NR, Kielland K, Boone RD (2009) Soil amino acid composition across a boreal forest successional sequence. *Soil Biology and Biochemistry,* **41**, 1210-1220.
Wershaw RL (1986) A new model for humic materials and their interactions with hydrophobic organic chemicals in soil-water or sediment-water systems. *Journal of Contaminant Hydrology,* **1**, 29-45.
Williams MA, Jangid K, Shanmugam SG, Whitman WB (2013) Bacterial communities in soil mimic patterns of vegetative succession and ecosystem climax but are resilient to change between seasons. *Soil Biology and Biochemistry,* **57**, 749.
Zonneveld KaF, Versteegh GJM, Kasten S *et al.* (2010) Selective preservation of organic matter in marine environments; processes and impact on the sedimentary record. *Biogeosciences,* **7**, 483-511.

Chapter 3. Seasonal dynamics of soil organic nitrogen across a boreal-temperate successional sequence

i. Authors: Jinyoung Moon[1], Kang Xia[2], Mark A. Williams[1]

ii. Institute:
[1]Soil Microbial Ecology and Biogeochemistry Laboratory, Department of Horticulture, Virginia Polytechnic Institute and State University, 312 Latham Hall, 220 Ag Quad Ln., Blacksburg, VA 24061
[2]Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, 1880 Pratt Dr., Blacksburg, VA 24061

iii. Corresponding Author: Mark A. Williams, Phone: 540-231-2547, FAX 540-231-3083, Email: markwill@vt.edu

iv. Keywords: Lake Michigan Chronosequence, pedogenesis, soil organic matter (SOM), soil organic nitrogen (SON), soil protein, hydrolysable amino acid, organo-mineral associations, microbial amino acid, soluble amino acid, HPLC

v. Type of paper: Primary Research Articles

Title: Seasonal dynamics of soil organic nitrogen across a boreal-temperate successional sequence

3.1. Abstract

The changing biotic and abiotic environment associated with the change of seasons underpins the dynamics of organic nitrogen (N). We previously investigated the accumulation patterns of dominant organic N molecular species (e.g., amino acid and amino sugar), and found significant changes in their relative distributions related to pedogenesis (soil ecosystem development). The distribution of proteinogenic amino acids of whole soil pool changed seasonally and these seasonal dynamics were independent of pedogenesis (PerMANOVA, p=0.0002). The seasonal variations of amino acid distribution in whole soil pool were accounted for 49% out of total 94% variation in NMS bi-plot, and those in mineral associated pool were accounted for 22% out of total 92% variation in NMS bi-plot. These seasonal variations were more dynamic than anticipated, regarding they were thought to be slow pool. This suggested dynamics and replenishment of proteinaceous compounds on mineral surfaces between seasons. The amino acid distributions were clearly clustered into three pools: whole soil, mineral associated, and soluble pools. Positively charged (histidine, arginine, and lysine), aromatic (phenylalanine and tyrosine), and sulfur containing (cysteine and methionine) amino acid groups were relatively enriched in the mineral associated fractions while some of neutral polar amino acids (glycine, glutamic acid+glutamine, and threonine) were enriched in the soluble fractions. This suggested that the interactions of amino acids with the mineral and soil solution provided selective partitioning for amino acids. The abundance and distribution of amino sugars were not affected by seasons.

Although relatively conservative microbial biomarkers such as amino sugars and amino acids were consistent in abundance between seasons, phospholipid fatty acids (PLFA), which is indicative of living microbes, were different in abundance between seasons. This suggested that changes in composition of soil organic matter between seasons were driven by living communities and their physiology. Overall, results support that seasonal changes play a significant role in soil organic matter formation and cycling of organic N.

3.2. Introduction

New paradigm of soil N cycle emphasizes that biodegradation of N-containing polymers through microbial extracellular enzymes is a key step to nutrient supply to plants and microorganisms rather than N mineralization (Schimel & Bennett, 2004). Organic N is especially significant in boreal regions where N mineralization and bulk organic matter decomposition are slow due to a cold, dry climate (Flanagan & Cleve, 1983, Van Cleve & Alexander, 1981). Cool temperate and boreal forests of Wilderness State Park, Lake Michigan chronosequence, thus, provided an ideal setting to study the organic N dynamics. The dynamics of organic N, especially soluble amino acid monomers, are shown to be rapid with seasons (Jämtgård *et al.*, 2010, Kielland, 1995, Weintraub & Schimel, 2005), largely due to physiology variations of plants and microbes in response to abiotic changes between seasons. It is, however, still unknown about the seasonal dynamics of transformation between monomer and polymer of organic N as well as partitioning mechanisms of organic N associated with soil constituents, such as mineral particles, soil water, and organic aggregates.

Mineral associated organic matter fraction is theoretically stable pool and radiocarbon dating on this pool indicates relatively slower turnover of decades to thousands of years than organic matters (OM) not retained to mineral particles (Kögel-Knabner *et al.*, 2008, Lützow *et al.*, 2006b) . The counter phenomena, nevertheless, were observed in topsoil mineral associated fractions, where continuous replenishment of organic matter occurs and interaction of OM with mineral surfaces rather dynamic (Mikutta *et al.*, 2010, Paul *et al.*, 2008). The previous study on amino acid dynamics has shown a clear pattern of change in amino acid distribution associated with mineral during 4000 year of development (Chapter 2). In this regards, we hypothesized that seasonal variations may influence displacement of amino acids on mineral associations along with shifts in the amino acid patterns in soluble pool and consequently in whole soil OM pool.

In addition to protection mechanism associated with the mineral matrix, the transformation of plant materials to microbial residues is evident to be critical process in SOM formation enhancing resistance to biodegradation (Cotrufo *et al.*, 2013, Kai *et al.*, 1973, Miltner *et al.*, 2012). Even though living microbial biomass contributes to ~2% of SOM, non-living microbial biomass, especially microbial cell wall debris, is rather slow in decomposition and significantly contributes to SOM formation (Miltner *et al.*, 2009). Amino sugars, the second dominant organic N, are structure units of fungal and bacterial cell wall, and often used as indicator of microbial biomarkers (Amelung, 2001, Hobara *et al.*, 2014). Bacterial community compositions were consistent between growing and dormant seasons in Michigan sand dune chronosequence based on 16s ribosomal RNA phylogenetic analysis, which represent total of living and non-living

bacterial community (Williams *et al.*, 2013). Due to slow turnover of amino sugars combined with the steady bacterial community composition between seasons, it is hypothesized that there will be relatively less variations in dynamics of amino sugar by seasonality compared to amino acid. The seasonal influence on biogeochemical process of amino sugars, however, is uncertain.

In this chapter, we focused on seasonal dynamics of organic N coupled with pedogenic gradient in cool temperate-boreal, sand dune ecosystem near Lake Michigan. In order to understand displacement of proteinaceous compounds among soil constituents, we determined seasonal effect on the shift patterns of proteinogenic amino acid associated with mineral associated, soluble, whole soil OM pools in the chronosequence. We, furthermore, investigated the seasonal dynamics of microbial biomass and biomarkers across the chronosequence, expecting to find variations in contribution of microbial groups to SOM formation.

3.3. Materials and methods

3.3.1. Site descriptions and sampling

In Wilderness State Park, Lake Michigan, the soil chronosequence of sediments have been derived from intermittent deposition of Lake Michigan for 4000 years. Estimation of ages of a series of beach dune ridges was conducted by using accelerator mass spectrometry radiocarbon dating technique (Lichter, 1995). Five replicates of top soil samples were collected from the incipient A-E horizon (0-15cm, 5-cm dia.) in nine dunes of age 105, 155, 210, 450, 845, 1475, 2385, 3210, and 4010 years by the same way as previous published literature (Williams *et al.*, 2013). Each replicate was

separated by 10-m intervals across transects along each dune's crest. The soil samples were stored in sterile Whirlpak bags, and frozen immediately in coolers with dry ice and kept in -20°C until the analysis. Samples were collected in August (summer) and December (winter) of 2008 to determine the seasonal effects. The average highest and lowest temperatures are 23.9°C and 15.0°C in August 2008, respectively, and -0.6°C and -6.1°C in December 2008. The precipitation is 6.9cm in August 2008 and 4.6cm in December 2008. The vegetation and soil properties have been characterized (Lichter, 2000, Williams *et al.*, 2013).

3.3.2. Whole soil hydrolysable amino acid analysis

The hydrolysable amino acids in the whole soil were acid digested, purified, and then analyzed using post-derivatization high performance liquid chromatography (HPLC). Two to five grams (dry weight) of moist soil was hydrolyzed in 10 ml of 6 M HCl with an internal standard (L-norvaline) at 110 °C for 24 h (Amelung & Zhang, 2001). After hydrolysis, the soil hydrolysates were centrifuged at 10,000 *Xg* for 10 min. The aliquot of the 400 µl supernatant was diluted in 55 ml ultra-pure water and cleaned on a preconditioned Dowex 50Wx8 resin (hydrogen form, 50-100 mesh; Alfa Aesar, Cat# B22109) (Küry & Keller, 1991, Norman & BOAS, 1953). The interfering metals were removed by rinsing with 0.1 M oxalic acid (pH 1.6-1.7). Amino acids retained on the resin were eluted with 30 ml 3M NH_4OH, filtered through a 0.22 µm polyvinylidene fluoride (PVDF) membrane syringe filter, vacuum-dried, reconstituted in 10 µl 0.05 M HCl, and finally derivatized using the AccQ Fluor^{TM} reagent kit (Fluorescent 6-Aminoquinoly-N-Hydroxysuccinimidyl Carbamate derivatizing reagent; Waters Co. Cat# WAT052880) following the standard protocol from Bosch *et al.* (2006) and Hou *et al.*

(2009). Chromatographic separations on the HPLC 1260 Infinity system (Agilent Technologies, USA) were carried out on a reversed phase column (Waters X-Terra MS C18, 3.5µm, 2.1X150mm). The mobile phase consisted of A: an aqueous solution containing 140 mM sodium acetate, 17 mM TEA, and 0.1% (g/L, w/v) EDTA-2Na (pH 5.05, adjusted with phosphoric acid solution) and B: ACN/water (60:40, v/v). The gradient conditions were 0 - 17 min 100 - 93% A, 17 - 21 min 93 - 90% A, 21 - 30 min 90 - 70% A, 30 - 35 min 70% A, 35 - 36 min 70 - 0% A, and 0% A for 4 min. The column was thermostated at 50 °C and operated at a flow rate of 0.35 ml/min. The sample injection volume was 5 µL. The analytes detection was carried out using a fluorescence detector (λ_{ex} = 250 nm and λ_{em} = 395 nm) (Bosch *et al.*, 2006, Hou *et al.*, 2009). Hydrolysable amino acids in the samples were qualified and quantified by comparison with amino acid standard solutions at different concentrations. Each amino acid standard solution contained 20 amino acids including alanine (Ala), arginine (Arg), aspartic acid (Asp), asparagine (Asn), cystine (Cys–Cys), glutamic acid (Glu), glutamine (Gln), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tyrosine (Tyr), tryptophane (Trp), and valine (Val). Because of the transformation of Asn to Asp and Gln to Glu and the destruction of Trp during acid hydrolysis, 17 amino acids except Asn, Gln, and Trp were quantified for hydrolysable proteinogenic amino acids. Non-protein amino acid, ornithine (Orn) was also quantified as an indicator of bacterial contribution in soil.

3.3.3. Soil mineral associated amino acid analysis

Soil mineral associated fraction was isolated by the density gradient fractionation method (Kaiser & Guggenberger, 2007), followed by amino acid analysis in the mineral associated fraction (heavy fraction). Air-dried soils (2.5 g) were fractionated using sodium metatungstate (SMT, $H_2\ Na_6\ O_{40}\ W_{12}$) solutions with a density of 2.4 g/cm^3. The mixture was vigorously agitated on a shaker until the soil was completely dispersed. After the dispersion, the sample was centrifuged and the floating particulate (light fraction) was carefully separated from the heavy fraction. The heavy fraction was thoroughly cleaned with distilled water and completely dried at 60^{o}C in an oven overnight. The dried heavy fraction was weighed and hydrolyzed by using the same procedure with the whole soil hydrolysable amino acid analysis as described. The heavy fraction is referred to as mineral associated OM fraction.

3.3.4. Soil water soluble amino acids analysis

Five grams (dry weight) of moist soils was placed into 50ml sterile polypropylene centrifuge tube. Five ml of 20 mM KCl containing 20 mM NaN_3 and internal standard, α-Aminobutyric acid (AABA), was added into the centrifuge tube. Distilled water was added to make the final volume of 10 ml and concentration of 10 mM KCl. The mixture then was shaken gently on a reciprocal shaker at room temperature for 15min, followed by centrifugation at room temperature for 15 min at 4000 *Xg*. After centrifugation, the supernatant was collected and filtered through 0.22 µm PVDF membrane syringe filter. An aliquot of 500 µL of the filtrate was taken for centrifugal vacuum drying. (Werdin-Pfisterer *et al.*, 2009). The dry aliquot was subjected to derivatization procedure followed by amino acid analysis using HPLC, previously described in section 3.3.2. This

fraction which did not undergo the hydrolysis procedure is referred as soluble free (monomer) amino acid pool. An exception is noted for Cys-cys (cysteine dimer). Cysteine (Cys) was only detected in form of dimer due to the formation of disulfide bond between two Cys under oxidized conditions. Twenty amino acids were quantified for soluble free amino acids. During HPLC separation, Lys and Leu were coeluted (Lys+Leu), Asn and Ser (Asn+Ser), and Gln and His (Gln+His).

Soluble hydrolysable amino acids were operationally defined as amino acids that were released by acid-hydrolysis from water soluble extract, including not only monomer amino acids but also soluble peptides and proteins. Hydrolysable amino acids also included amino acids released from mixed compounds such as peptidoglycans. In order to extract soluble hydrolysable amino acids, another aliquot of the filtrate above was evaporated, hydrolyzed, and derivatized as described at the section of 3.3.2. This fraction is referred to as soluble hydrolysable amino acid fraction or soluble hydrolysate.

3.3.5. Microbial (cytoplasmic) amino acid analysis

Five grams (dry weight) of moist soils was placed into 50ml sterile polypropylene centrifuge tube. Four milliliter of chloroform was added into the centrifuge tube and the tube was shook at 160 rpm rotatory shaker for 2 hours at room temperature. The background amino acids from the reaction between the polypropylene tube and chloroform were monitored and subtracted from the final concentration. Five milliliter of 20 mM KCl containing 20 mM NaN_3 and internal standard, α-Aminobutyric acid (AABA) was added into the centrifuge tube. Distilled water was added to make the final volume of 10 ml and concentration of 10 mM KCl. The mixture then was shaken gently on a reciprocal shaker at room temperature for 15 min, followed by centrifugation at room

temperature for 15 min at 4000 *Xg*. After centrifugation, the supernatant was collected and filtered through 0.22 µm PVDF membrane syringe filter. An aliquot of 500 µL of the filtrate was taken for centrifugal vacuum drying. The dry aliquot was subjected to derivatization procedure previously described at 3.3.2. This fraction is referred as microbial amino acid fraction. Twenty amino acids were quantified for microbial free amino acids.

3.3.6. Whole soil hydrolysable amino sugar analysis

The hydrolysable amino sugars of whole soil were determined by acid digestion, purification, derivatization and then analysis using gas chromatograph (GC) (Amelung, 2001). Two grams (dry weight) of moist soil was hydrolyzed in 4ml of 6 N HCl with an internal standard (N-methyl-D-glucamine) at 110°C for 24h (Zhang & Amelung, 1996). After hydrolysis, the soil hydrolysates were centrifuged at 2,000 *Xg* for 10min. After centrifugation, the supernatant was collected and filtered through 0.22µm PVDF membrane syringe filter. An aliquot of 500µl of the filtrate was dried completely at 45°C heating bath with gentle stream of N_2 gas. The dry pellet was reconstituted in 5ml of DI water and pH adjusted to 6.5-7 with 2 N NaOH. The precipitate was removed by centrifugation (2000 Xg for 10 min) and the supernatant was centrifugal vacuum dried at 45°C. The residue was dissolved with 3 ml of HPLC grade methanol and centrifuged (2000 Xg for 10 min) to remove the salts. The supernatant was transferred to a 10ml-deactivated glass vial, centrifugal vacuum dried at45^0C, and finally derivatized following the procedure of Aldononitrile acetate derivatization from Guerrant and Moss (1984). Briefly, 300 µl of the reagent contained 32 mg hydroxylamine hydrochloride ml^{-1} and 40 mg 4-(dimethylamino)pyridine ml^{-1} in pyridine-methanol (4: 1 v/v) was added to a vial

containing a dry sample. After capping the vial and shaking for a few seconds, the solution was heated for 30 min at 75-80°C (during heating, the vial was shaken once more). Then, the vial was cooled to room temperature, and 1 ml of acetic anhydride was added. The vial was closed, shaken again, and reheated for 20 min. After cooling, 2 ml of dichloromethane was added. Excess derivatization reagents were removed by two washing steps. First, after 1 ml of 1 M HCl was added and strongly shaken for 30 sec the upper aqueous phase was removed. Secondly, in the same manner, the organic phase was extracted 3 times with distilled water (1 ml each). In the last washing step, the water was removed as completely as possible. The final organic phase was dried with gentle stream of N_2 gas at 60°C and finally, dissolved in 500 ul of ethyl acetate-hexane (1:1). Chromatographic separations on the GC 7890A system (Agilent Technologies, USA) equipped with flame ionization detector were carried out on a 25 m x 0.2 mm ID (0.33 pm) fused silica column (Ultra-2) with split ratio of 30:1. N_2 gas was used as a carrier gas with the column head pressure at 110 kPa. The temperature program was set as follows: the initial column temperature of 120°C was held for 1 min and then temperature was increased at 10°C min^{-1} to 250°C for 2.5 min. Thereafter, the temperature was increased again at 20°C min^{-1} to 270”C, held for 2 min. The injector temperature was 250°C and the temperature of detector was 300°C. Hydrolysable amino sugars in the samples were qualified and quantified by comparison with amino sugar standard solutions at different concentrations. Each amino sugar standard solution contained the internal standard, N-methyl-D-Glucamine (MeGluC), and 4 amino sugars including glucosamine (GlcN), mannosamine (ManN), galactosamine (GalN), and muramic acid (MurA).

3.3.7. Phospholipid Fatty acid (PLFA) analysis

Total lipids were extracted according to the procedure of Ringelberg *et al.* (1997) as modified by Butler *et al.* (2003). Briefly, ~10 g (dry weight) of soil samples were transferred to sterilized 160 ml serum bottles. The soils were extracted overnight using a mixture of 50mM phosphate buffer (pH 7.1), chloroform and methanol (0.8:1:2). The samples were centrifuged the following day at 1000 rpm for 5 min and filtered using Whatman # 1 filter paper. The filtrate was added with 3 M NaCl solution and a pinch of Na_2SO_4 salt and allowed to stand for ~8 h for phase separation. The chloroform phase was collected into separate glass tubes and dried completely under gentle stream of N_2 gas. Dried lipids were re-dissolved and fractionated into neutral, glyco- and phospholipids using silicic acid bonded phase extraction columns (Supelco, cat. No. 505048). The neutral and glyco lipids were eluted using chloroform and acetone respectively. Phospholipids were finally eluted with methanol into separate tubes and completely dried under N_2. The dried phospholipids fraction was transformed into fatty acids methyl esters under alkaline conditions and extracted twice in 1:4 chloroformehexane solution. The chloroform-hexane mixture was completely evaporated under stream of N_2 gas and the residue was re-suspended in 500 ml of hexane for GC analysis. The fatty acids were quantified and detected on Agilent 6890 Series gas chromatograph (Santa Clara, CA) equipped with a flame ionization detector, an Ultra-2 column (19091B-102; 0.2 mm by 25 m). H_2 was the carrier gas at a column head pressure of 20 kPa, septum purge of 5 ml min^{-1}, a split ratio of 40:1, injection temperature of 300 ^{0}C, injection volume of 2 ml. The oven temperature ramps from 170 ^{0}C to 288 ^{0}C at 28 ^{0}C min-1 and the analysis time of each sample was 6 min. Peak

identification was carried out by the Microbial Identification System (MIDI, Inc.) following calibration with a standard mixture of 17 fatty acid methyl esters (1300A calibration mix). The PLFA's i15:0, a15:0, i16:0, a16:0, i17:0, a17:0 (gram-positive), 16:1ω9, 16:1 ω 7, 18:1 ω 7 and cy19:0 (gram-negative) were considered as bacterial biomarkers, 10:Me 16:0 and 10:Me 18:0 for actinomycetes and 18:1 ω 9 and 18:2 ω 6 as fungal biomarkers (Frostegård & Bååth, 1996, Liang *et al.*, 2008, Zhang *et al.*, 2005). The ratio of fungal to bacteria biomarker fatty acids were used to indicate change in the fungal to bacterial biomass ratio (Bossio *et al.*, 1998).

3.3.8. Statistics

For the multivariate comparison, molecular species of amino acid and amino sugar concentration were transformed by using the general relativization to remove the potentially strong influence of absolute abundance on distribution. Multi-Response Permutation Procedures (MRPP), Two way-factorial Permutation based Multivariate analysis of variance (PerMANOVA) and Nonmetric multidimensional scaling (NMS) ordination were performed using the PC-ORD software version 6.0 (MjM Software, Gleneden Beach, OR, USA) to compare the effect of season and soil age on the relative abundance (mol%) of 17 proteinogenic amino acids in whole soil OM, mineral associated OM, and soluble OM hydrolysates as well as soluble and microbial monomer extracts. Those analyses were performed for the relative abundance of amino sugar in whole soil hydrolysate as well. The cutoff of statistical significance in relative abundance data was p=0.01. Univariate comparisons were conducted by using two way-factorial Analysis of Variance (ANOVA) and Student's t-test on the absolute abundance of amino acid, amino sugar, and PLFA and using SAS JMP pro11 (SAS Institute Inc., SAS

Campus Drive, Cary, NC, USA). The cutoff of statistical significance in absolute abundance data was p=0.05. SigmaPlot version 11.0 (Systat Software, San José, CA, USA) was used to make graphs.

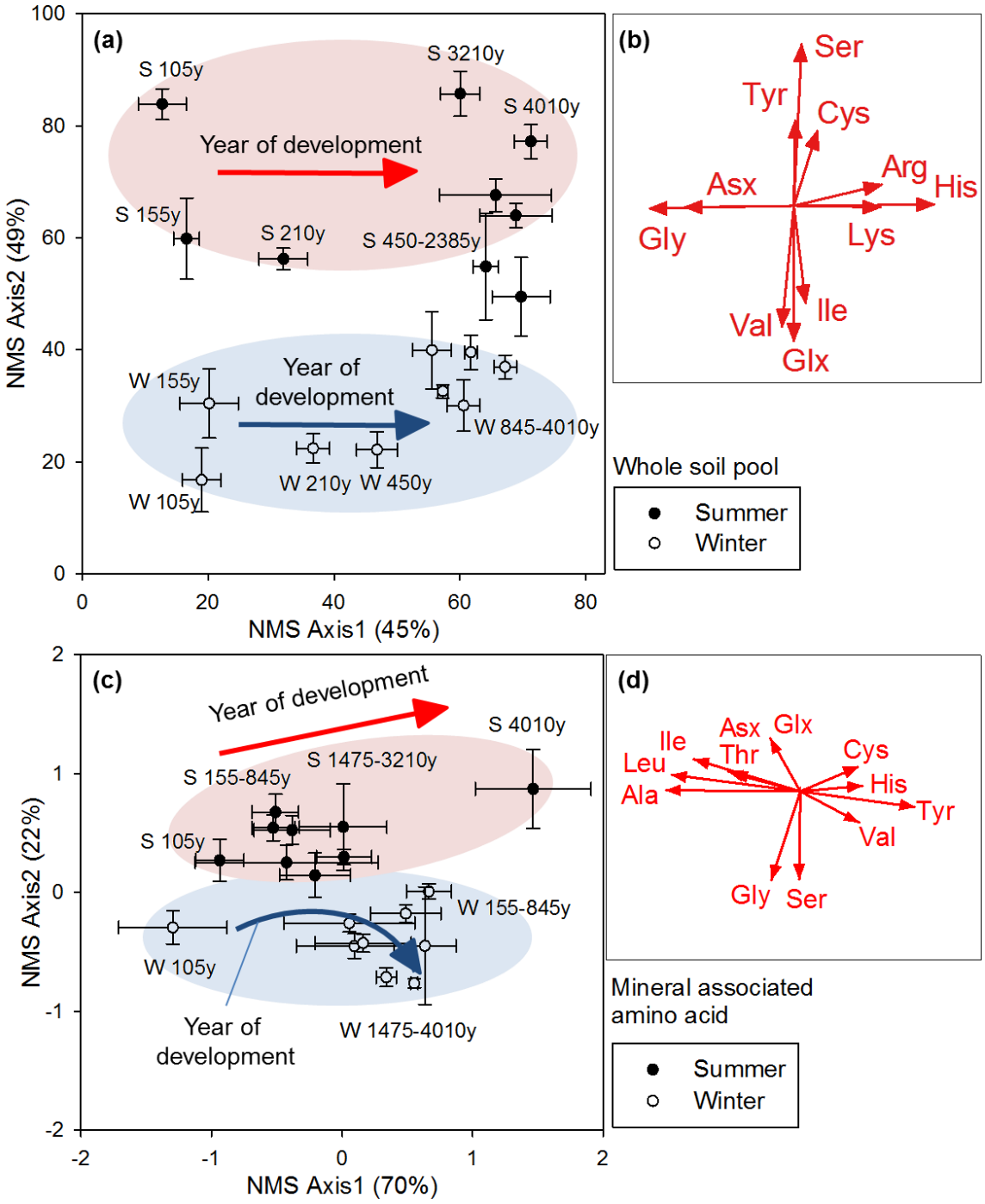


Figure.3. 1. Relative distribution of 17 proteinogenic amino acids from the whole soil pool **(a)** and mineral associated pool **(c)** between summer and winter during soil ecosystem development across Lake Michigan chronosequence, plotted by Nonmetric multidimensional scaling (NMS) ordination. Correlations of variables with ordination with r^2>0.3 were shown in bi-plot vector where length and direction represent the magnitude and directions of the correlation, respectively **(b)** and **(d)**. The distributions of whole soil pool were tested by Two way-PerMNOVA between summer and winter (p=0.0002); among site ages (p=0.0002); interaction term (p=0.0106). Due to unbalanced sample number, the distributions of mineral associated pool were tested by MRPP between summer and winter (p<0.0001); among site ages (p<0.0001). Error bars in **(a)** and **(c)** represent standard error (n=5 and n≤5 respectively). Percentages on each axis on **(a)** and **(c)** denote the amount of variability associated with each axis. The final stress for 2-d NMS was 11 and 14 for **(a)** and **(c)** respectively. The Pearson and Kendall correlations of the vectors are provided in appendix (Table B3.3 and 3.4)

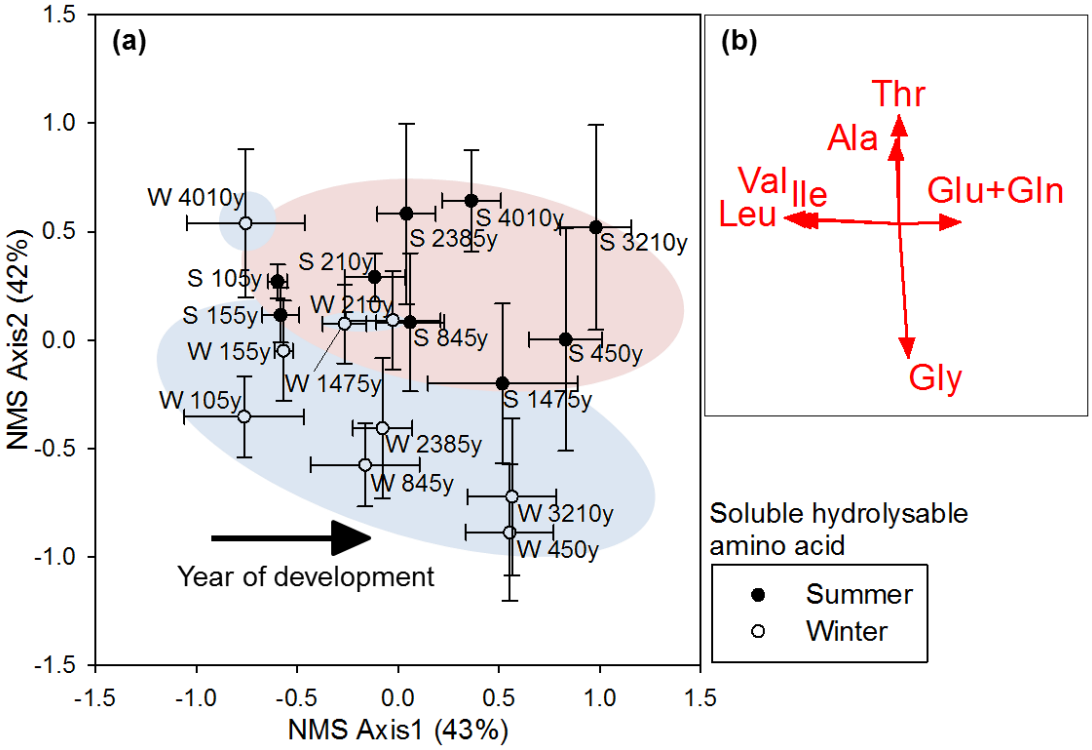


Figure.3. 2 Relative distribution of 17 proteinogenic amino acids from the hydrolysates in the soluble pool between summer and winter during soil ecosystem development across Lake Michigan chronosequence , plotted by Nonmetric multidimensional scaling (NMS) ordination **(a)**. Correlations of variables with ordination with r^2>0.3 were shown in bi-plot vector where length and direction represent the magnitude and directions of the correlation **(b)**. The distributions of whole soil pool were tested by Two way-PerMNOVA between summer and winter (p=0.0002); among site ages (p=0.0002); interaction term (p=0.0002). Error bars in **(a)** represent standard error (n=5 and n≤5 respectively). Percentages on each axis on **(a)** denote the amount of variability associated with each axis. The final stress for 2-d NMS was 19 for **(a)** The Pearson and Kendall correlations of the vectors are provided in appendix (Table B3.5)

3.4. Results

There were distinct seasonal shifts in the relative abundance of the molecular species with amino acid, accounting largely for 49% of NMS variation (axis2 in Fig.3.1 a) and having greater dynamics than amino sugar (Fig.3.7) in the whole soil hydrolysate. The pronounced dynamics in relative abundance of amino acid between summer and winter also were shown in the sub-pools associated with mineral, soil water, and microbes, accounting largely for 22%, 42%, and 51% respectively of NMS variations (axis2, Fig.3.1.b, Fig.3.2, and Fig.3.6.c). In contrast, monomer amino acid in the soluble

pool had no shift in pattern of distribution between seasons (Fig.3.6.a) although the relative abundances of soluble amino acid hydrolysates were different between seasons (Fig.3.2). The distribution of the mineral associated sub-pool were expected to be indicative of a longer timeframe and representing, on average, many decades or more of SOM formation. Although this has been shown to be true in many systems (Kögel-Knabner *et al.*, 2008, Lützow *et al.*, 2006b), the results suggested that a portion of the mineral associated pool was seasonally dynamic. The fraction, related with seasonal change, was thus unlikely to be part of the most stable SOM pools with slow turnover.

3.4.1. Amino acid in whole soil hydrolysable OM pool

The relative abundance (mol%) of amino acid in the whole soil pool was significantly different between seasons (Two way-PerMANOVA p=0.0002) as a clear separation between summer and winter by axis 2 accounting 49% of variation was shown in Fig.3.1.a. Ser, Tyr, and Cys had strong preferential distribution in soil collected in summer, while Glx, Val, and Ile were more abundant in soil collected in winter, indicating distinct patterns of amino acid distribution by seasonal influence.

With the year of development, the distribution of amino acid in the whole soil pool in both seasons dynamically changed (axis1 in Fig.3.1.a). Positively charged amino acids (His, Arg, Lys) and Pro were positively- and simplest alkyl amino acid (Gly and Ala), and Asx were negatively- correlated with developmental age (Appendix_Table B3.3). Despite of the distinct distributions of amino acid between seasons, it is notable that there were common trends between seasons in the change of amino acid distribution associated with the year of development. In other words, positively charged amino acids relatively increased with the year of development in summer so as in winter,

whereas Gly and Asx decreased with the year of development in both summer and winter. This might indicate that the seasonal and pedogenic influences on amino acid distribution could be somewhat separated with relatively weak interaction between two factors (Season*Age interaction by Two way-PerMANOVA p=0.0106 with the significance cutoff of 0.01). This thus suggested that the amino acids may comprise a biological or chemical fraction important to the process of soil C and N cycling.

3.4.2. Hydrolysable amino acid associated with mineral

The relative abundance (mol%) of amino acid in mineral associated OM fraction was affected by season accounting 22% of multivariate variation (Fig.3.1.c, MRPP $p<0.0001$). Glx was positively correlated with summer, while Ser and Gly were positively correlated with winter in the mineral associated fraction (Fig.3.1.c, Appendix_Table B3.4). It is notable that Glx in the mineral associated sub-pool was more abundant in summer, while Glx in whole soil pool was rather abundant in winter. Ser, on the contrary, exhibited the opposite trend to Glx between seasons in mineral associated sub-pool and whole soil pool.

The pedogenic variation was shown to be larger in amino acid distribution of mineral associated fraction, accounting 70% of variation in axis1 (Fig.3.1.c, MRPP $p<0.0001$) compared to seasonal variations (22% in axis2). Tyr, Cys, Met, His, and Val were positively- and Ala, Leu, Ile, Asx, and Thr were negatively- correlated with site age in both seasons in mineral associated fraction. There were somewhat common trends of change in amino acid distribution along the year of development. For example, the amino acids preferentially distributed in younger sites were consistently shown in summer and winter, but the amino acids positively correlated with developmental age

were not in common between summer and winter. The common traits between two seasons were not observed as strong as those in whole soil pool (Fig.3.1). It seemed to have some interaction between seasonal and pedogenic factors on amino acid distribution.

3.4.3. Hydrolysable amino acid dissolved in water

The seasonal variation was significantly great in the relative abundance (mol%) of amino acid, accounting 42% of NMS variation (Fig.3.2, Two way-PerMANOVA p=0.0002). There was clear separation between seasons in NMS bi-plot with the exception of 4010y-winter sites which were rather similar to 105y-summer site. Overall, Thr and Ala were more abundant in summer, while Gly were strongly correlated with winter in soluble fraction (Fig.3.2).

The pedogenic changes of amino acid distribution in hydrolysate from soluble fraction were presented with 43% of variation in Fig.3.2 (Two way-PerMANOVA p=0.0002). The pedogenic patterns of amino acid distribution in whole soil- and mineral associated- OM hydrolysates were correlated with axis 1 in NMS bi-plots throughout the year of development (Fig3.1). The correlations with axis 1 in NMS bi-plot of those in soluble OM hydrolysate, however, were partially exhibited during the early development (Fig3.2). There was strong correlation between amino acid distribution and axis1 in NMS bi-plot from 105y to 450y, but after 450y there was no clear pedogenic pattern of change with axis1. Plotted summer and winter together, Leu, Ile, and Val were strongly correlated with the sites younger than 210y, while Glx was relatively more abundant in older sites (>210y). No common trend between summer and winter was detected by pedogenic change in amino acid distribution, which indicates the strong interaction

between season and pedogenesis on amino acid distribution (Season*Age interaction by Two way-PerMANOVA, p=0.0.0002).

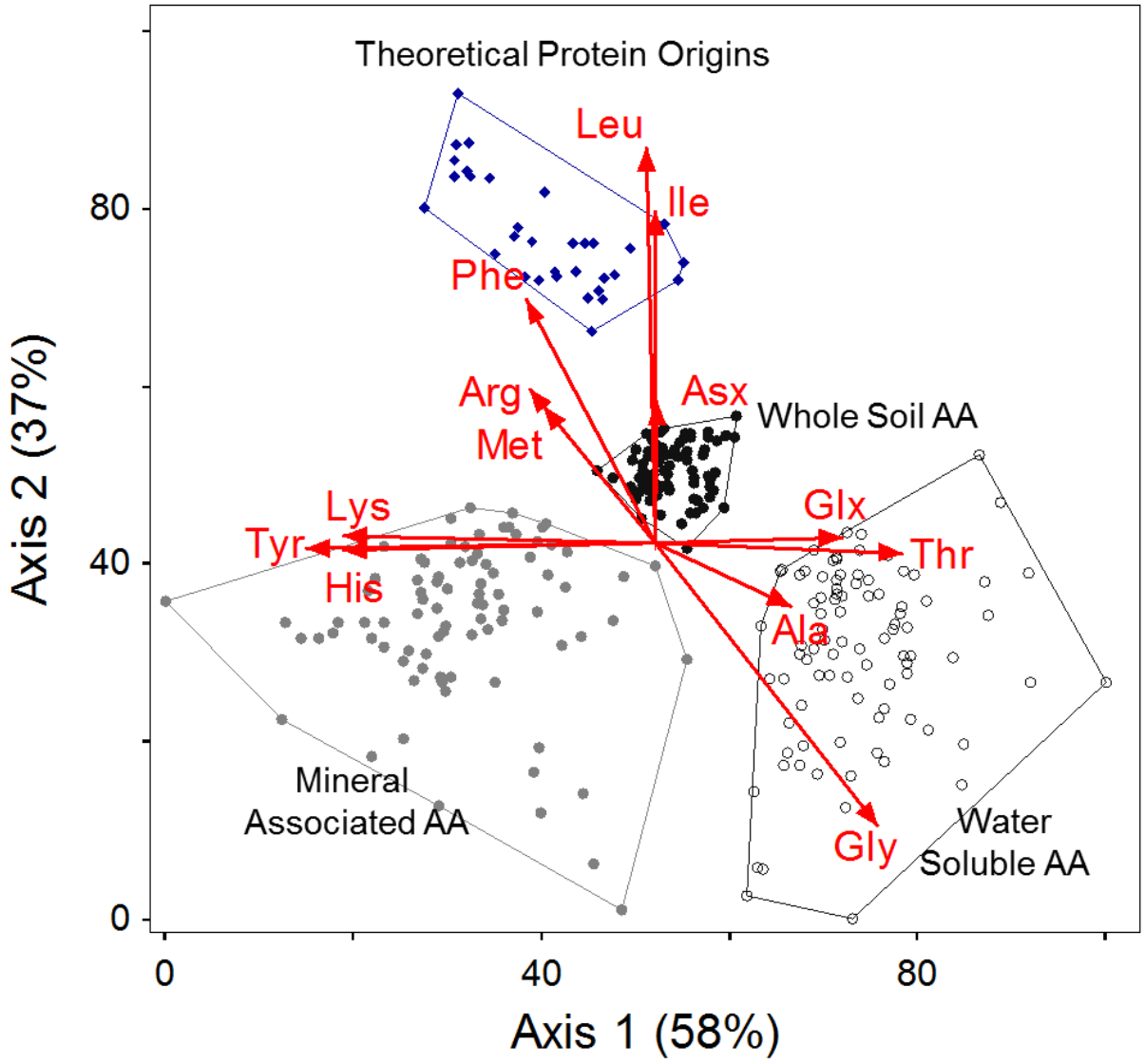


Figure.3. 3. Relative distribution of 17 proteinogenic amino acids from the theoretical protein origins (retrieved from Chen et al., 2013), and whole soil, mineral associated, and soluble OM hydrolysates across Lake Michigan chronosequence, plotted by Nonmetric multidimensional scaling (NMS) ordination. Theoretical protein origins was based on genome database (NCBI) and averaged by phylum level. Red vectors show the correlations of variables with ordination with r^2>0.3 where length and direction represent the magnitude and directions of the correlation. The distributions of amino acids were tested by MRPP between fractions (p<0.0001). The final stress for 2-d NMS was 10. Percentages on each axis denote the amount of variability associated with each axis. The Pearson and Kendall correlations of the vectors are provided in appendix (Table B3.6)

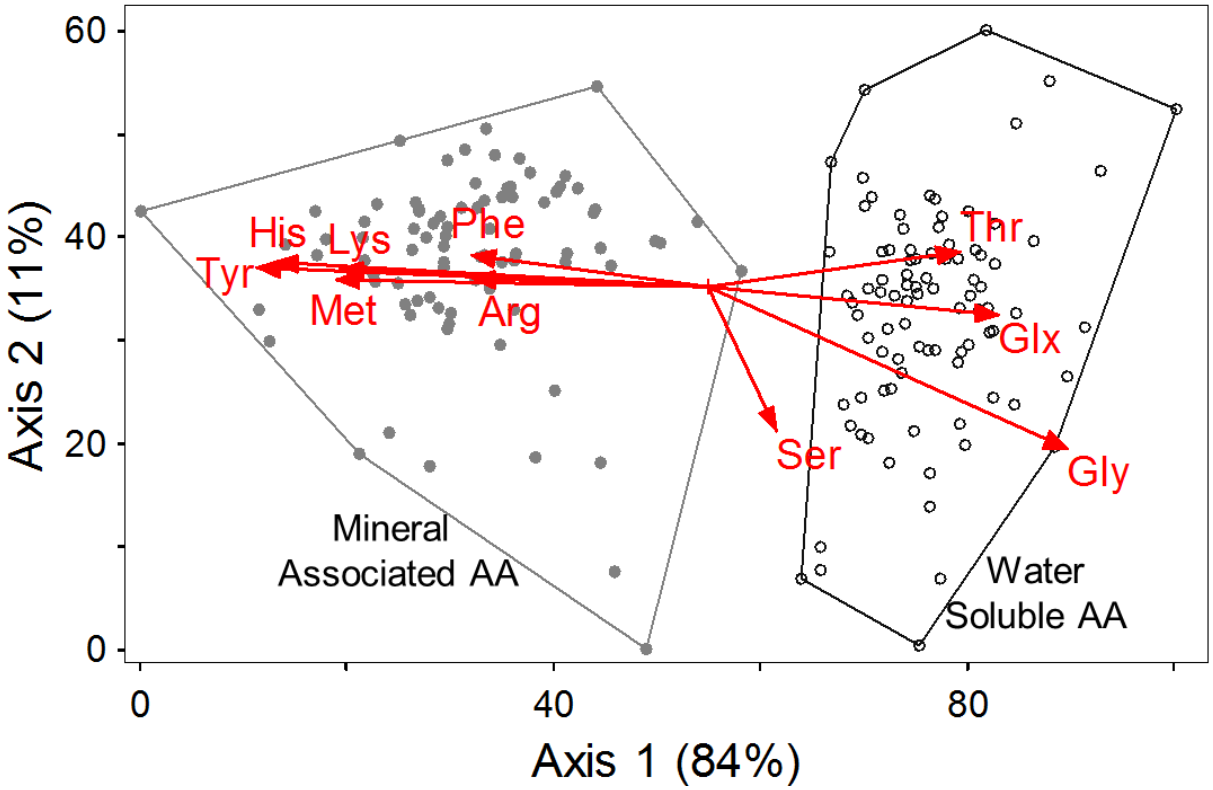


Figure.3. 4. Comparison in relative distribution of 17 proteinogenic amino acids between mineral associated and soluble OM pools across Lake Michigan chronosequence, plotted by Nonmetric multidimensional scaling (NMS) ordination. Red vectors show the correlations of variables with ordination with r^2>0.3 where length and direction represent the magnitude and directions of the correlation. The distributions of amino acids were tested by MRPP between fractions ($p<0.0001$). Percentages on each axis denote the amount of variability associated with each axis. The final stress for 2-d NMS was 11. The Pearson and Kendall correlations of the vectors are provided in appendix (Table B3.7)

3.4.4. Comparison in amino acid distribution among different OM hydrolysates

The amino acid distributions in soil were evidently clustered into three groups: whole soil, mineral associated, and soluble OM hydrolysates and each group was distinct from theoretical protein origins (Fig.3.3). It is distinguished that the position of whole soil cluster in distribution of amino acid was in the center among others, and relatively closer to the other sub-pools. Mineral associated and soluble OM pools were defined as sub-pools of whole soil OM pool, where OM were related to different soil components. In comparison between two soil components (mineral particle vs. water), positively charged amino acids (His, Lys, and Arg), metal binding amino acids (His, Met,

and Cys) and aromatic amino acids (His, Tyr, and Phe) were relatively enriched in the mineral associated OM fractions compared to the soluble OM fractions (Fig.3.4). Gly, Glx, and Thr were relatively enriched in the soluble OM fraction than mineral associated OM fraction. Among hydrophilic amino acids, neutral and negatively charged amino acids (amino acids that have hydroxyl side chain groups and carboxyl/amide side chain groups) were preferentially distributed in the soluble pool, while positively charged amino acids were preferentially distributed in the mineral associated pool. Among hydrophobic amino acids, Gly, which is the least hydrophobic among akyl group amino acids though and sometimes classified as polar, was predominant in soluble pool. Other hydrophobic amino acids, however, were distributed more in mineral associated pool (Appendix_Table B3.7). Chemically and physically very distinct pools were strongly related with the molecular species of amino acid; the change of amino acid species across SOM pools, thus, may provide possible mechanisms involved in amino acids localization regarding the nature of soil components such as water, mineral, and other organic compounds.

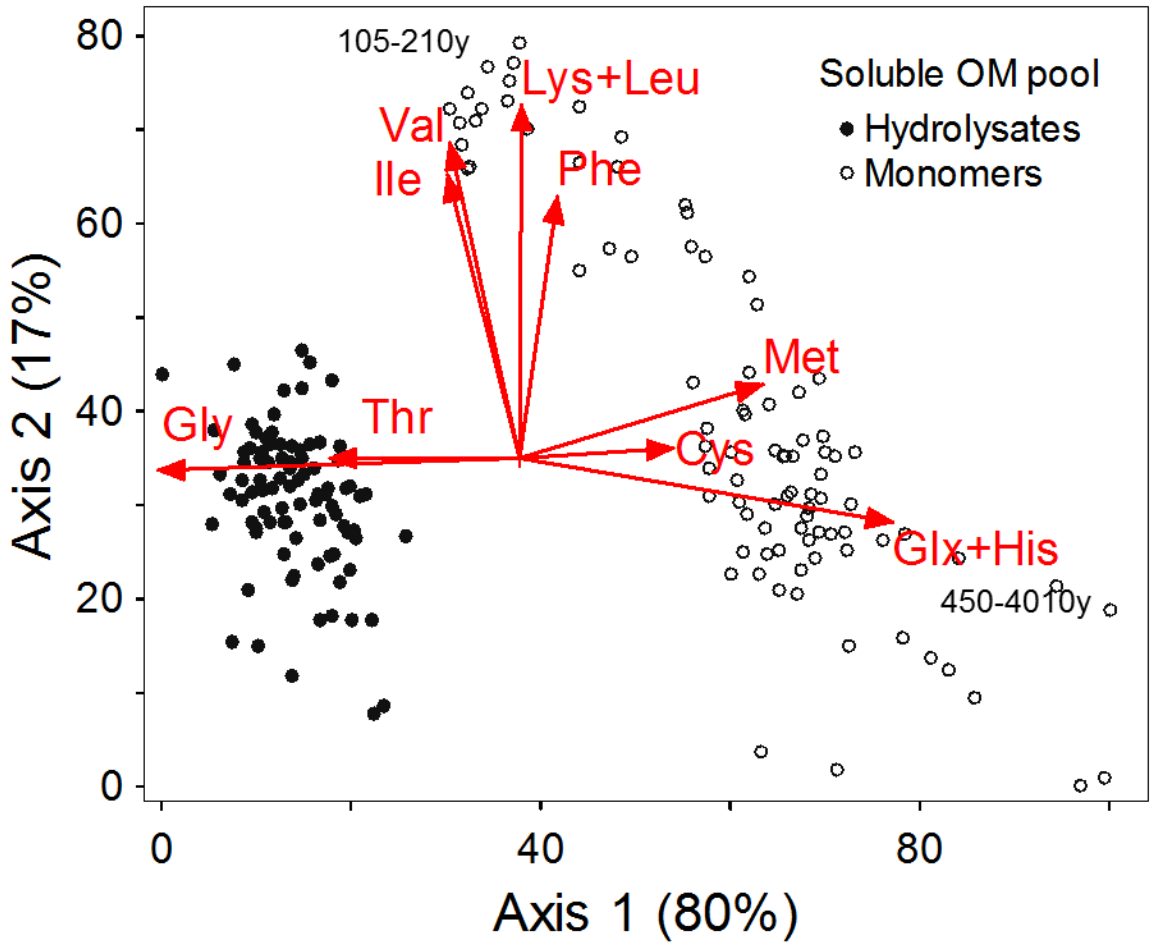


Figure.3. 5. Comparison in relative distribution of 17 proteinogenic amino acids between hydrolysates (polymers) and monomers within the soluble OM pool across Lake Michigan chronosequence, plotted by Nonmetric multidimensional scaling (NMS) ordination. Red vectors show the correlations of variables with ordination with r^2>0.3 where length and direction represent the magnitude and directions of the correlation. The distributions of amino acids were tested by MRPP between fractions (p<0.0001). Percentages on each axis denote the amount of variability associated with each axis. The final stress for 2-d NMS was 8. The Pearson and Kendall correlations of the vectors are provided in appendix (Table B3.8)

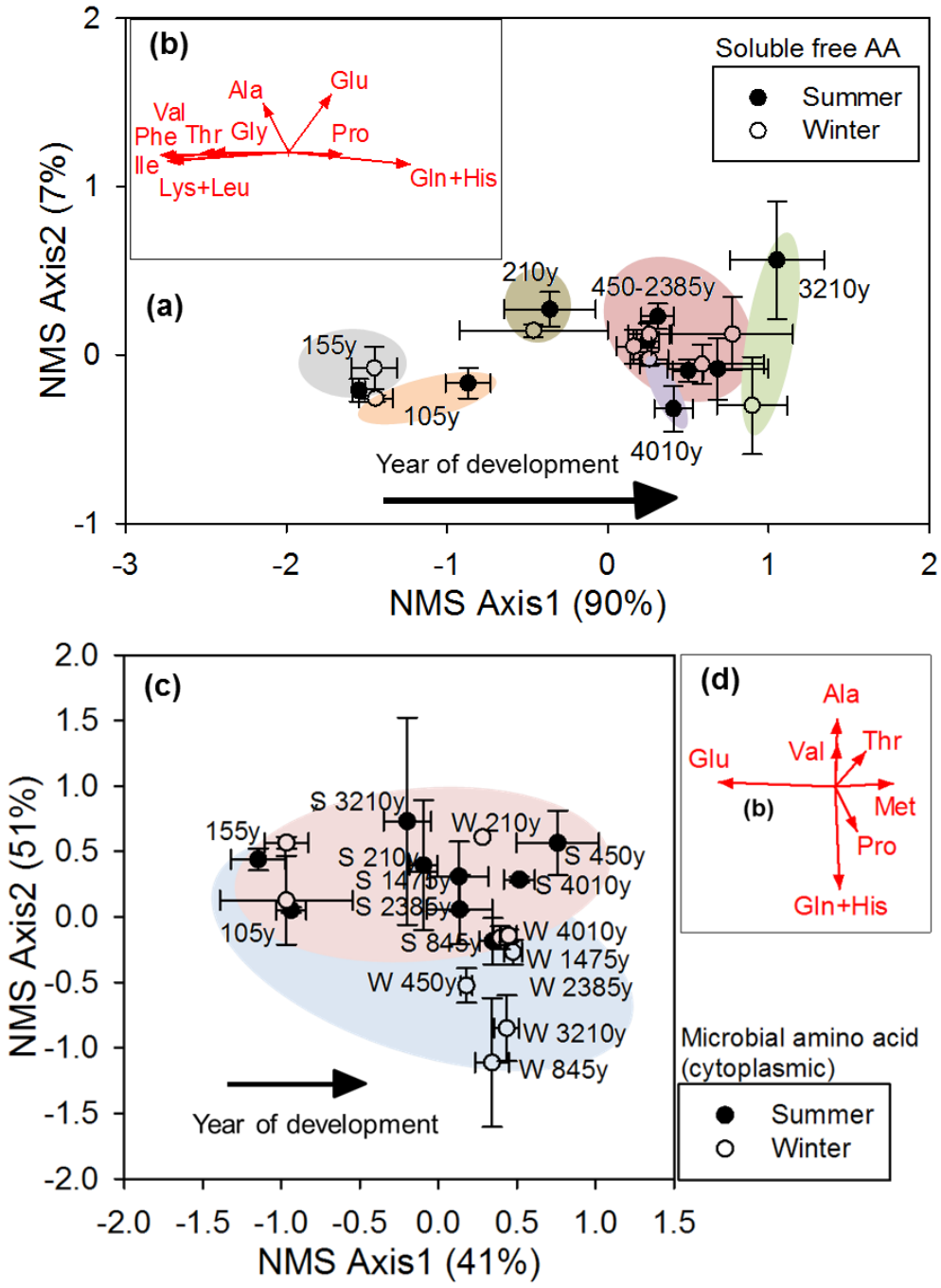


Figure.3. 6. Relative distribution of 19 proteinogenic amino acids from the soluble free (monomer) pool **(a)** and microbial (cytoplasmic) pool **(c)** between summer and winter during soil ecosystem development across Lake Michigan chronosequence, plotted by Nonmetric multidimensional scaling (NMS) ordination. Correlations of variables with ordination with r^2>0.3 were shown in bi-plot vector where length and direction represent the magnitude and directions of the correlation, respectively **(b)** and **(d)**. The distributions of soluble monomer amino acids were tested by Two way-PerMNOVA between summer and winter (p=0.4576); among site ages (p=0.0002); interaction term (p=0.0874). Due to unbalanced sample number, the distributions of microbial cytoplasmic amino acids were tested by MRPP between summer and winter (p<0.0001); among site ages (p<0.0001). Error bars in **(a)** and **(c)** represent standard error (n=5 and n≤5 respectively). Percentages on each axis on **(a)** and **(c)** denote the amount of variability associated with each axis. The final stress for 2-d NMS was 8 for **(a)** and 15 for **(c)**. The Pearson and Kendall correlations of the vectors are provided in appendix (Table B3.9 and 10)

3.4.5. Monomers vs. hydrolysates of amino acid in soluble OM fraction

The distribution of monomer amino acid in soluble pool was consistent between seasons (Fig.3.6.a, Two way-PerMNOVA p=0.4576). The change of amino acid distribution in soluble monomer extracts during ecosystem development was very clear accounting 90% of NMS variation (Fig.3.6.a, Two way-PerMANOVA p=0.0002). Most of the aliphatic amino acids, Gly ,Thr and Phe were more abundant in younger sites (105y-210y), whereas Gln+His and Pro were preferentially distributed in older sites (>450y).

The distributions of amino acid were very distinct between monomer extracts and hydrolysates in the soluble pool (Fig.3.5, MRPP p<0.0001). Gly and Thr were enriched in hydrolysates, while most amino acids that have hydrophilic side chain were positively correlated with monomer extracts in soluble pool. Monomer of amino acid in soluble pool showed clear pedogenic patterns of change in relative distribution with NMS axis1 (Fig3.6.a). Compared to hydrolysates, Val, Ile, Phe, and Lys+Leu were strongly related to the monomer extracts at younger sites (105y-210y), and Glx+His, Met and Cys were positively correlated with that at older sites (>450y) (Fig.3.5).

Polymers of amino acids were determined by subtracting hydrolysable pool by monomer pool of amino acid. This subtracted pool may include amino acid polymers that linked by peptide bonding or/and amino acids bound to other organic compounds and the bonding can be cleaved by chemical hydrolysis, where peptides and organic complexes were soluble in water and smaller than 0.22 μm in diameter. Polymers or organic complexes were dominant form of amino acid in soluble OM fraction (avg. 87%) and there was no difference in the percentage of polymer amino acids to total soluble

hydrolysable amino acid between summer and winter (Appendix_Fig.B3.1, Two way-ANOVA p=0.8224).

3.4.6. Microbial amino acid

The distribution of amino acid released from microbial lysis, so-called microbial amino acid in this paper, from summer was different from that from winter (Fig.3.6.a. MRPP p<0.0001). Ala, Val, and Thr were positively correlated with summer and Gln+His, and Pro were correlated with winter. The relative distribution of microbial amino acid was also dynamic during the early 200 years of development (MRPP p<0.0001). Glu was relatively more abundant at the younger sites (105-210y), while Met was positively correlated with older sites (>450y) in both summer and winter. Although in both seasons Glu and Met showed similar trends related with pedogenesis, the interaction between season and pedogenesis seemed to appear.

3.4.7. Amino sugar in whole soil hydrolysable OM pool

The distribution of amino sugar in the whole soil hydrolysate was similar between summer and winter (Fig.3.6.c, Two way-MANOVA p=0.0484 respectively), dominated by Glucosamine (GlcN: avg. 72% of total amino acid abundance, Appendix_Fig.B3.4). The distribution of amino sugar, however, was significantly different with the year of development (Two way-MANOVA p=0.0002), mostly derived from change of GlcN. There was no interaction between season and pedogenic change, meaning that the shifts in amino sugar distribution with year of development were not affected by season (Season*Age interaction by Two way-PerMANOVA p=0.2014).

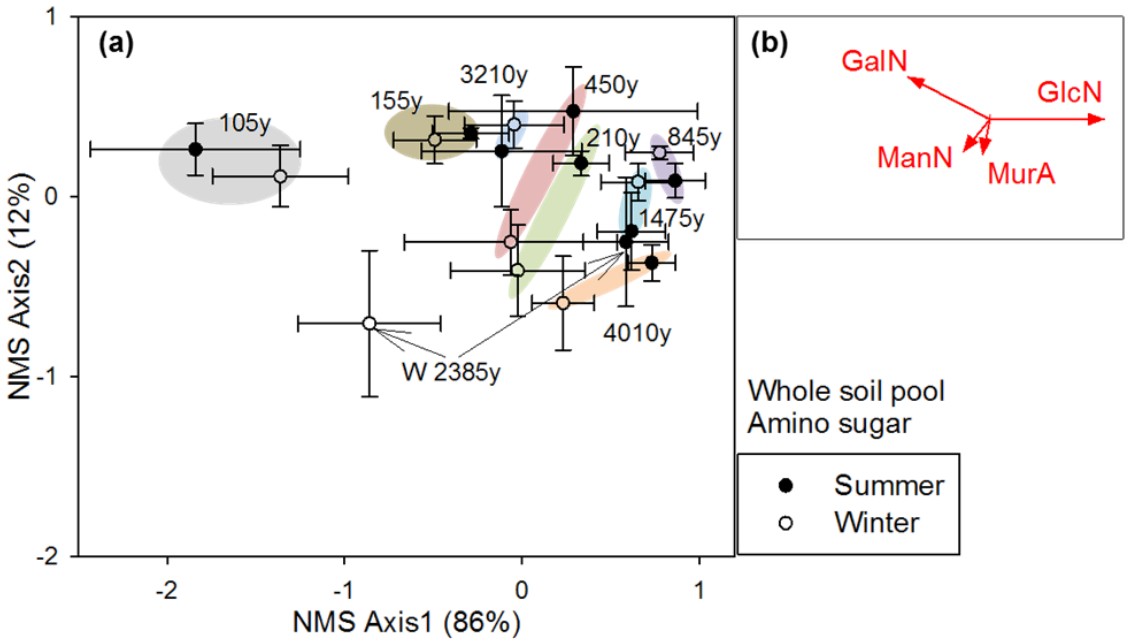


Figure.3. 7. Relative distribution of 4 amino sugars from the whole soil pool **(a)** between summer and winter during soil ecosystem development across Lake Michigan chronosequence, plotted by Nonmetric multidimensional scaling (NMS) ordination. Correlations of variables with ordination with r^2>0.3 were shown in bi-plot vector where length and direction represent the magnitude and directions of the correlation **(b)**. The distributions of whole soil pool were tested by Two way-PerMNOVA between summer and winter (p=0.0484); among site ages (p=0.0002); interaction term (p=0.2014). Error bars in **(a)** represent standard error (n=5). Percentages on each axis on **(a)** denote the amount of variability associated with each axis. The final stress for 2-d NMS was 7.

3.4.8. Microbial biomarkers: PLFA, amino sugars, and Orn

The abundance of total phospholipid fatty acids (PLFA) changed with season and pedogenesis. Seasonal dynamics of PLFA were more pronounced (Fig.3.8, ANOVA p=<0.0001) compared to those of amino sugar and amino acid (Appendix_Fig.B3.2, Two way-ANOVA p=0.0.4428 and p=0.3073, respectively). PLFA abundance was significantly great in summer compared to winter, whereas abundances of amino sugar and amino acid were slightly greater in winter despite no significant difference in amino sugar and amino acid by season. In comparison with amino sugar, PLFA immediately responded to aboveground vegetative changes during the ecosystem development. The abundance of PLFA peaked approximately at 719 μmol/kg-soil at 210y, and then about 30% decreased at 450y once the mixed pine forest was developed. Amino sugar, on the

other hand, kept accumulated up to 2610 μmol/kg-soil until 450y and began to decrease about 30% at 845y. Overall, the PLFA abundance shifted with age in a way to precede the change of amino acid abundance during ecosystem development.

The ratio of fungal to bacterial PLFA biomarker was consistent with seasonal changes (Fig.3.8.b, ANOVA p=0.1376), which agree with the shift in ratio of fungal to bacterial amino sugar biomarkers (GlcN/GalN) (Appendix_Fig.B3.6.b, ANOVA p=2322) and in ratio of amino sugar-C to amino acid-C (Appendix_Fig.B3.6.a, ANOVA p=0.3231). Nonetheless, the pedogenic pattern of change in the ratio of fungal to bacterial PLFA was different from those of the other ratios. The ratio of fungal to bacterial PLFA decreased during the early development, while the others increased gradually with age. The PLFA tracers behaved differently from amino acid and amino sugar

The abundance of ornithine (Orn) was different between seasons (Appendix_Fig.B3. 6.d, ANOVA p<0.0001) with relatively greater abundance in summer. Pedogenic change of Orn abundance was pretty much reflected by pedogenic shift of amino acid abundance in whole soil pool.

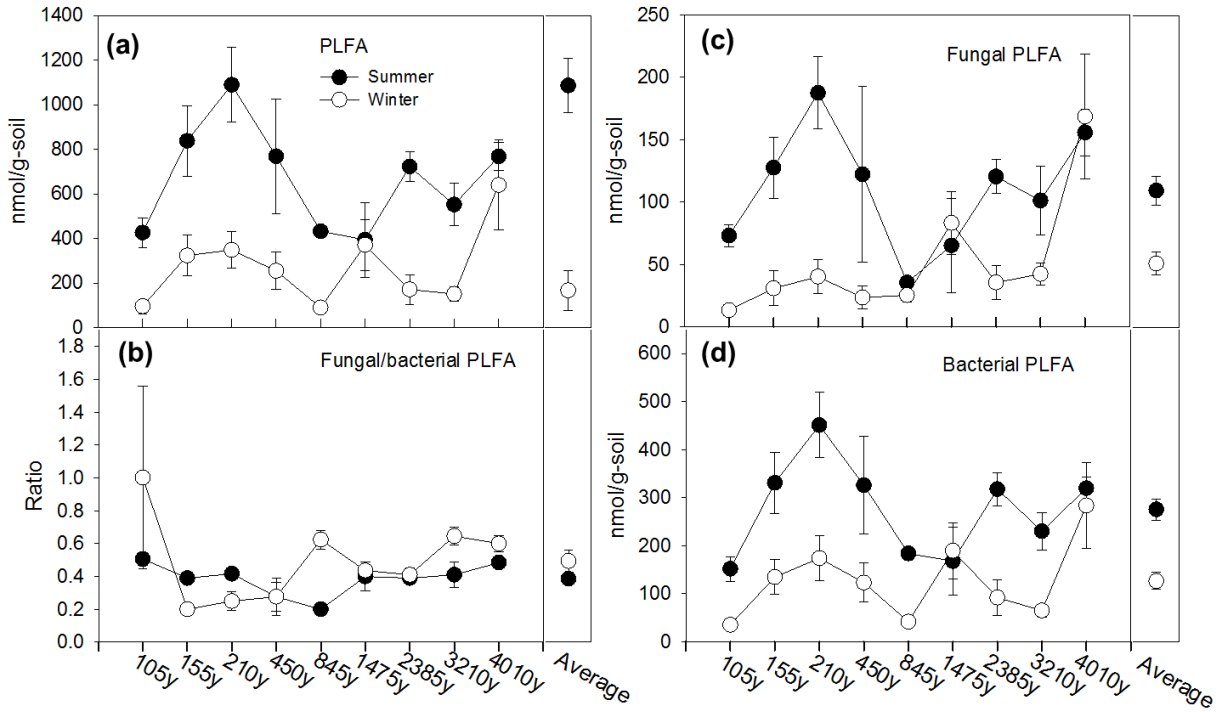


Figure.3. 8. Comparisons of the total PLFA (a), ratio of fugal to bacterial PLFA (b), fungal PLFA (c), bacterial PLFA (d) in whole soil pool between summer and winter across Lake Michigan chronosequence. Error bars represent standard error (n=5). The abundances were tested by Two way-ANOVA and the p-values show below

Abundance	Season	Age	Season*Age
PLFA	0.0513	**0.0008**	0.1022
F/B PLFA	**0.0166**	0.8914	0.2147
F PLFA	0.1216	**0.0054**	**0.0428**
B PLFA	0.1074	**0.0010**	0.1089

3.4.9. Abundance of amino acid

The abundance of amino acid in the whole soil pool did not differ between summer and winter, averaging of 948 mg/kg-soil (Appendix_Fig.B3.2.a, Two way-ANOVA p=0.3073). Mineral associated OM sub-pool was on average 13%- amino acid in whole soil pool (Appendix_Fig.B3.3.e). The abundance and fraction size of amino acid in the mineral associated fraction were affected by season (Appendix_Fig.B3.3.a and e, Two way-ANOVA p=0.0002 for abundance, p=0.0011 for fraction size). Water soluble OM pool contributed to avg. 1%-amino acid in whole soil pool (Appendix_Fig.B3.3.f). Soluble fraction was consistent in fraction size between summer and winter, but some differences of absolute abundance appeared (Fig.3.9, Two way-ANOVA p=0.0744 for fraction size and p=0.0181 for abundance). The abundance and fraction size of monomer amino acid in soluble pool were consistent between seasons (Appendix_Fig.B3.3.c and 3.3.g, Two way-ANOVA p=0.4806 and p=0.3114 respectively). The pool size of amino acid released from microbial lysis, so-called microbial amino acid in this paper, was 174% larger than that of soluble free amino acid. The abundance of amino acid in microbial pool was different between summer and winter (Appendix_Fig.B3.3.d). Soils collected in winter were significantly larger in microbial free amino acid compared to those in summer (Two way-ANOVA p=0.0071). No difference in fraction size of microbial amino acid was shown (Appendix_Fig.B3.3.h, Two way-ANOVA p=0.0879). The abundance of amino sugar in the whole soil pool were similar between summer and winter (Appendix_Fig.B3.2.b, Two way-ANOVA p=0.0.4428), averaging 319 mg/kg-soil, dominated by Glucosamine (GlcN: avg. 72% of total amino acid abundance, Appendix_Fig.B3.4).

3.5. Discussion

This is the first report to describe proteinogenic amino acids related to soil organic N dynamics and formation associated with whole soil, mineral associated, and soluble pools over 4000 years of pedogenesis. The results showed that Gly, Ala, Asx, Glx, and Ser are always among the most dominant amino acids across SOM pools. The distributions of proteinogenic amino acids across SOM pools have displayed both agreement and disagreement with similar previous report. For example, proteinogenic amino acids were characterized as obtaining similar distribution with minor variation regardless seasons and successional stages (Werdin-Pfisterer *et al.*, 2009), as well as horizons (Werdin-Pfisterer *et al.*, 2012), climates (Campbell *et al.*, 1991, Senwo & Tabatabai, 1998, Sowden *et al.*, 1977), and organic amendment (Gotoh (Gotoh *et al.*, 1986) et al., 1986). The consistent distribution of amino acids is thought to be due to the commonality on dominant amino acid originated from a variety of sources or on similar biochemical process in soil regardless affecting factors tested previously (Werdin-Pfisterer *et al.*, 2012). The current study in details, nonetheless, revealed, remarkable variations in amino acid distribution between growing and dormant seasons, as well as those with pedogenic changes related to soil ecosystem development, by using multivariable statistic methods.

3.5.1. Origins and transformation of amino acids in soil

It is supported that the relative distribution of amino acids in whole soil OM pool has shown a distinct fingerprint compared to that of their origins (Fig.3.1), such as plant (Marumoto *et al.*, 1972) or microbial amino acid (Chen *et al.*, 2013, Friedel & Scheller, 2002). So as, in soluble and mineral associated OM pools, the relative distributions of

amino acid species were different from those of their biological origins. This implies that the physicochemical properties of proteinaceous compounds and their biochemical mechanisms associated with soil solution or mineral surfaces might be responsible for selective distribution of amino acids on different SOM pools and consequently for their turnover. Moreover, regarding the observation of that amino acid distribution of whole soil OM was in between those of other OM pools and biological sources, the amino acid distribution found in whole soil OM might be the average of source and sink of proteinogenic amino acids, resulting in the consistency in distribution (also found in Chapter4. Fig,4,4).

3.5.2. Selective partitioning of amino acids associated with soil constituents

The compositional differences among whole soil, soluble and mineral associated OM can help to explain accumulation patterns related to different SOM pools. The degree of hydrophobicity or polarity and side chain properties of amino acid might be controlling factors to distribute amino acids among water soluble, mineral associated, and occluded SOM pools (Allard, 2006, Fu *et al.*, 2014, Knicker & Hatcher, 1997, Murphy *et al.*, 1990, Sollins *et al.*, 2006, Wershaw, 1986). Polar amino acids, for example, are presumed to be distributed more in soluble monomer form rather than non-polar amino acids due to aqueous nature (Werdin-Pfisterer *et al.*, 2012). We have shown that, in general, polar amino acids indeed more frequently appeared in soluble monomer form. In addition to degree of polarity, the charge of side chain on amino acids, furthermore, tended to associate with the chemistry of soil matrix components such as mineral surface and soil solution. Most of the amino acids that contain neutral or negative charge on their side chains preferentially associated with aqueous component,

whereas those are positively charged have shown to associate more with mineral matrix, especially for silicate minerals dominated by negative charge on the surfaces.

Mineral associated fraction, which is theoretically stabilized with little mobility and limited accessibility, is thought to remain in soil for a longer term from many decades to thousands of years (Kögel-Knabner *et al.*, 2008, Lützow *et al.*, 2006a). There is evidence of at least some portion of mineral associated PM does turn over. The compositional variations explained by approximately 70% (Fig.3.1.c) was highly related to 4000 years of ecosystem development, suggesting that the OM associated with mineral may be rather dynamic by replenishing OM to and alternately adsorbing OM from the whole soil pools which may result in shifting SOM distribution. About 22% variation in amino acid distribution, furthermore, was associated with seasonality which represents relatively short term timeframe response. The fraction size of mineral associated OM changed by seasons although season did not affect the amino acid content in whole soil OM pool. This indicates that mineral associated OM, thought to be long residence pool, might be partially exchanged during the annual cycling in response to seasonality. More interestingly, Ser in the whole soil OM was preferentially distributed in summer, but was found to be more associated with mineral in summer. Glx in the whole soil OM, in contrast, was greater in relative abundance in winter, possibly due to its large input after litter decomposition after fall, and was shown to accumulate in mineral associated fraction in summer. These opposite trends exhibited in two OM pools with seasons suggest that there might be dynamics of OM among the pools.

3.5.3. Microbial contribution to SOM formation

The quantitative and compositional changes of organic N in soil approaching microorganisms may be resulted from the transformation of plant materials to microbial residues (Cotrufo *et al.*, 2013, Miltner *et al.*, 2012). Most plant carbohydrates in litter are thought to be rapidly degraded, whereas microbial cell wall fragments appear to remain in soil for long time (Angers, 1992, Chantigny *et al.*, 1997, Foster *et al.*, 1983, Tisdall & Oades, 1980). When plant material are decomposed in soil, its distribution of amino compounds is exchanged through the microbial metabolisms and approaches what is contained in the soil natively, and that amino sugar compounds are synthesized newly and accumulated in the soil due to their resistance to decomposition (Kai *et al.*, 1973). For example, fungi and bacteria contain avg. 11% and 56% C as an amino acid-C respectively, and both contain 1-2% C as an amino sugar-C (Hobara *et al.*, 2014). Hobara *et al.* (2014) have shown that the proportion of amino acid-C to total organic C was initially 2-4% on the plant litters and increased to ~ 9%, so as amino sugar-C twice increased over 3 years of decomposition. Accordingly, across the 4000 years of soil ecosystem development, the percentages of amino-acid-C and amino sugar-C to the total organic C were ranged from 6-27% and 1-3% respectively in our study site. This indicates that amino compounds in soil might approach those in microorganisms through microbial mediated biogeochemical processes. Relatively greater degree of amino sugar accrual compared to amino acid, however, might be due to their longer persistence as resistant cell wall fragments (Chenu & Cosentino, 2011, Linhares & Martin, 1978, Miltner *et al.*, 2012).

Fungi cell residues likely become relatively more important in contributing to SOM formation with soil ecosystem development compared to bacteria, supported by molecular biomarkers, especially cell wall constituents. GlcN occurs in various microbial cell wall membrane complexes: fungal chitins, bacterial peptidoglycans. GalN, on the other hand, occurs in bacterial cell wall and membrane complexes and in archaeal glycoproteins (Amelung, 2001, Giroldo & Vieira, 2002, Parsons, 1981, Schäffer & Messner, 2001). Since the GlcN yields from fungi are higher than those from bacteria, the ratio GlcN to GalN is used for the indicator of fungal to bacterial contribution (Amelung, 2003, Joergensen & Wichern, 2008). MurA is found in bacterial peptidoglycan, so the ratio of GlcN to MurA is also indicator of fungal to bacterial contribution (Amelung, 2001). In our study site, both ratios increased for the first 1000 years of ecosystem development (Appendix_Fig.B3.6.b and c), suggesting the increase in fungal contribution to SOM during this time period. In addition, due to larger content of amino acid-C but similar level of amino sugar-C content in bacteria compared to fungi (Appendix_Fig.B3.5), the ratio of amino sugar-C to amino acid-C can reflect fungi to bacteria contribution to organic C content in soil. The ratio of amino sugar-C to amino acid-C increased during the earlier ecosystem development (Appendix_Fig.B3.6.a), which further supports the trends of other molecular tracers. In addition, Orn is non-protein amino acid and occurs in bacterial peptidoglycan and in Orn-containing lipids (Lehninger, 1979, Ratledge & Wilkinson, 1988). The relative abundance of Orn compared to proteinogenic amino acids can be indicative of relative bacterial contribution to SOM compared to other SOM origins, and it decreased over time during the ecosystem development. Biodegradation and stability of these molecules in soil,

however, are known to be relatively different (Amelung, 2003, Engelking *et al.*, 2007, Kawasaki & Benner, 2006, Tremblay & Benner, 2006), and so these ratios does not merely indicate difference in contribution by two microbial groups, but also the degree of stabilization of SOM. For example, amino sugar rather than amino acid and GlcN rather than MurA are relatively resistant to decomposition in soil, suggesting relatively longer persistence of amino sugars especially GlcN.

Another molecular tracer, PLFA, was measured and the ratio of fungal to bacterial PLFA was rather consistent across the chronosequence. Unlike to other molecular tracers, PLFA is rapidly degraded in soil once the death and lysis of microbial cells occur; thus, it represents living organisms. The signature fatty acids are used to determine different microbial groups; for example, PLFA's i15:0, a15:0, i16:0, a16:0, i17:0, a17:0 (gram-positive), 16:1ω9, 16:1 ω 7, 18:1 ω 7 and cy19:0 (gram-negative) were considered as bacterial biomarkers, 10:Me 16:0 and 10:Me 18:0 for actinomycetes and 18:1 ω 9 and 18:2 ω 6 as fungal biomarkers. Based on the consistency in ratio of fungal to bacterial living biomass combined with the increase in ratio of fungal to bacterial SOM contribution with age, cell debris derived from fungi are likely preserved in soil and contributing to SOM formation with little seasonal variations.

Together with GlcN, MurA, and Orn, several proteinogenic amino acids occur in bacterial peptidoglycan. D-form amino acids such as D-Ala and D-Gln in peptideoglycan are one of tracers, since they are rarely utilized by other organisms. In addition, Gly occurs as Gly-pentapeptide inter-bridge in peptidoglycan structure (Kai *et al.*, 1973). Our results have shown that Gly was predominated as a peptide form, possibly repeating strucuture, in the soluble pool. This might relate to large production of the

bacteria derived Gly-pentapeptide in soluble SOM pool during the biodegradation of peptidoglycan. We hypothesized that the sizable portion of bacterial cell debris is likely dissolved in soil water during their decomposition, and consequently relatively decline in their contribution to preserved SOM compared to fungi. Because acid hydrolysis amino acid analysis does not provide sequences of amino acids, however, further investigation on peptide mass fingerprinting of soluble pool is needed to confirm our hypothesis.

3.6. Conclusions

Overall, there was significant seasonal effect on shifting the relative distribution of organic N, especially amino acid although season had little or no influence on the abundance and pool size of organic N. Seasonal dynamics of amino acids in whole soil and mineral associated OM pools indicated that at least some part of operationally defined as slow pools was seasonally cycled. Ser, one of amino acids showed the strong accumulation in the summer of whole soil pool, was relatively abundant in the winter of mineral associated pool. The opposite trends were found for Glx. Mineral associated fraction is not composed solely of so-called stable OM. The relative abundances of the amino acids preferentially distributed in mineral associated fraction (e.g. His, Arg, Lys, and Phe) gradually increased in whole soil fraction with year of ecosystem development, while those positively correlated with soluble fraction (e.g. Gly) declined in the whole soil fraction over time during pedogenesis. Compared to biological origins of amino acids, the signature of distribution in amino acid in soil was distinct, and the amino acid distribution of neither mineral associated nor soluble fraction resembled biological sources. This suggested that the interactions of amino acids with the mineral and soil solution provide selective partitioning for amino acids.

References

Allard B (2006) A comparative study on the chemical composition of humic acids from forest soil, agricultural soil and lignite deposit: Bound lipid, carbohydrate and amino acid distributions. *Geoderma,* **130**, 77-96.

Amelung W (2001) Methods using amino sugars as markers for microbial residues in soil. *Assessment methods for soil carbon*, 233-272.

Amelung W (2003) Nitrogen biomarkers and their fate in soil. *Journal of Plant Nutrition and Soil Science,* **166**, 677-686.

Amelung W, Zhang X (2001) Determination of amino acid enantiomers in soils. *Soil Biology and Biochemistry,* **33**, 553-562.

Angers D (1992) Changes in soil aggregation and organic carbon under corn and alfalfa. *Soil Science Society of America Journal,* **56**, 1244-1249.

Bosch L, Alegría A, Farré R (2006) Application of the 6-aminoquinolyl-N-hydroxysccinimidyl carbamate (AQC) reagent to the RP-HPLC determination of amino acids in infant foods. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences,* **831**, 176-183.

Bossio D, Scow K, Gunapala N, Graham K (1998) Determinants of soil microbial communities: effects of agricultural management, season, and soil type on phospholipid fatty acid profiles. *Microbial Ecology,* **36**, 1-12.

Butler JL, Williams MA, Bottomley PJ, Myrold DD (2003) Microbial community dynamics associated with rhizosphere carbon flow. *Applied and Environmental Microbiology,* **69**, 6793-6800.

Campbell CA, Zentner RP, Knipfel JE, Schnitzer M, Lafond GP (1991) Thirty-Year Crop Rotations and Management Practices Effects on Soil and Amino Nitrogen. *Soil Science Society of America,* **55**, 739.

Chantigny MH, Angers DA, Prévost D, Vézina L-P, Chalifour F-P (1997) Soil aggregation and fungal and bacterial biomass under annual and perennial cropping systems. *Soil Science Society of America Journal,* **61**, 262-267.

Chen W, Shao Y, Chen F (2013) Evolution of complete proteomes: guanine-cytosine pressure, phylogeny and environmental influences blend the proteomic architecture. *BMC Evolutionary Biology,* **13**, 219.

Chenu C, Cosentino D (2011) Microbial regulation of soil structural dynamics. *The architecture and biology of soils: life in inner space*, 37-70.

Cotrufo MF, Wallenstein MD, Boot CM, Denef K, Paul E (2013) The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Global Change Biology,* **19**, 988-995.

Engelking B, Flessa H, Joergensen RG (2007) Shifts in amino sugar and ergosterol contents after addition of sucrose and cellulose to soil. *Soil Biology and Biochemistry,* **39**, 2111-2118.

Flanagan P, Cleve KV (1983) Nutrient cycling in relation to decomposition and organic-matter quality in taiga ecosystems. *Canadian Journal of Forest Research,* **13**, 795-817.

Foster R, Rovira A, Cock T (1983) *Ultrastructure of the root-soil interface*, American Phytopathological Society.

Friedel JK, Scheller E (2002) Composition of hydrolysable amino acids in soil organic matter and soil microbial biomass. *Soil Biology & Biochemistry,* **34**, 315-325.
Frostegård Å, Bååth E (1996) The use of phospholipid fatty acid analysis to estimate bacterial and fungal biomass in soil. *Biology and Fertility of Soils,* **22**, 59-65.
Fu IW, Markegard CB, Chu BK, Nguyen HD (2014) Role of Hydrophobicity on Self-Assembly by Peptide Amphiphiles via Molecular Dynamics Simulations. *Langmuir,* **30**, 7745-7754.
Giroldo D, Vieira AA (2002) An extracellular sulfated fucose-rich polysaccharide produced by a tropical strain of Cryptomonas obovata (Cryptophyceae). *Journal of Applied Phycology,* **14**, 185-191.
Gotoh S, Araragi M, Koga H, Ono S-I (1986) Hydrolyzable organic forms of nitrogen in some rice soil profiles as affected by organic matter application. *Soil Science and Plant Nutrition,* **32**, 535-550.
Hobara S, Osono T, Hirose D, Noro K, Hirota M, Benner R (2014) The roles of microorganisms in litter decomposition and soil formation. *Biogeochemistry,* **118**, 471-486.
Hou S, He H, Zhang X, Zhang W, Xie H (2009) Determination of soil amino acids by high performance liquid chromatography-electro spray ionization-mass spectrometry derivatized with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate. *Talanta,* **80**, 440-447.
Jämtgård S, Näsholm T, Huss-Danell K (2010) Nitrogen compounds in soil solutions of agricultural land. *Soil Biology and Biochemistry,* **42**, 2325-2330.
Joergensen RG, Wichern F (2008) Quantitative assessment of the fungal contribution to microbial tissue in soil. *Soil Biology and Biochemistry,* **40**, 2977-2991.
Kai H, Ahmad Z, Harada T (1973) Factors affecting immobilization and release of nitrogen in soil and chemical characteristics of the nitrogen newly immobilized III. Transformation of the nitrogen immobilized in soil and its chemical characteristics. *Soil Science and Plant Nutrition,* **19**, 275-286.
Kaiser K, Guggenberger G (2007) Distribution of hydrous aluminium and iron over density fractions depends on organic matter load and ultrasonic dispersion. *Geoderma,* **140**, 140-146.
Kawasaki N, Benner R (2006) Bacterial Release of Dissolved Organic Matter during Cell Growth and Decline: Molecular Origin and Composition. *Limnology and Oceanography,* **51**, 2170-2180.
Kielland K (1995) Landscape Patterns of Free Amino Acids in Arctic Tundra Soils. *Biogeochemistry,* **31**, 85-98.
Knicker H, Hatcher PG (1997) Survival of Protein in an Organic-Rich Sediment: Possible Protection by Encapsulation in Organic Matter. *Naturwissenschaften,* **84**, 231-234.
Kögel-Knabner I, Ekschmitt K, Flessa H, Guggenberger G, Matzner E, Marschner B, Von Lützow M (2008) An integrative approach of organic matter stabilization in temperate soils: Linking chemistry, physics, and biology. *Journal of Plant Nutrition and Soil Science,* **171**, 5-13.
Küry D, Keller U (1991) Trimethylsilyl-O-methyloxime derivatives for the measurement of [6, 6-2 H 2]-d-glucose-enriched plasma samples by gas chromatography—

mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications,* **572**, 302-306.
Lehninger AL (1979) Biochemistry, 1975. *Worth, New York*, 268-271.
Liang C, Fujinuma R, Balser TC (2008) Comparing PLFA and amino sugars for microbial analysis in an Upper Michigan old growth forest. *Soil Biology and Biochemistry,* **40**, 2063-2065.
Lichter J (1995) Lake Michigan beach-ridge and dune development, lake level, and variability in regional water balance. *Quaternary Research,* **44**, 181-189.
Lichter J (2000) Colonization constraints during primary succession on coastal Lake Michigan sand dunes. *Journal of Ecology,* **88**, 825-839.
Linhares L, Martin J (1978) Decomposition in soil of the humic acid-type polymers (melanins) of Eurotium echinulatum, Aspergillus glaucus sp. and other fungi. *Soil Science Society of America Journal,* **42**, 738-743.
Lützow MV, Kögel-Knabner I, Ekschmitt K, Matzner E, Guggenberger G, Marschner B, Flessa H (2006a) Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions – a review. *European Journal of Soil Science,* **57**, 426-445.
Lützow MV, Kögel-Knabner I, Ekschmitt K, Matzner E, Guggenberger G, Marschner B, Flessa H (2006b) Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions–a review. *European Journal of Soil Science,* **57**, 426-445.
Marumoto T, Furukawa K, Yoshida T, Kai H, Harada T (1972) Effect of the application of rye-grass on the contents of individual amino acids and amino sugars contained in the organic nitrogen in soil. *Kyushu Univ Fac Agr J*.
Mikutta R, Kaiser K, Dörr N *et al.* (2010) Mineralogical impact on organic nitrogen across a long-term soil chronosequence (0.3–4100 kyr). *Geochimica et Cosmochimica Acta,* **74**, 2142-2164.
Miltner A, Bombach P, Schmidt-Brücken B, Kästner M (2012) SOM genesis: microbial biomass as a significant source. *Biogeochemistry,* **111**, 41-55.
Miltner A, Kindler R, Knicker H, Richnow H-H, Kästner M (2009) Fate of microbial biomass-derived amino acids in soil and their contribution to soil organic matter. *Organic Geochemistry,* **40**, 978-985.
Murphy EM, Zachara JM, Smith SC (1990) Influence of mineral-bound humic substances on the sorption of hydrophobic organic compounds. *Environmental Science & Technology,* **24**, 1507-1516.
Norman F, Boas F (1953) Method of determination of hexosamines in tissues. *J. biol. Chem*, 553-563.
Parsons JW (1981) Chemistry and distribution of amino sugars in soils and soil organisms. *Soil biochemistry,* **5**.
Paul S, Veldkamp E, Flessa H (2008) Differential response of mineral-associated organic matter in tropical soils formed in volcanic ashes and marine Tertiary sediment to treatment with HCl, NaOCl, and Na 4 P 2 O 7. *Soil Biology and Biochemistry,* **40**, 1846-1855.
Ratledge C, Wilkinson S (1988) Fatty acids, related and derived lipids. *Microbial lipids,* **1**, 23-52.

Ringelberg DB, Sutton S, White DC (1997) Biomass, bioactivity and biodiversity: microbial ecology of the deep subsurface: analysis of ester-linked phospholipid fatty acids. *FEMS Microbiology Reviews,* **20**, 371-377.
Schäffer C, Messner P (2001) Glycobiology of surface layer proteins. *Biochimie,* **83**, 591-599.
Schimel JP, Bennett J (2004) Nitrogen Mineralization: Challenges of a Changing Paradigm. *Ecology,* **85**, 591-602.
Senwo Z, Tabatabai M (1998) Amino acid composition of soil organic matter. *Biology and Fertility of Soils,* **26**, 235-242.
Sollins P, Swanston C, Kleber M *et al.* (2006) Organic C and N stabilization in a forest soil: Evidence from sequential density fractionation. *Soil Biology and Biochemistry,* **38**, 3313-3324.
Sowden FJ, Chen Y, Schnitzer M (1977) The nitrogen distribution in soils formed under widely differing climatic conditions. *Geochimica et Cosmochimica Acta,* **41**, 1524-1526.
Tisdall J, Oades J (1980) The management of ryegrass to stabilise aggregates of a red brown earth. *Soil Research,* **18**, 415-422.
Tremblay L, Benner R (2006) Microbial contributions to N-immobilization and organic matter preservation in decaying plant detritus. *Geochimica et Cosmochimica Acta,* **70**, 133-146.
Van Cleve K, Alexander V (1981) Nitrogen cycling in tundra and boreal ecosystems. *Ecological Bulletins (Sweden)*.
Weintraub MN, Schimel JP (2005) The Seasonal Dynamics of Amino Acids and Other Nutrients in Alaskan Arctic Tundra Soils. *Biogeochemistry,* **73**, 359-380.
Werdin-Pfisterer NR, Kielland K, Boone RD (2009) Soil amino acid composition across a boreal forest successional sequence. *Soil Biology and Biochemistry,* **41**, 1210-1220.
Werdin-Pfisterer NR, Kielland K, Boone RD (2012) Buried organic horizons represent amino acid reservoirs in boreal forest soils. *Soil Biology & Biochemistry,* **55**, 122-131.
Wershaw RL (1986) A new model for humic materials and their interactions with hydrophobic organic chemicals in soil-water or sediment-water systems. *Journal of Contaminant Hydrology,* **1**, 29-45.
Williams MA, Jangid K, Shanmugam SG, Whitman WB (2013) Bacterial communities in soil mimic patterns of vegetative succession and ecosystem climax but are resilient to change between seasons. *Soil Biology and Biochemistry,* **57**, 749.
Zhang W, Parker K, Luo Y, Wan S, Wallace L, Hu S (2005) Soil microbial responses to experimental warming and clipping in a tallgrass prairie. *Global Change Biology,* **11**, 266-277.
Zhang X, Amelung W (1996) Gas chromatographic determination of muramic acid, glucosamine, mannosamine, and galactosamine in soils. *Soil Biology and Biochemistry,* **28**, 1201-1206.

Chapter 4. Similarity in selecting patterns of protein amino acid during pedogenesis in two disparate chronosequences located in Lake Michigan, USA and Haast River, New Zealand

i. Authors: Jinyoung Moon[1], Kang Xia[2], Benjamin L. Turner[3], Mark A. Williams[1]

ii. Institute:
[1]Soil Microbial Ecology and Biogeochemistry Laboratory, Department of Horticulture, Virginia Polytechnic Institute and State University, 312 Latham Hall, 220 Ag Quad Ln., Blacksburg, VA 24061
[2]Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, 1880 Pratt Dr., Blacksburg, VA 24061
[3]Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancon, Republic of Panama

iii. Corresponding Author: Mark A. Williams, Phone: 540-231-2547, FAX 540-231-3083, Email: markwill@vt.edu

iv. Keywords: Mineral associated organic matter, organo-mineral association, soil organic matter (SOM) formation, soil organic nitrogen (SON), soil protein, hydrolysable amino acid, HPLC, primary succession, proteinaceous compounds

v. Type of paper: Primary Research Articles

Title: Similarity in selecting patterns of protein amino acid during pedogenesis in two disparate chronosequences located in Lake Michigan, USA and Haast River, New Zealand

4.1. Abstract

The emerging evidence of preferential accumulation and long residence time of proteinaceous compounds in soil are counter to the traditional view that their structure is readily broken down by soil microbial activities. Knowledge of the residence time of these compounds in soil organic matter (SOM) pools is for understanding global biogeochemical nitrogen, and ultimately carbon cycles. We tested (1) whether proteinaceous compounds are either randomly or selectively accumulated, (2) whether proteinaceous compounds are selectively associated with mineral particles, and (3) if patterns of change can be explained and confirmed in two independent pedogenesis and ecosystem development gradients. To accomplish the objectives, we determined the distribution of amino acids – structure unit of proteinaceous compounds – in whole soil organic matter (OM) pool and mineral associated OM sub-pool. Soils were sampled from two geologically separated and climatically different sand dune chronosequences where primary successions had been progressed: adjacent to Lake Michigan, USA (~4010y) and Haast River and Tasman Sea, New Zealand (~6500y). We found the consistency of selecting patterns of proteinaceous compounds of two disparate locations in three major ways: (i) similarity of proteinogenic amino acid fingerprints in whole soil pools, (ii) resemblance of strong selection of proteinogenic amino acid by mineral associated fractions, and (iii) simultaneous change patterns of proteinogenic amino acids along with biological community successions. The similarity in

transformation of sources to whole soil pools in these two locations provided evidence that a mixed pool of plant and microbial derived OM that has gone through the process of selective preservation, enriching (glycine, alanine, serine, and aspartic acid+asparagine). The silicate mineral associated fractions showed evidence for a strong selection of positively charged (histidine, arginine, and lysine), aromatic (phenylalanine and tyrosine), or sulfur containing (methionine and cysteine) amino acids (referring as sink selection). With soil ecosystem development, both locations showed that the long-term accumulation patterns of amino acids were closely related with shifts in their biological sources (referring as source selection) (r^2=0.71, p<0.0001 for Michigan and r^2=0.71, p=0.0002 for Haast). The consistency of the results at two locations in the southern and northern hemispheres is strong evidence that SOM formation processes and dynamics associated with pedogenesis and ecosystem development are parsimonious and predictable.

4.2. Introduction

Although more and more evidence is reported that supports the importance of proteinaceous compounds as a pool of both labile and the recalcitrant soil organic nitrogen (SON), it is uncertain which factors are causing their degradation and which mechanisms are responsible for their sequestration. These compounds can be decomposed incorporating into global N and C cycling. However, a considerable part escapes complete mineralization and the residues are entering the stabilized soil organic matter (SOM) pool (Cotrufo *et al.*, 2013), prolonging their residence time in soil system. The residence time of proteinaceous compounds such as polypeptides was estimated several hundred years (Amelung *et al.*, 2006) or more than ten thousand

years (Curry *et al.*, 1994). Thus, they largely contributes to the sequestration of N into soil reservoir, which was estimated total about 60 Tg N per year accumulating in soil system (Galloway *et al.*, 2004).

Several possible mechanisms to explain the preservation of proteinaceous compounds in soil have been suggested. It is unclear which soil factor is more important between mineral particles or other organic matter constituents (Knicker, 2011). The hypothesis of association with minerals to protect these compounds from enzymatic attacks is supported by the evidence of variability on the amount and distribution of proteinaceous compounds recovered from the different soil particle size fractions (Ding & Henrichs, 2002) and from soils with different mineral constituents (Mikutta *et al.*, 2010). Proteinaceous compounds adsorb strongly to mineral surface and they are physically protected in mesopores <10 nm particle size that are too small for degrading enzyme to enter (Aufdenkampe *et al.*, 2001, Ding & Henrichs, 2002, Wang & Lee, 1993). The proteinaceous compounds adsorbed to mesopore size mineral are rather small peptides than larger proteins and this is agreed with the sorption behavior of amino acid monomers and polymers onto fabricated mesoporous alumina and silica, studied by Zimmerman *et al.* (2004). Schnitzer and Kodama (1992) reported that non-crystalline inorganics separated from the prairie soils were rich in Si, which appeared to contribute to the preferential accumulation of neutral amino acid, while non-crystalline components from the soils from eastern Canada were rich in Al, which may have been associated with the accumulation of acidic amino acids in these soils. Also, it has been shown that basic amino acids are typically enriched in environments with negatively charged aluminosilicate minerals (Aufdenkampe *et al.*, 2001, Keil *et al.*, 1998), while

sorption to metal oxides is selective for acidic amino acids (Matrajt & Blanot, 2004). Based on these findings, the nature of inorganic soil components appears to influence the type of SON that was formed and accumulated in the soil environment. Alternatively, but also based on adsorption to mineral surfaces, Sollins *et al.* (2006) suggested that proteinaceous compounds may form a stable inner organic layer around a mineral surface and this inner layer may help less polar organic compounds sorb more readily to the mineral surfaces. Nonetheless, similar or even longer residence time of proteinaceous compounds is observed in mineral-poor soils such as sapropels and peats compared with in mineral soils (Knicker & Hatcher, 2001).

Since accruals of proteinaceous compounds have shown to be ubiquitous regardless soil mineral contents, the hypothesis of the biopolymer interactions for proteinaceous compounds stabilization was proposed; the proteinaceous compounds are connected to resistant aliphatic polymers (hydrophobic macromolecules) and surrounded by these polymers, and therefore they are protected from biological degradation (Knicker & Hatcher, 1997, Zang *et al.*, 2000). The mechanisms also include chemical incorporations and reactions of proteinaceous compounds with reducing sugars (Maillard reaction), polyphenols, quinones, and tannins (Espeland & Wetzel, 2001, Fan *et al.*, 2004). Allard (2006) has observed that the relative distribution of neutral polar amino acids to total amino acids was significantly larger in lignite deposit (at the higher degree of humification) than that in soil. This suggests a preferential preservation of polar amino acids and proteinaceous compounds associated rich in these amino acids would be retained in internal voids of three dimensional structure of other organic matter by hydrogen bonds. In addition, based on the amino acid studies

by acid hydrolysis, basic amino acids have lower concentrations in general, which might have to do with their greater ability to react with reducing saccharides and quinones (Swift and Posner, 1972; Szajdak and Österberg, 1996). Another suggested hypothesis is intrinsic stabilization of peptide/protein by modification of their key groups that are recognized by enzymes or conformational restrictions such as amyloid aggregates and fibrils that efficiently protect them in soil ecosystem (Nelson et al., 2008; Rillig et al., 2007). Although a lot of possible mechanisms to stabilize proteinaceous compounds have been suggested up to now, we are far from a satisfactory understanding.

In this study, we compared proteinaceous compounds between two geologically separate and climatically different sites with gradients of soil ecosystem development in order to determine the commonality of dynamics of proteinaceous compounds during ecosystem development. The objectives are (1) whether proteinaceous compounds are either randomly or selectively accumulated, (2) whether proteinaceous compounds are selectively associated with mineral particles, and (3) if patterns of change can be explained and confirmed in two independent pedogenesis and ecosystem development gradients. To accomplish the objectives, we determined the distribution of amino acids – structure unit of proteinaceous compounds – in whole soil organic matter (OM) pool and mineral associated OM sub-pool.

4.3. Materials and methods

4.3.1. Study sites

4.3.1.1. Lake Michigan Chronosequence, U.S.A. (Michigan site)

4.3.1.1.1. Location and climate

The study site consists of a series of beach-dune ridges bordering Lake Michigan (N 45.72729, W84.94076), and located in Wilderness State Park in Emmet County of the northern lower peninsula of Michigan. The park lies between 177 and 225 m elevation (0–48 m above lake level). There is >108 eolian deposited dune ridges running parallel to the shoreline with depositional ages from present day to w4500 years (Lichter, 1995). The site is under temperate and boreal climate resgion. Temperature and precipitation averaged 6.28°C and 77.2 cm per year, respectively, between 1951 and 1980 at Mackinaw City (Nurnberger, 1996), 15 km to the east.

4.3.1.1.2. Dune formation and parent materials

The park consists of lake plains that developed during and since the mid-Holocene Nipissing lake stages (3800–5500 years B.P.). Nipissing-aged features at the site include a series of high parabolic dunes and a well-marked beach (Leverett & Taylor, 1915, Spurr & Zumberge, 1956). Post-Nipissing features consist of an extensive 5-km strandplain containing approximately 108 shore-parallel dune-capped beach ridges, which have formed, on average, every 32.4 years over the past ~4500 years (Lichter, 1995b). The ridges are approximately 2.5 km long, 10–30 m wide, and vary between 3 and 5 m in height above the basal foreshore deposits except where episodes of shore erosion destabilized ridges and produced slowly moving parabolic dunes

reaching 15 m height. The dune ridges have a parent material originating from glacial deposits and Paleozoic bedrock underlying the lake basin. The parent material is assumed to be similar across the dune sequence. Fine sands deposited on the lake shore are dominated by quartz but contain numerous other minerals in minor quantities (Lichter, 1995a).

4.3.1.1.3. Soil types and properties

The youngest soils (<100 y) are mapped as dunes which then develop into Deer Park sands (soil series) and described taxonomically as mixed, frigid, Spodic Udipsamments. The oldest soils (>1475 y) tend to be mapped to the Roscommon series, and are mixed, frigid Mollic Psammaquents. Soil Ca and Mg levels decreased in a log-linear pattern and were concurrent with declining pH (7.6-3.5) as soils aged from younger to older soils across the chronosequence. Soil organic matter and total soil organic C (but not mineralizable C) decreased along the chronosequence from younger to older soils ($r^2 = 0.76$; $P < 0.05$). Soil Na (~149 mg/g) and P (~4 mg/g), in contrast, did not change with soil development (Lichter, 1998).

4.3.1.1.4. Vegetation

The change in plant community structure was greater during early compared to late ecosystem development. Generally speaking, dune-building grass species were replaced by evergreen shrubs and these were then replaced by mixed pine forests. This shift in early-succession to late-succession plant species happened at 450 years of soil and ecosystem development, when the early-succession species began to disappear and the mixed pine forest began to develop. Early succession was thus defined by considerable turnover of plant species. Indeed, plant community composition in the

young dunes (105-155 y) was completely different from communities observed at 210 y, which were again taxonomically different from those >450 y of ecosystem development. Once the forest matured, the plant species composition stabilized and there was no major change in the plant community structure during late ecosystem development (P = 0.59) (Williams *et al.*, 2013).

4.3.1.1.5. Bacterial community

Bacterial communities showed patterns of change across the chronosequence during early ecosystem development (<845 y) but changed little during latter (845-4010 y) ecosystem development. The chronosequence gradient showed a number of changes in phyla but were generally dominated by the abundance and dynamics of Acidobacteria, Actinobacteria, and Alphaproteobacteria, comprising 71% of all the sampled sequences. Other less abundant phyla (<4%) were Bacteroidetes, Cyanobacteria, Firmicutes, Planctomycetes, Betaproteobacteria, and Gammaproteobacteria. Between early (<450 y) and late (>450 y) ecosystem development, Acidobacteria increased approximately 6-fold from around 4% to w30%. Actinobacterial abundance declined, in contrast, from around 60 to w35% during this same time. The gradient of ecosystem development also was described by changes in low abundance taxa, with Bacteroidetes and Firmicutes, for example declining and Planctomycetes and Gammaproteobacteria increasing 4-fold. Cyanobacterial abundance declined from 5% to less than 0.5% following 210 y of ecosystem development (Williams *et al.*, 2013).

4.3.1.2. Haast River Chronosequence, New Zealand (Haast site)

4.3.1.2.1. Location and Climate

The study site consists of set in a foredune barrier system of NE- to SW-aligned shore-parallel coastal dunes (beach ridges) on a prograding coastal plain west of the Southern Alps, northwest of the Haast River, on the west coast of the South Island of New Zealand (43°43′20″ S, 169°4′30″ E) (Eger *et al.*, 2011). The Haast dune system extends ~10 km alongshore and 5 km inland, with dunes 20–100m long rising up to 20 m above adjacent dune slacks. There are a total of seventeen dune ridges that occur as generally continuous features across the length of the system (Turner *et al.*, 2012). The site is under lowland temperate rain forest. Temperature and precipitation averaged 11.3 °C and 345.5 cm per year, respectively based on the 36 year period between 1941 and 1976 at Haast Beach (New Zealand Meteorological Service, 1983). Relative humidity averages 83%.

4.3.1.2.2. Formation and parent material

The oldest dune is ca 6000–7000 year old, forming after the culmination of the post-glacial sea level rise (Chappell & Shackleton, 1986, Gibb, 1986). At least for the six youngest dunes (age range AD1826 to AD1230), dune building has been shown to be associated with episodic sediment pulses brought down the Haast River after Alpine Fault earthquakes (Wells & Goff, 2007). Dune ridges are interspaced by poorly drained swales creating a relief difference of <5 m to 20m within the dune system. To the north and south of the dune sequence are large alluvial fans of the Waita and Haast Rivers, respectively. Glacially sculptured outcrops of quartzo-feldspathic gneiss of the Greenland Group (Late Cambrian–Ordovician) also occur near the site (Rattenbury *et*

al., 2010). Parent material is uniform quartzo-feldspathic dune sand derived from well-foliated schist. The mineralogy of the unweathered sand appears relatively uniform across the chronosequence. being 40–50% quartz, with the remainder feldspar, mica and chlorite (Palmer *et al.*, 1985).

4.3.1.2.3. Soil types and properties

Short lived phases of Entisols followed by Inceptisols culminate in the formation of persistent Spodosol forms within 1000 to >30,000 y depending on rainfall. Soils develop rapidly to podzols (Spodosols) under the super-humid climate of the west coast of New Zealand. Eluvial horizons are reflected by low pH in the upper part of the soil (<4.5) and illuvial horizons by accumulation of poorly or noncrystalline Fe, Al and occasional Si together with organic matter. These trends are accompanied by decreasing base saturation, increasing C/N ratios, and depletion of alkaline cations and apatite phosphorus. These processes are often promoted by acid litter-producing conifer vegetation. Impeded drainage is typical of more advanced stages of Spodosol pedogenesis resulting in in-situ formation of low permeable, massive silt loam horizons and the decline of coarser fractions and the formation of cemented iron pans or Bs horizons as a result of iron translocation (Eger *et al.*, 2011).

4.3.1.2.4. Vegetation

Forests in the region are mixed conifer–broadleaf temperate rain forest, which have persisted in the lowlands since 7700 B.P., and probably since 11,400 B.P. (Li *et al.*, 2008). The conifers consist of members of the family Podocarpaceae, which occur widely throughout New Zealand forests, (Coomes & Bellingham, 2011). Prominent species include *Dacrydium cupressinum* (rimu), *Prumnopitys ferruginea* (miro),

Podocarpus hallii (montane totara), and *Phyllocladus alpinus* (celery pine). Woody angiosperms in the area include *Weinmannia racemosa* (kamahi), *Coprosma spp.*, *Metrosideros umbellata* (southern rata), and *Nothofagus menziesii* (silver beech), as well as the tree ferns *Dicksonia squarrosa* (wheki) and *Cyathea smithii* (pateke). The youngest dune has been largely cleared of forest and converted to pasture, but some low stature forest remains on the seaward dune crest. Detailed analysis of vegetation changes along the sequence will be reported elsewhere (Turner *et al.*, 2012).

4.3.1.2.5. Bacterial community

Bacterial communities showed patterns of change during pedogenesis, with the largest change during the first several hundred years after dune stabilization. The most abundant bacterial taxa were Alphaproteobacteria, Actinobacteria and Acidobacteria. These include taxa most closely related to nitrogen-fixing bacteria, and suggest heterotrophic nitrogen input may be important throughout the chronosequence. Changes in bacterial community structure were related to changes in several soil properties, including total phosphorus, C:N ratio, and pH. The Bacteroidetes, Actinobacteria, Cyanobacteria, Firmicutes, and Betaproteobacteria all showed a general decline in abundance as pedogenesis proceeded, while Acidobacteria, Alphaproteobacteria, and Plantctomycetes tended to increase as soils aged. Conclusions There were trends in the dynamics of bacterial community composition and structure in soil during ecosystem development. Bacterial communities changed in ways that appear to be consistent with a model of ecosystem progression and retrogression, perhaps indicating fundamental processes underpin patterns of below and above-ground community change during ecosystem development (Jangid *et al.*, 2013).

4.3.2. Soil sampling

Five replicates of top soil samples were collected from the incipient A-E horizon (0-15cm, 5-cm dia.) in nine dunes of age 105, 155, 210, 450, 845, 1475, 2385, 3210, and 4010 years at the Michigan site by the same way as previous published literature (Williams *et al.*, 2013). Each replicate was separated by 10-m intervals across transects along each dune's crest. Five replicates of freshly deposited beach sands were also sampled to assess the community composition of parent material expected to be similar to the source material that formed the eolian deposits of the dune soils. Thus, 50 plots were sampled. The soil samples were stored in sterile Whirlpak bags, and frozen immediately in coolers with dry ice and kept in -20°C. Soil from each plot was collected in August, 2008.

Six dunes of age 181, 392, 517, 1826, 4422, and 6500 years at the Haast site were sampled by the same manner as previous published literature (Jangid *et al.*, 2013). Four replicate plots (5×10 m), separated by ~50 m were established along the crest of each dune. Ten locations within each plot were chosen and soil collected from mineral soil layer (0 to 20 cm depth) with the use of a 2.5 cm diameter soil probe. The sample bags were frozen immediately in cooler packed filled with dry ice. Thus, 24 plots were sampled. Upon arrival in the laboratory, soils were thawed for~30 min, homogenized through a 2-mm sieve, extraneous roots and organic materials were removed, and the samples were kept in –20 °C.

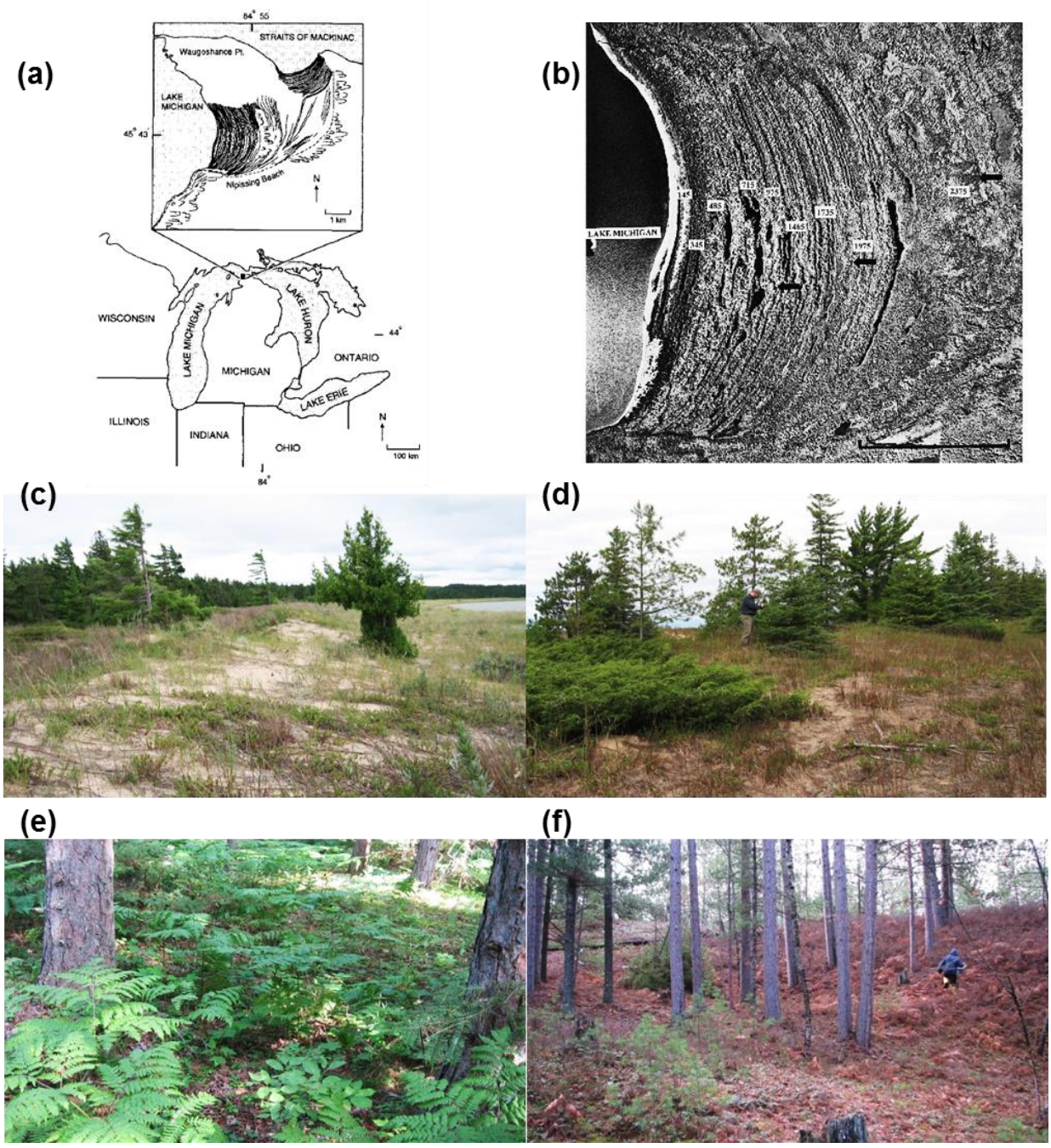


Figure.4. 1 **(a)** Map showing the location of Wilderness State Park in Ernmet County, northern lower Michigan, (Lichter, 1005) **(b)** Aerial photograph of the beach-ridge chronosequence. Arrows indicate parabolic-dune development, with youngest dunes on the left close to the beach, and oldest dunes on the right. Scales1 km. (Lichter 1998).**(c)** Vegetation in 105 year development site; **(d)** Vegetation in 155 year development site; **(e)** Vegetation in 450 year development site; **(f)** Vegetation in 1475 year development site. (Pictures taken by Williams' lab)

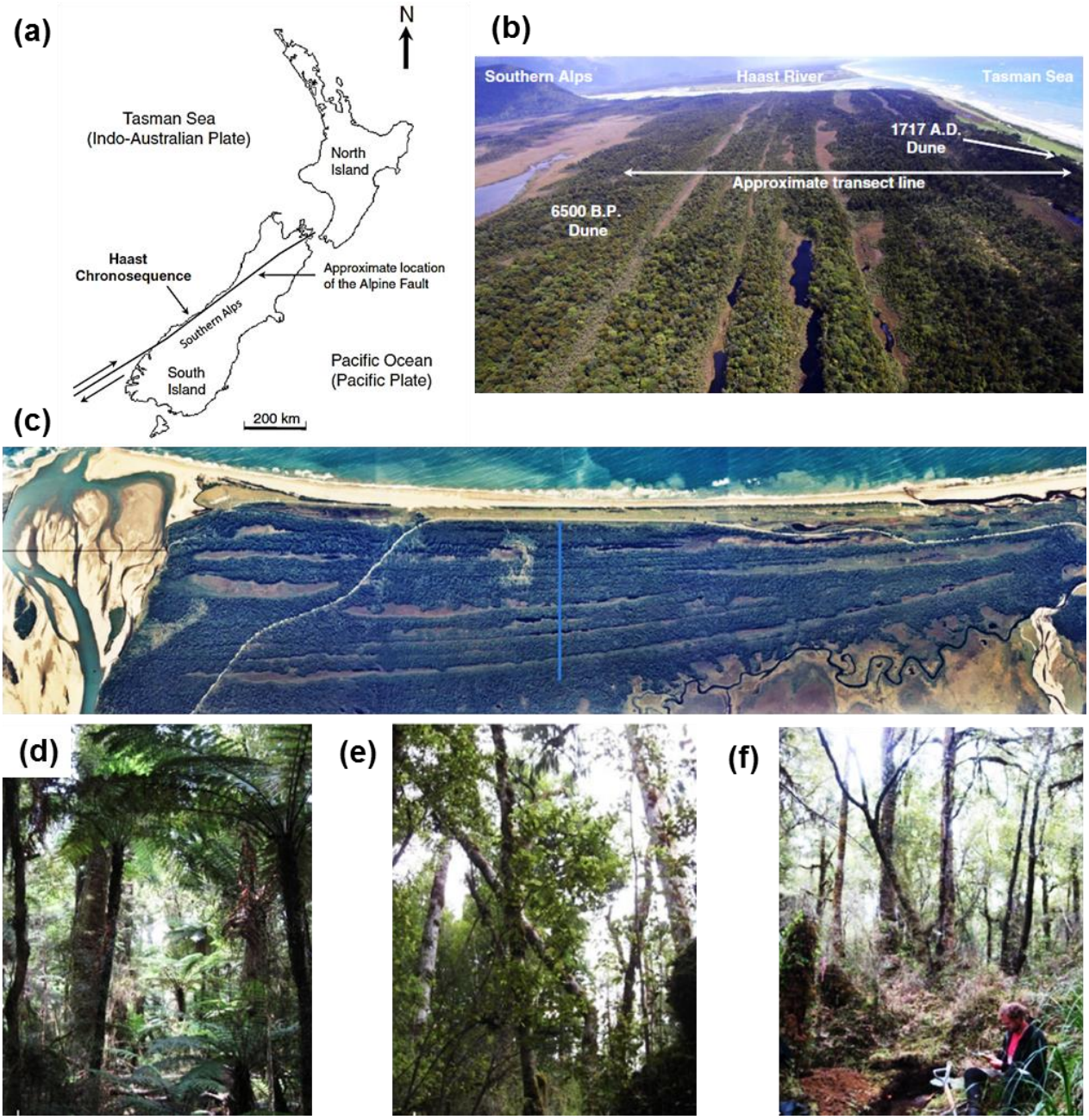


Figure.4. 2. **(a)** The location of the Haast chronosequence, South Island, New Zealand (cite). **(b)** Aerial view of the Haast Chronosequence looking south towards the Haast River in the distance, with the youngest dunes on the right close to the ocean, indicated by Dune 2 formed following the 1717 A.D. earthquake, and the oldest dunes furthest inland, indicated by the 6500 B.P. dune (Turner et al., 2012). **(c)** The Haast chronosequence, showing a an aerial image of the entire sequence with the approximate transect line indicated by the blue bar, with youngest dunes on the top close to the road, and oldest dunes on the bottom. **(d)** Vegetation in 517 year development site; **(e)** Vegetation in 1,826 year development site; **(f)** 3,903 year development site (cite).

4.3.3. Whole soil hydrolysable amino acid analysis

The hydrolysable amino acids in the whole soil were acid digested, purified, and then analyzed using post-derivatization high performance liquid chromatography (HPLC). Two to five grams (dry weight) of moist soil was hydrolyzed in 10 ml of 6 M HCl with an internal standard (L-norvaline) at 110 °C for 24 h (Amelung & Zhang, 2001). After hydrolysis, the soil hydrolysates were centrifuged at 10,000 *Xg* for 10 min. The aliquot of the 400 µl supernatant was diluted in 55 ml ultra-pure water and cleaned on a preconditioned Dowex 50Wx8 resin (hydrogen form, 50-100 mesh; Alfa Aesar, Cat# B22109) (Küry & Keller, 1991, Norman & BOAS, 1953). The interfering metals were removed by rinsing with 0.1 M oxalic acid (pH 1.6-1.7). Amino acids retained on the resin were eluted with 30 ml 3M NH_4OH, filtered through a 0.22 µm polyvinylidene fluoride (PVDF) membrane syringe filter, vacuum-dried, reconstituted in 10 µl 0.05 M HCl, and finally derivatized using the AccQ Fluor™ reagent kit (Fluorescent 6-Aminoquinoly-N-Hydroxysuccinimidyl Carbamate derivatizing reagent; Waters Co. Cat# WAT052880) following the standard protocol from Bosch *et al.* (2006) and Hou *et al.* (2009). Chromatographic separations on the HPLC 1260 Infinity system (Agilent Technologies, USA) were carried out on a reversed phase column (Waters X-Terra MS C18, 3.5µm, 2.1X150mm). The mobile phase consisted of A: an aqueous solution containing 140 mM sodium acetate, 17 mM TEA, and 0.1% (g/L, w/v) EDTA-2Na (pH 5.05, adjusted with phosphoric acid solution) and B: ACN/water (60:40, v/v). The gradient conditions were 0 - 17 min 100 - 93% A, 17 - 21 min 93 - 90% A, 21 - 30 min 90 - 70% A, 30 - 35 min 70% A, 35 - 36 min 70 - 0% A, and 0% A for 4 min. The column was thermostated at 50 °C and operated at a flow rate of 0.35 ml/min. The sample

injection volume was 5 µL. The analytes detection was carried out using a fluorescence detector (λ_{ex} = 250 nm and λ_{em} = 395 nm) (Bosch *et al.*, 2006, Hou *et al.*, 2009). Hydrolysable amino acids in the samples were qualified and quantified by comparison with amino acid standard solutions at different concentrations. Each amino acid standard solution contained 20 amino acids including alanine (Ala), arginine (Arg), aspartic acid (Asp), asparagine (Asn), cystine (Cys–Cys), glutamic acid (Glu), glutamine (Gln), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tyrosine (Tyr), tryptophane (Trp), and valine (Val). Because of the transformation of Asn to Asp and Gln to Glu and the destruction of Trp during acid hydrolysis, 17 amino acids except Asn, Gln, and Trp were quantified for hydrolysable proteinogenic amino acids. Non-protein amino acid, ornithine (Orn) was also quantified as an indicator of bacterial contribution in soil.

4.3.4. Soil mineral associated amino acid analysis

Soil mineral associated fraction was isolated by the density gradient fractionation method (Kaiser & Guggenberger, 2007), followed by amino acid analysis in the mineral associated fraction (heavy fraction). Air-dried soil (2.5 g) were fractionated using sodium metatungstate (SMT, $H_2\ Na_6\ O_{40}\ W_{12}$) solutions with a density of 2.4 g/cm^3. The mixture was vigorously agitated on a shaker until the soil was completely dispersed. After the dispersion, the sample was centrifuged and the floating particulate (light fraction) was carefully separated from the heavy fraction. The heavy fraction was thoroughly cleaned with distilled water and completely dried at 60^{o}C in an oven overnight. The dried heavy fraction was weighed and hydrolyzed by using the same procedure with the whole soil

hydrolysable amino acid analysis as described. The heavy fraction is referred to as mineral associated OM fraction.

4.3.5. Statistics

For the multivariate comparison, molecular species of amino acid concentration were transformed by using the general relativization to remove the potentially strong influence of absolute abundance on distribution. Multi-Response Permutation Procedures (MRPP) and Nonmetric multidimensional scaling (NMS) ordination were performed using the PC-ORD software version 6.0 (MjM Software, Gleneden Beach, OR, USA) to compare the effect of soil age on the relative abundance (mol%) of 17 proteinogenic amino acids in whole soil and mineral associated OM hydrolysates. The cutoff of statistical significance in relative abundance data was p=0.01. Univariate comparisons were conducted by using One-way Analysis of Variance (ANOVA) and Student's t-test on the absolute abundance of amino acid, using SAS JMP pro11 (SAS Institute Inc., SAS Campus Drive, Cary, NC, USA). The cutoff of statistical significance in absolute abundance data was p=0.05. SigmaPlot version 11.0 (Systat Software, San José, CA, USA) was used to make graphs.

4.4. Results

4.4.1. Abundance of amino acids

Overall, amino acid abundance of Haast site was higher than that of Michigan site as total organic matter content was approximately ~20% higher in Haast site than Michigan site (Fig. 4.3). About 12 times more of amino acid in the whole soil extracts and about 18 times more in the mineral associated extracts were found in Haast site

compared to Mighigan site. The proportion of the mineral associated amino acid in the whole soil was also higher in Haast site, ranged from 12% to 50% of amino acid, while from 8% to 22% of amino acid was associated with mineral in Michigan site. However, in both sites, the mineral associated amino acid had low variations in abundance during the year of development, whereas the abundance of amino acid in whole soil extracts varied with dune age.

The abundances of Ornithine (non-protein amino acid) showed correlation with total abundance of proteinogenic amino acids (r^2=0.5374, $p<0.0001$). The abundance of Orn in Haast site was greater than that in Michigan site. However, Orn abundance relative to amino acid abundance in Michigan site was two to three folds higher than that in Haast site. Compared to Orn in whole soil extracts, Orn was about six times and eight times enriched in mineral associated fractions in Michigan and Haast site respectively.

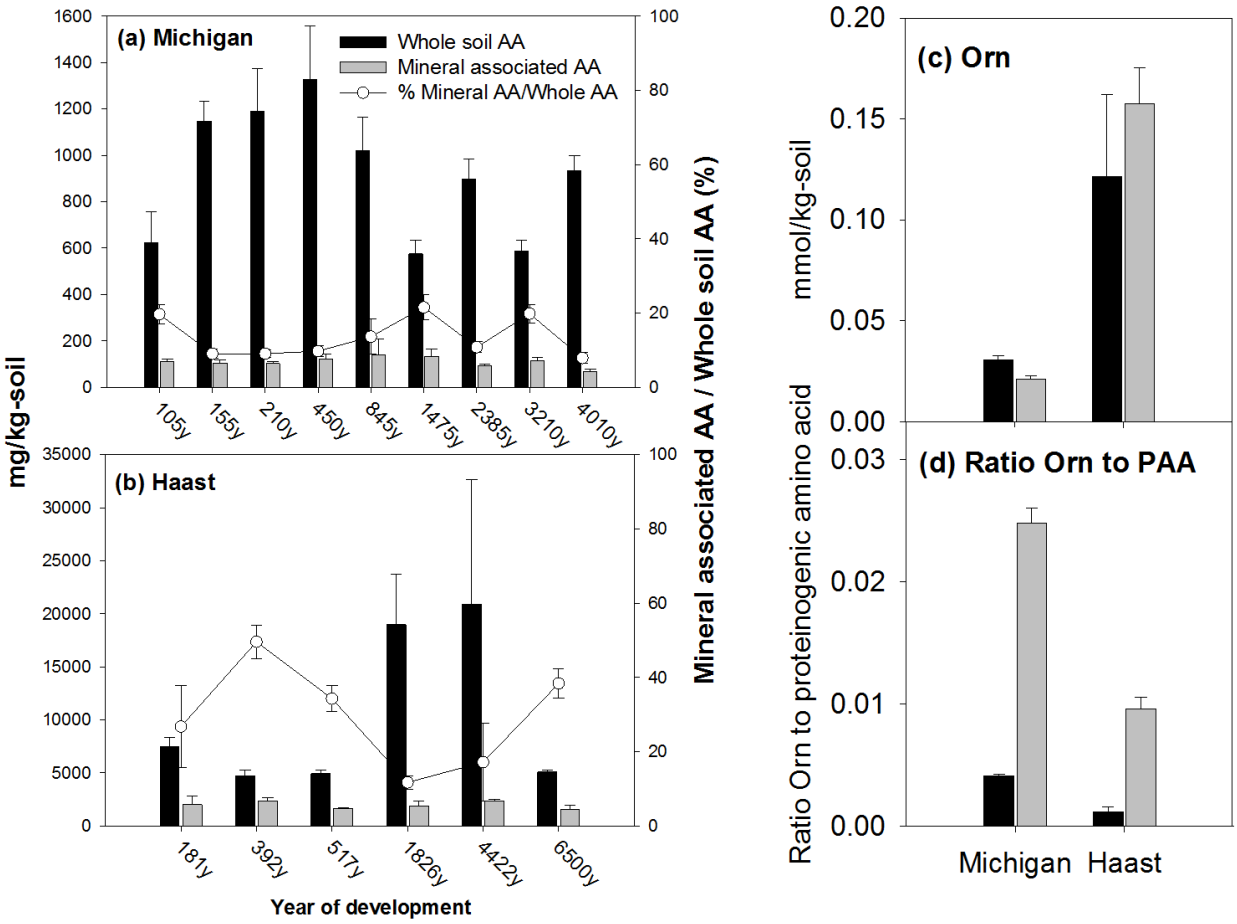


Figure.4. 3 Absolute amount of amino acid in whole soil extract (black bar), mineral associated fraction (grey bar), and the proportion of mineral associated amino acid (open circle and line) in Michigan site **(a)** and in Haast site **(b)**. Absolute amount of non- protein amino acid, Ornithine (Orn) **(c)**, and ratio of Orn to total proteinogenic amino acid **(d)**. Error bars represent standard error (n=5 for Michigan **(a)** and n=4 for Haast **(b)**; and n=45 for Michigan, n=24 for Haast in **(c)** and **(d)**).

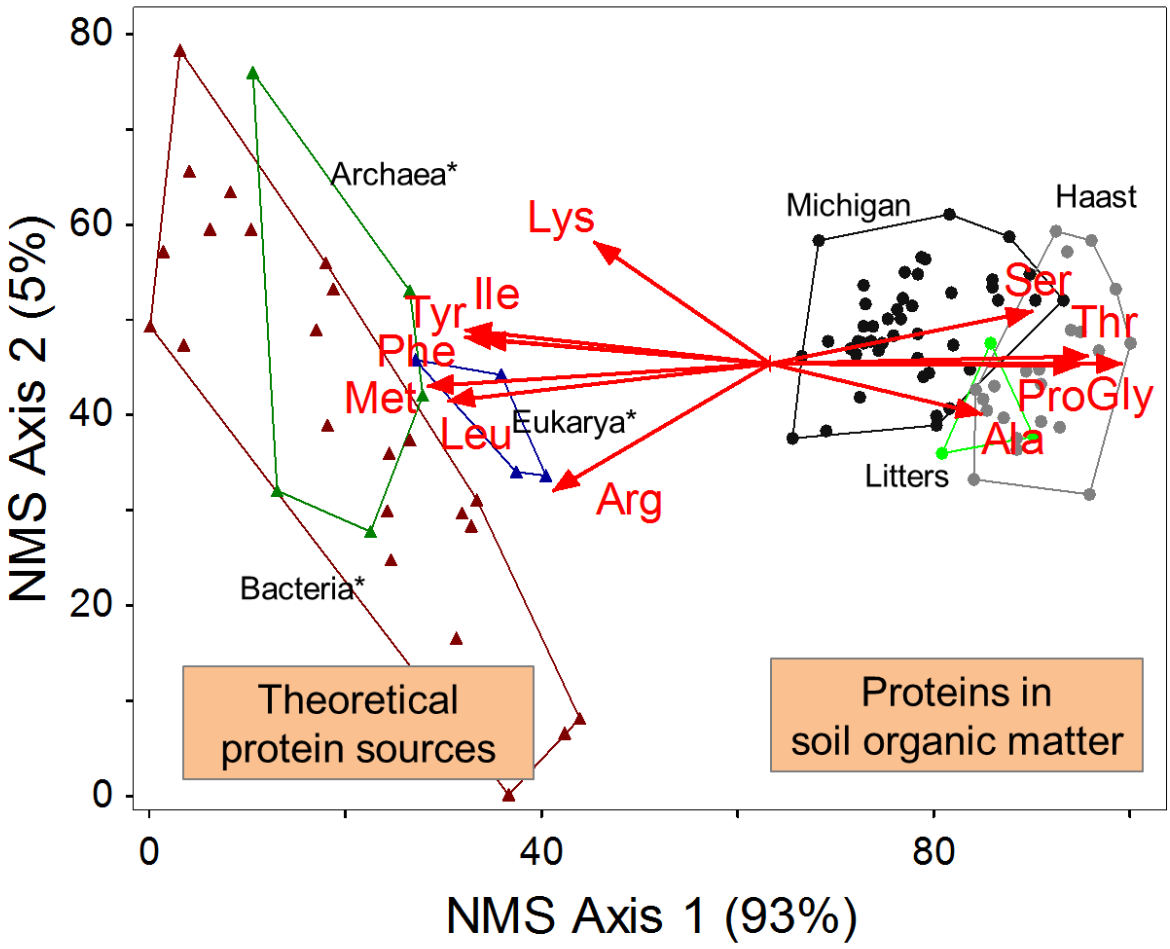


Figure.4. 4. Comparisons of amino acid distribution between theoretical biological sources and soil organic matters from Michigan and Haast chronoseuqnces. Nonmetric multidimensional scaling (NMS) ordination plot of the relative distribution of 17 proteinogenic amino acids. Black circle (●) is amino acid distribution in whole soil OM from Michigan site; grey circle (●) is from Haast site; neon green (●) is amino acid distribution in litters collected from 1826y dune in Haast site; *Brown triangle (▲) is amino acid distribution of bacteria; *green triangle (▲) is of archaea; *blue triangle (▲) is of eukarya, based on NCBI genome database (Chen et al., 2013). Correlations of variables with ordination with r2>0.5 were shown in bi-plot vector (**red arrow**) where length and direction represent the magnitude and directions of the correlation, respectively. Percentages on each axis denote the amount of variability associated with each axis.

4.4.2. Composition of amino acids

The distributions of amino acid in whole soil extracts in both Michigan and Haast sites were relatively similar to each other. Amino acid distribution in whole soil OM differed from the theoretical amino acid distribution of organisms including three taxonomic domains: bacteria, archaea, and eukarya (Fig.4.4 and Table.4.1). The simplest amino acids (Gly and Ala), neutral polar amino acids (Ser and Thr), and Pro were preferentially accumulated in soil. Gly, for example, was approximately twice more abundant in whole soil OM compared to theoretical proteins. On the contrary, Phe, Tyr, Leu, Ile, Met, Lys, and Arg were associated more in theoretical biological sources. In comparison to the biological sources based on genomic database, the relative abundances of amino acids in soil have shown to be shifted. The amino acid patterns in the whole soil OM fractions and the litter from Haast site were relatively uniform (MRPP A=0.0673, p=0.0290). Since amino acid pattern in litters was similar to that in whole soil OM at 181y, rather than 1826y where actually the litters were collected, the distribution of amino acid is subjective to change over time with the limited variation.

The amino acid distributions in the mineral associated fractions were distinct to those in the whole soil extracts in both sites (Fig. 4.5). Compared to the mineral associated fractions, the distributions of amino acid in the whole soil extracts had relatively smaller variations; they were somewhat overlapped between two sites. However, the amino acid distributions in the mineral associated fractions were relatively more distinct than those of whole soil pool between two sites. In comparison between whole soil and mineral associated extracts, the magnitude of variations in the distribution of amino acid with year of development was opposite to that in the

abundance of amino acid. Although the abundance of amino acid in the mineral associated fractions was consistent across the year of development (Fig. 4.3), relatively higher variations in the distribution of amino acid during the ecosystem development were pronounced in mineral associated fractions than those in whole soil OM fractions (Fig.4.5). On the other hand, even if there were great variations in the abundance of amino acid with age in whole soil OM pools, the distributions of amino acid have shown to have less variation compared to mineral associated fractions.

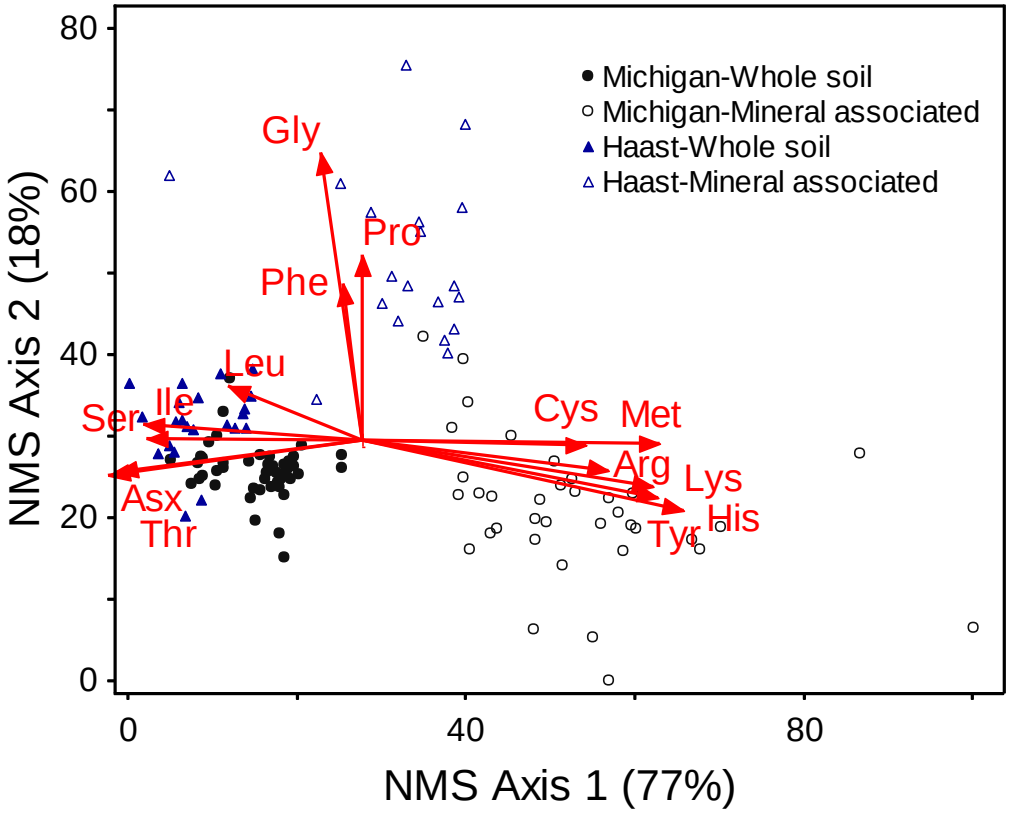


Figure.4. 5. Comparisons of 17 proteinogenic amino acid distribution in whole soil and mineral associated OM fractions in Michigan and Haast chronosequences, plotted by nonmetric multidimensional scaling (NMS) ordination. Correlations of variables with ordination with r^2>0.3 were shown in bi-plot vector (red arrow) where length and direction represent the magnitude and directions of the correlation, respectively. Percentages on each axis denote the amount of variability associated with each axis

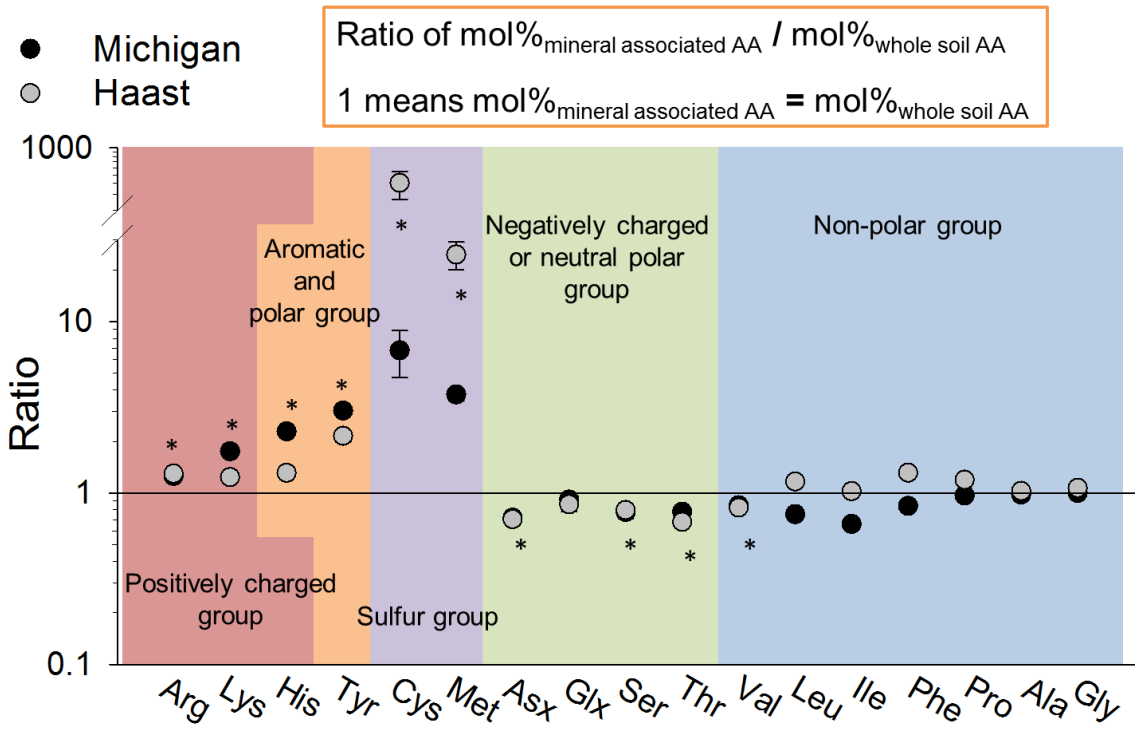


Figure.4. 6. Ratio of mineral associated amino acids to whole soil amino acids from Michigan site (black circle) and Haast site (grey circle). Ratio of each amino acid was calculated by dividing mol% of mineral associated amino acid by mol% of whole soil amino acid. Thus, ratio value one indicates the equality of mol% of mineral associated amino acid and mol% of whole soil amino acid. Ratio value higher than one is indicative of enrichment of amino acid on the mineral associated fraction, while Ratio value lower than one indicates the depletion of amino acid on the mineral associated fraction compared to the whole soil OM hydrolysate. "*" indicates the ratios that obtain common trends between two sites (both ratios are higher than one or both are lower than one) and the both ratios are significantly different from one tested by t-test ($p<0.05$). Amino acids were separated into five groups: (1) positively charged group, (2) aromatic and polar group, (3) sulfur group, (4) negatively charged or neutral polar group, and (5) non-polar group. His belongs to both (1) and (2).

4.4.3. Mineral associated vs. whole soil amino acids

Majority of the amino acid (eleven out of seventeen amino acids) have shown the similar trends either enrichment or depletion on mineral surfaces from both sites (Fig. 4.6 and Fig.4.7). The common enriched amino acid on mineral associated fractions in both Michigan and Haast sites were positively charged amino acid (His, Lys, and Arg) and amino acids containing aromatic and hydroxyl side chain group (Tyr and His), and sulfur group amino acids (Met and Cys). On the other hand, the relative abundance of

negatively charged group amino acid (Asx and Glx), neutral polar goup (Ser and Thr), and Val were lower in the mineral associated fractions in both sites compared to the whole soil OM fractions. Most of the amino acids that showed the common trends associated with mineral in both sites were related to polar interactions and redox reactions (Brosnan and Brosnan 2006). Amino acids seemed to selectively associate with mineral surfaces and the physico-chemistry of amino acids may be related to interaction with mineral.

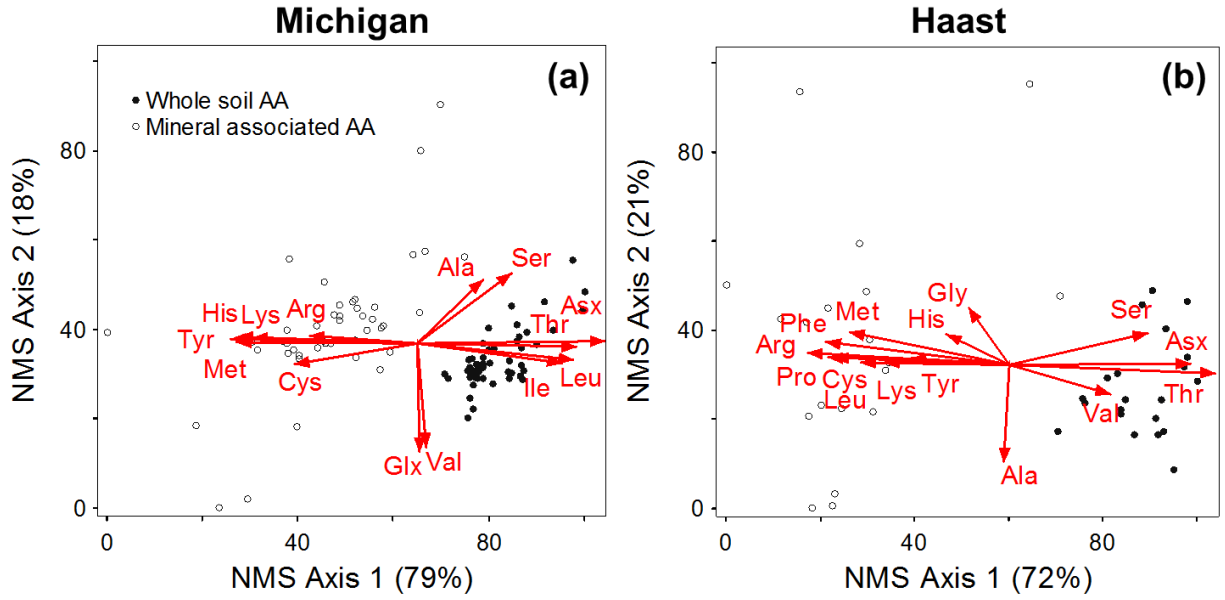


Figure.4. 7. Comparisons of 17 proteinogenic amino acid distribution between whole soil and mineral associated extracts in Michigan (a) and Haast sties (b), plotted by nonmetric multidimensional scaling (NMS) ordination. Correlations of variables with ordination with r^2>0.3 were shown in bi-plot vector (red arrow) where length and direction represent the magnitude and directions of the correlation, respectively. Percentages on each axis denote the amount of variability associated with each axis.

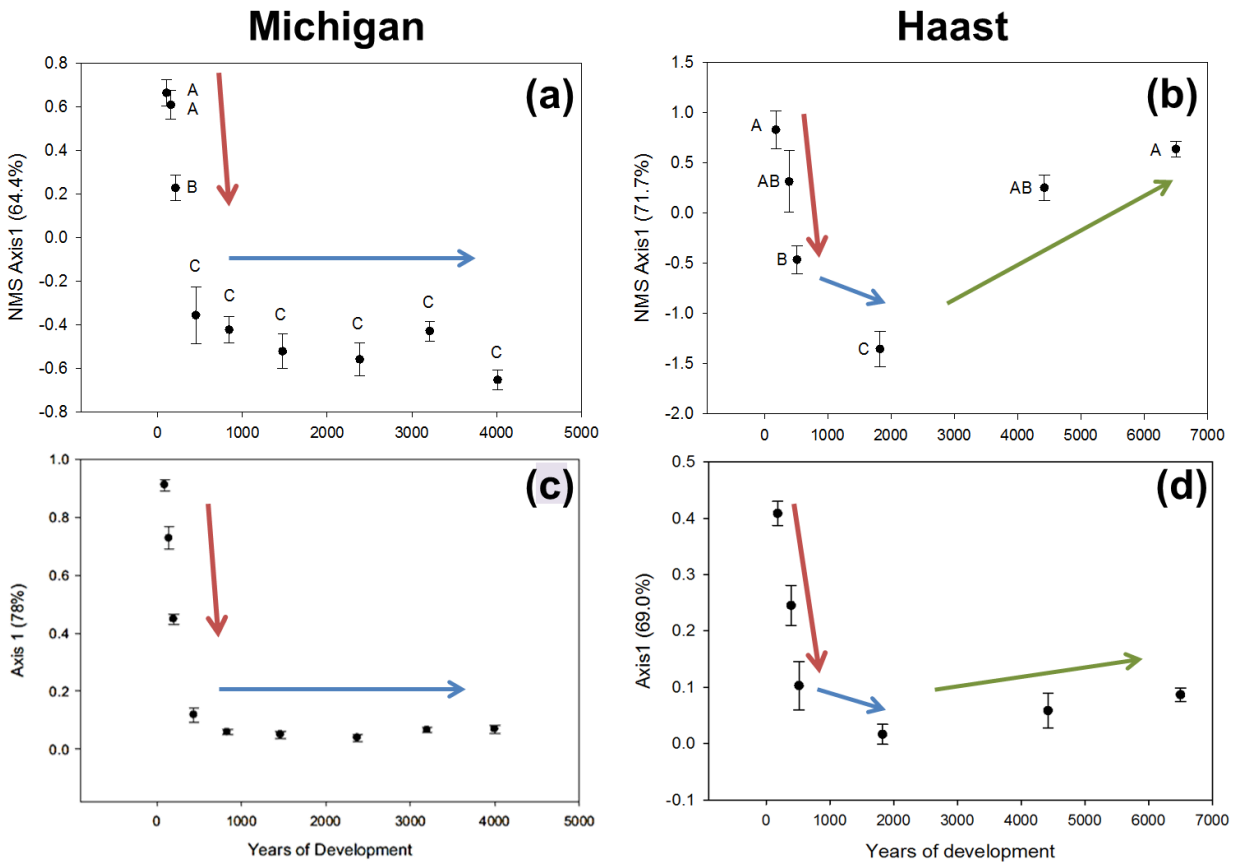


Figure.4. 8. Comparison between the changes of amino acid distributions and the changes of bacterial community distributions by year of development in Michigan and Haast chronosequences: (a) the relationship between Axis1 (r2= 64.4%) from 2-dementional Nonmetric multidimensional scaling (NMS) ordination of 17 protein amino acid relative distribution from whole soil hydrolysate and the year of development in Lake Michigan chronosequence; (b) the relationship between Axis1 (r2= 71.7%) from 2-d NMS ordination of 17 protein amino acid relative distribution from whole soil hydrolysate and the year of development in Haast chronosequence; (c) the relationship between Axis1 (r2=78.0%) from Bray-Curtis ordination on bacterial community structure based on relative proportion of 200 most abundant OTUs (Williams et al., 2013); (d) the relationship between Axis1 (r2=69.0%) from Bray-Curtis ordination on bacterial community structure based on relative proportion of 120 most abundant OTUs (Jangid et al., 2013). Error bars represent standard error (n=10 for (a) and (c); n=4 for (b) and (d)). Comparison of Axis1 scores among different years of development was conducted by Tukey test and letters denote significant difference (P< 0.05) in (a) and (b). Colored arrows denote concurrent shifts in amino acid distribution and bacterial communities with progressive (red), steady (blue), retrogressive (green) stages of ecosystem development

4.4.4. Relationship between dynamics of amino acid distribution and bacterial community composition

Both Michigan and Haast sites showed the patterns of the change of amino acid distribution mimicked their patterns of bacterial community change with years of development (Fig4.8.a and c; b and d). In the Michigan site, two major trends were apparent in both amino acid and bacterial community compositions. (1) From 155y to 450y, the shifts have shown to be relatively dynamic, but (2) less varying after 450y and during thousands of years at the later stage of development. However, due to relatively faster weathering process, in the Haast site, three major trends were appeared in plots of both amino acid and bacterial community composition. There were (1) rapid and steep shift was shown during several hundreds of years from 181y to 517y, (2) the change during a few thousands of years of development after 517y was relatively small, and (3) between 1826y and 2200y to 6500y, the change patterns appeared to reverse to the earlier trends. In other words, the y axis scores of the first and second trends are decreasing, but the scores of the third trends are increasing. Although the shifts of bacterial community compositions were shown to be similar to amino acid distribution changes in the Haast site, the third trend of bacterial community was less pronounced than amino acid distribution. Bacterial community compositions of 4422y and 6500y were similar to those from 1826y, while the amino acid distributions of the same time periods were rather similar to those from earlier stage of development (181y-392y) than those from 1826y. Overall the distribution change in amino acid appeared to parallel the change in belowground bacterial community composition with pedogenic progress related to ecosystem development.

4.4.5. Pedogenic patterns of amino acid distribution

There was no clear pattern with age in common between two sites. The relative distributions of amino acid in the whole soil OM fractions changed with age to the different direction and pattern between Michigan and Haast sties, shown as blue and red arrows in Fig.4.9.a. Like whole soil OM fractions, the amino acid distributions in the mineral associated OM fractions were separated between Michigan and Haast sites, as well as the change patterns in amino acid distribution exhibited the opposite directions (Fig.4.9.b). It is notable that the dynamic in amino acid distributions associated with mineral appeared to be more conspicuous during the late ecosystem development as tectosilicates primary minerals were weathered slowly and the changes of OM associated with mineral were reflected at the later stage of pedogenesis in our study sites. The amino acid distribution in 4010y was different from those in younger sites in Michigan chronosequence. Likely, the dynamic of amino acid distribution began to appear considerably from 4422y gradually thereafter. The dynamic of amino acid distribution related to ecosystem development have shown to be rather completed with the difficulty to compare two geologically separated and climatically very distinct ecosystems.

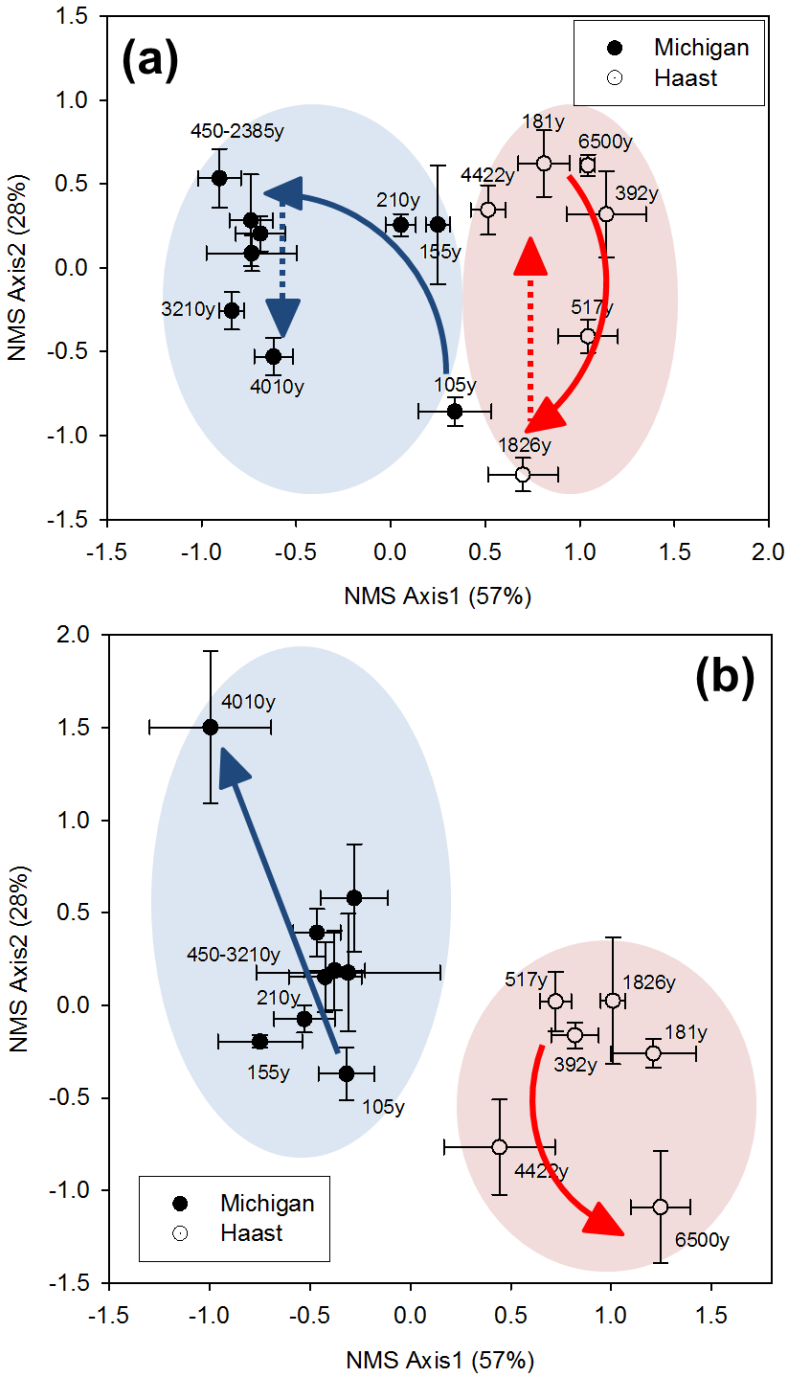


Figure.4. 9. The directions of change in17 proteinogenic amino acid distribution with year of development in Michigan (blue cluster) and Haast (red cluster) sites, comparing within the same pools: whole soil (a) and mineral associated (b) extracts, plotted by nonmetric multidimensional scaling (NMS) ordination. The blue arrows indicate the direction of the increase in age in Michigan site and red arrows are for Haast site. Percentages on each axis denote the amount of variability associated with each axis

4.5. Discussion

The broad findings of this study indicate that there are predictable and recurrent patterns of SOM composition at two independent and ecologically distinct ecosystems in USA and New Zealand. The research, furthermore, provide new evidence of SOM formation in support of a mixed pool of plant and microbial derived organic matter that has gone through the process of selective fractionation and preservation. In large part, the whole soil OM pools strongly resemble that of plant litter and thus suggest that much of the proteins and amino acids in soil remained unchanged during litter breakdown. The mineral associated organic matter, in contrast, showed evidence for strong selection of positively charged, aromatic, and sulfur containing proteinogenic amino acids. The consistency of the results at two disparate locations in the southern and northern hemispheres is strong evidence that the processes of pedogenesis and ecosystem development are parsimonious and predictable.

4.5.1. Bacterial contribution to SOM formation

To determine the bacterial contributions to organic matter, ornithine (Orn), a non-proteinogenic amino acid, was determined and found to change during the process of pedogenesis and ecosystem development. Non-protein biomarkers, such as hydroxyproline have been used trace plant derived organic matter into soil pools (Saharinen & Schnitzer, 1989, Sowden *et al.*, 1977, Szajdak *et al.*, 2003, Szajdak & Österberg, 1996) but fewer studies have used non-protein amino acids. Orn, a molecule sometimes found in bacterial cell wall peptidogylcans and ornithine-containing lipids (Nelson *et al.*, 2008, Ratledge & Wilkinson, 1988) have been used. These authors suggest that microbial origins to SOM formation were greater than those of

plants. In this dissertation, it was found that the bacterial contribution to SOM might be relatively greater in low SOM ecosystems like those found at the Michigan site. In addition, Orn was highly enriched in the soil mineral associated fractions in both ecosystems, a result consistent with the conceptual model of microbial cell wall debris stabilization on mineral surface (Miltner *et al.*, 2012). The positively charged side chain group of Orn might facilitate its adsorption to mineral surfaces through electrostatic forces (Aufdenkampe *et al.*, 2001, Keil *et al.*, 1998, McBride, 1994).

Arguments suggesting that microbial proteinaceous amino acids play an important role in SOM formation (Guggenberger *et al.*, 1999, Liang & Balser, 2011, Lichtfouse *et al.*, 1995), The results from the research in this dissertation provide new evidence that other mechanisms, including selective association with mineral surfaces, and litter inputs also appeared to be equally or more important than a model that emphasizes direct incorporation of microbial amino acids and proteins.

4.5.2. Origins and transformation of amino acid in soil

It still unclear that the proteinaceous compounds in soil are derived from incompletely decomposed plant residues or non-living microbial biomass. By comparing the relative distribution of amino acids, there was no resemblance of proteinaceous compounds in whole soil extracts to proteins from bacterial origins *per se*. In this regard, the results are consistent with those of Friedel and Scheller (2002), showing that soil and its microbial community composed of different proteinogenic amino acid fingerprints. The result from the comparison between SOM and theoretical protein origins (Fig.4.4), furthermore, showed that the amino acid distribution in whole soil OM extracts were slightly closer to Eukarya domain including plant and fungi rather than Bacteria domain

with the similarity of relatively high abundance in Thr, Ser, and His. The similarity in amino acid distribution between whole soil OM fraction and litters is, furthermore, consistent to the finding from Friedel and Scheler 2002, where the amino acid distribution of leaf litters was alike to that in SOM in mineral soil. This might indicated that the shifts in distribution of amino acid happen at the early stage of litter decomposition and assimilated amino acids during the litter decomposition migrate down to the mineral layer. Otherwise, the decomposition processes of leaves on the soil surface and roots in subsoil can lead the similarity in amino acid distribution through biogeochemical cycling. We have shown a strong evidence of shift in amino acid distribution, suggesting the selectivity in amino acid accumulated in soil. However, the assumption that we take here is that protein profiles based genome database will have high correlation with proteins potentially expressed. It is thus noted that the caution is needed to interpret the comparison between theoretical protein sources and proteins in soil, which the protein profiles based on genome database do not reflect the physiology of organisms in the soil habitat where the large heterogeneity and rapid nutrient dynamics exist.

4.5.3. Selection for amino acid associated with minerals

Comparing two geologically separated and climatically very different chronosequences was somewhat advantageous for this study because the similarity of soil textures and parent materials. Although the two have very distinct aboveground vegetative communities, the dominant bacterial taxonomic groups belowground and the trends of their change have in common. We expected to see very differences in amino acid dynamics in two sites; however, we also expected to find some commonality in

relatively younger developed and undisturbed soil ecosystems. Due to the quartz dominated parent materials, permanent negatively charges by isomorphic substitution are predominant. Thus, the preferential accrual of positively charged amino acids on mineral associated fraction in both sites does make sense. Negatively charged amino acid, in contrast, depleted on the mineral surface, possibly due to the repulsion by negative charge.

In addition, other amino acids that had common trends in association with minerals were gained attention. Cys and Met containing sulfur side chain group play important roles in binding to metal ions and redox reactions (reductant) on the active sides exposing on the surface of proteins (Brosnan and Brosnan 2006, Russell et al., 2003). Cys is more reactive than Met due to hydrogen atom connected to sulphur atom in Cys. Their enrichment on mineral surfaces could be achieved through the selective interaction between metal ion on the mineral surface and these amino acids. Similarly, His is commonly found in metal binding motifs. The protons of His can be transferred on and off easily and this is ideal for charge relay systems, such as those found within catalytic triads in proteases. His, on top of that, is one of amino acids containing aromatic ring, which is involved in stacking interactions with other aromatic side chains. As weathering progressed, aromatic compounds accumulated on mineral surfaces through reactive OH sites such as short-range order minerals (Kramer *et al.*, 2012). Because non-crystalline and secondary mineral formation has little or not occurred in both Michigan and Haast sites, however, the interaction between aromatic ring and hydroxyl group is less likely responsible for accumulation of His and Tyr in mineral associated fractions. It is notable that amino acids that contain both aromatic ring and

hydrophilic side chain such as His and Tyr were preferentially associated with minerals. There is possibility of important roles of these amino acids in binding to minerals.

Neutral polar amino acids (Ser and Thr) were less associated with mineral associated fractions in both sites. This can be explained by two reasons. One is due to their hydrophilic nature and strong interaction with water molecules, favoring their presence in soil solution. This was supported by abundance of neutral amino acids in soluble pools in chapter3 (3.4.4). The other reason is that the selective preservation of these amino acids possibly by interaction with organic aggregates through hydrogen bonds may restrict their chemical accessibility (Ahmed *et al.*, 2015, Schulten & Schnitzer, 1997, Senesi *et al.*, 2009). Allard (2006) has shown the evidence of increasing neutral polar amino acids in lignite deposit where organic matters are preserved for a long time. This indicates that non-charged polar amino acids may undergo the preservation pathway to interaction with other organic matter rather than association with mineral components.

In spite of the commonality associated with minerals in two sites, the proportion of amino acids associated with mineral in whole soil amino acids was a lot higher in Haast sites in general (Fig.4.1). This means that minerals in Haast site have greater capacity to retain amino acids either due to larger surface area of mineral or due to the adsorption of micro-aggregates or multiple layers of organic matters to mineral. The latter was suggested by onion layering conceptual model (Sollins *et al.*, 2006), where multiple layers of organic matters are constructed by initiative inner organic layer around mineral surface through electrostatic interaction and then less polar organic layers on top of the inner layer via non-polar interactions. The NMS bi-plot in Fig4.5 showed the

distribution of mineral associated amino acids from the Haast site in between whole soil amino acid and mineral associated amino acids from the Michigan site. This might be reflected by that the amino acids accumulated on mineral particles become more similar in distribution to amino acids represented in whole soil pool, where organic compounds are more likely stabilized within the form of aggregates. This also implies the difficulty to differentiate the characteristics of amino acids intact with mineral surfaces from those associated with outer organic layers in soil obtaining such high sorption capacity.

4.5.4. Selection for amino acid in relation to life strategy of soil microbes

The tight relationship between amino acid dynamics and microbial community change implies the potential for direct and indirect input of microbial community to amino acid turnover. As described in the chapter2, phylum level of microbial community change possibly explain some of shift patterns of amino acid distribution. In addition to microbial contribution to amino acid dynamics over the long term, the characteristics of different amino acid distribution can, alternatively, explain some of bacterial community change as a result of their life strategies and ecological functions in response to shifting available organic matter pools. Bacterial phyla that shifted along with the ecosystem development were relevant to the ecological classifications based on life-strategies, either r- or K-strategists. In both sites, the relative abundance of oligotrophic taxa (mainly Acidobacteria) increased during the ecosystem development (r^2=0.66, p=0.007 for Michigan site ((Williams *et al.*, 2013) and r^2=0.56, p=0.08 for Haast site (Jangid *et al.*, 2013)). Copiotrophic taxa (Actinobacteria, Bacteroidetes, Betaproteobacteria, and Fimicutes), in contrast, relatively decreased with year of development; for example, Bacteroidetes was negatively correlated with age of sites in both sites (r^2=0.67, p=0.007

for Michigan site ((Williams *et al.*, 2013) and r^2=0.85, p=0.008 for Haast site (Jangid *et al.*, 2013)). With ecosystem development, in general, organic compounds with long residence time relatively accumulate and they are often characterized as chemically and physically recalcitrant pool with limited accessibility. At the later developed site, therefore, the available OM pool size relatively declines which accords with the accumulation of His in both sites. His, particularly, seems to be important indicator of stabilized OM despite the fact that it is among the minor amino acids (Stevenson, 1956). The relative distribution of His fitted well with ecosystem development dynamic model (Fig4.10). It is unclear how His is involved in the stabilization mechanisms, but it has potential to play significant roles in preservation of proteinaceous compounds in soil. Its reactivity to mineral surfaces and metal ions is certainly one possible reason. Its amphiphilic side chain groups, furthermore, allow occurring in buried and surface moieties of protein three dimensional structures, which is assumed to have possibility of various interactions related to its persistence.

4.6. Conclusions

The broad findings of this study indicate that there are predictable and recurrent patterns of SOM change that show consistency between two ecologically discrete ecosystems. We found the consistency of selecting patterns of proteinaceous compounds by two disparate locations in the three major ways: (1) similarity of proteinogenic amino acid fingerprint in whole soil pools, (2) resemblance of strong selection of proteinogenic amino acid by mineral associated fractions, and (3) simultaneous change patterns of proteinogenic amino acids along with biological community successions. The research, furthermore, provide new evidence of SOM

formation in support of a mixed pool of plant and microbial derived organic matter that has gone through the process of selective fractionation and preservation. In large part, the whole soil OM pools strongly resemble that of plant litter and thus suggest that much of the proteins and amino acids in soil remained unchanged during litter breakdown. The consistency of the results at two disparate locations in the southern and northern hemispheres is strong evidence that the processes of pedogenesis and ecosystem development are parsimonious and predictable.

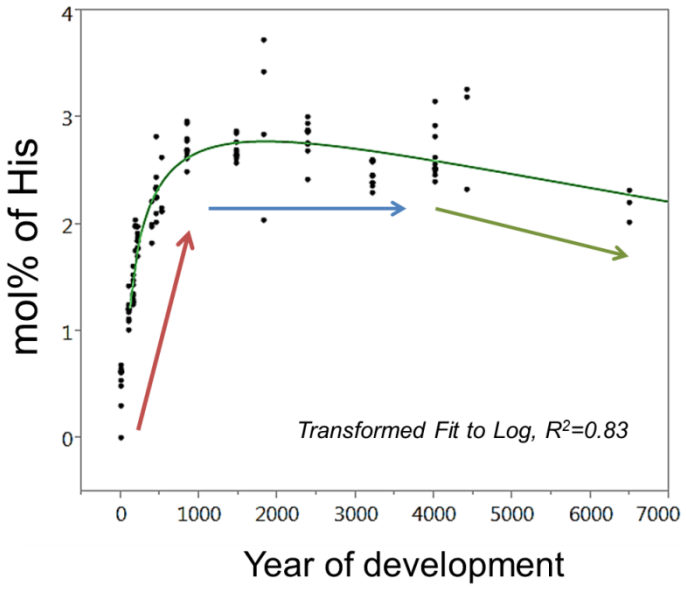


Figure.4. 10. Mol% change of His with year of development combined Michigan and Haast sites (P<0.0001). Colored arrows denote concurrent shifts in amino acid distribution and bacterial communities with progressive (**red**), steady (blue), retrogressive (**green**) stages of ecosystem development

References

Ahmed AA, Thiele-Bruhn S, Aziz SG *et al.* (2015) Interaction of polar and nonpolar organic pollutants with soil organic matter: Sorption experiments and molecular dynamics simulation. *Science of the Total Environment,* **508**, 276-287.

Allard B (2006) A comparative study on the chemical composition of humic acids from forest soil, agricultural soil and lignite deposit: Bound lipid, carbohydrate and amino acid distributions. *Geoderma,* **130**, 77-96.

Amelung W, Zhang X (2001) Determination of amino acid enantiomers in soils. *Soil Biology and Biochemistry,* **33**, 553-562.

Amelung W, Zhang X, Flach KW (2006) Amino acids in grassland soils: Climatic effects on concentrations and chirality. *Geoderma,* **130**, 207-217.

Aufdenkampe AK, Hedges JI, Richey JE, Krusche AV, Llerena CA (2001) Sorptive fractionation of dissolved organic nitrogen and amino acids onto fine sediments within the Amazon Basin. *Limnology and Oceanography,* **46**, 1921-1935.

Bosch L, Alegría A, Farré R (2006) Application of the 6-aminoquinolyl-N-hydroxysccinimidyl carbamate (AQC) reagent to the RP-HPLC determination of amino acids in infant foods. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences,* **831**, 176-183.

Chappell J, Shackleton NJ (1986) Oxygen isotopes and sea level.

Coomes DA, Bellingham PJ (2011) Temperate and tropical podocarps: how ecologically alike are they. *Ecology of the Podocarpaceae in tropical forests*, 119-140.

Cotrufo MF, Wallenstein MD, Boot CM, Denef K, Paul E (2013) The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Global Change Biology,* **19**, 988-995.

Curry GB, Theng BK, Zheng H (1994) Amino acid distribution in a loess-palaeosol sequence near Luochuan, Loess Plateau, China. *Organic Geochemistry,* **22**, 287-298.

Ding X, Henrichs SM (2002) Adsorption and desorption of proteins and polyamino acids by clay minerals and marine sediments. *Marine Chemistry,* **77**, 225-237.

Eger A, Almond PC, Condron LM (2011) Pedogenesis, soil mass balance, phosphorus dynamics and vegetation communities across a Holocene soil chronosequence in a super-humid climate, South Westland, New Zealand. *Geoderma,* **163**, 185-196.

Espeland E, Wetzel R (2001) Complexation, stabilization, and UV photolysis of extracellular and surface-bound glucosidase and alkaline phosphatase: implications for biofilm microbiota. *Microbial Ecology,* **42**, 572-585.

Fan TM, Lane A, Chekmenev E, Wittebort R, Higashi R (2004) Synthesis and physico-chemical properties of peptides in soil humic substances. *The Journal of peptide research,* **63**, 253-264.

Friedel JK, Scheller E (2002) Composition of hydrolysable amino acids in soil organic matter and soil microbial biomass. *Soil Biology & Biochemistry,* **34**, 315-325.

Galloway JN, Dentener FJ, Capone DG *et al.* (2004) Nitrogen cycles: past, present, and future. *Biogeochemistry,* **70**, 153-226.

Gibb J (1986) A New Zealand regional Holocene eustatic sea-level curve and its application to determination of vertical tectonic movements. *Royal Society of New Zealand Bulletin,* **24**, 377-395.
Guggenberger G, Frey SD, Six J, Paustian K, Elliott ET (1999) Bacterial and Fungal Cell-Wall Residues in Conventional and No-Tillage Agroecosystems. *Soil Science Society of America Journal,* **63**, 1188-1198.
Hou S, He H, Zhang X, Zhang W, Xie H (2009) Determination of soil amino acids by high performance liquid chromatography-electro spray ionization-mass spectrometry derivatized with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate. *Talanta,* **80**, 440-447.
Jangid K, Whitman WB, Condron LM, Turner BL, Williams MA (2013) Progressive and retrogressive ecosystem development coincide with soil bacterial community change in a dune system under lowland temperate rainforest in New Zealand. *Plant and Soil*, 1-13.
Kaiser K, Guggenberger G (2007) Distribution of hydrous aluminium and iron over density fractions depends on organic matter load and ultrasonic dispersion. *Geoderma,* **140**, 140-146.
Keil RG, Tsamakis E, Giddings JC, Hedges JI (1998) Biochemical distributions (amino acids, neutral sugars, and lignin phenols) among size-classes of modern marine sediments from the Washington coast. *Geochimica et Cosmochimica Acta,* **62**, 1347-1364.
Knicker H (2011) Soil organic N - An under-rated player for C sequestration in soils? *Soil Biology and Biochemistry,* **43**, 1118-1129.
Knicker H, Hatcher PG (1997) Survival of Protein in an Organic-Rich Sediment: Possible Protection by Encapsulation in Organic Matter. *Naturwissenschaften,* **84**, 231-234.
Knicker H, Hatcher PG (2001) Sequestration of organic nitrogen in the sapropel from Mangrove Lake, Bermuda. *Organic Geochemistry,* **32**, 733-744.
Kramer MG, Sanderman J, Chadwick OA, Chorover J, Vitousek PM (2012) Long-term carbon storage through retention of dissolved aromatic acids by reactive particles in soil. *Global Change Biology,* **18**, 2594-2605.
Küry D, Keller U (1991) Trimethylsilyl-O-methyloxime derivatives for the measurement of [6, 6-2 H 2]-d-glucose-enriched plasma samples by gas chromatography—mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications,* **572**, 302-306.
Leverett F, Taylor FB (1915) *The Pleistocene of Indiana and Michigan and the history of the Great Lakes*, US Government Printing Office.
Li X, Rapson G, Flenley JR (2008) Holocene vegetational and climatic history, Sponge Swamp, Haast, south-western New Zealand. *Quaternary International,* **184**, 129-138.
Liang C, Balser TC (2011) Microbial production of recalcitrant organic matter in global soils: implications for productivity and climate policy. *Nature Reviews. Microbiology,* **9**, 75;-author reply 75.
Lichter J (1995a) Lake Michigan beach-ridge and dune development, lake level, and variability in regional water balance. *Quaternary Research,* **44**, 181-189.

Lichter J (1998) Rates of weathering and chemical depletion in soils across a chronosequence of Lake Michigan sand dunes. *Geoderma,* **85**, 255-282.
Lichter JP (1995b) Mechanisms of plant succession in coastal Lake Michigan sand dunes. Unpublished 9612986 Ph.D., University of Minnesota, Ann Arbor, 252-252 p. pp.
Lichtfouse É, Berthier G, Houot S, Barriuso E, Bergheaud V, Vallaeys T (1995) Stable carbon isotope evidence for the microbial origin of C14□C18 n-alkanoic acids in soils. *Organic Geochemistry,* **23**, 849-852.
Matrajt G, Blanot D (2004) Properties of synthetic ferrihydrite as an amino acid adsorbent and a promoter of peptide bond formation. *Amino Acids,* **26**, 153-158.
Mcbride MB (1994) *Environmental chemistry of soils,* New York, Oxford University Press.
Mikutta R, Kaiser K, Dörr N *et al.* (2010) Mineralogical impact on organic nitrogen across a long-term soil chronosequence (0.3–4100 kyr). *Geochimica et Cosmochimica Acta,* **74**, 2142-2164.
Miltner A, Bombach P, Schmidt-Brücken B, Kästner M (2012) SOM genesis: microbial biomass as a significant source. *Biogeochemistry,* **111**, 41-55.
Nelson DL, Lehninger AL, Cox MM (2008) *Lehninger principles of biochemistry*, Macmillan.
Norman F, Boas F (1953) Method of determination of hexosamines in tissues. *J. biol. Chem*, 553-563.
Palmer RWP, Doyle R, Grealish G, Almond P (1985) Soil studies in South Westland 1984/1985.
Ratledge C, Wilkinson S (1988) Fatty acids, related and derived lipids. *Microbial lipids,* **1**, 23-52.
Rattenbury M, Jongens R, Cox S (2010) Geology of the Haast Area. Institute of Geological & Nuclear Sciences 1: 250,000 Geological Map 14. GNS Science Lower Hutt.
Saharinen M, Schnitzer M (1989) Nitrogen in finnish agricultural soil and it's humic acid compared to two Canadian soils. *Science of the Total Environment,* **81–82**, 459-463.
Schnitzer M, Kodama H (1992) Interactions between Organic and Inorganic Components in Particle-Size Fractions Separated from Four Soils. *Soil Science Society of America Journal,* **56**.
Schulten HR, Schnitzer M (1997) The chemistry of soil organic nitrogen: a review. *Biology and Fertility of Soils,* **26**, 1-15.
Senesi N, Xing B, Huang PM (2009) *Biophysico-chemical processes involving natural nonliving organic matter in environmental systems,* Hoboken, N.J, Wiley.
Sollins P, Swanston C, Kleber M *et al.* (2006) Organic C and N stabilization in a forest soil: Evidence from sequential density fractionation. *Soil Biology and Biochemistry,* **38**, 3313-3324.
Sowden FJ, Chen Y, Schnitzer M (1977) The nitrogen distribution in soils formed under widely differing climatic conditions. *Geochimica et Cosmochimica Acta,* **41**, 1524-1526.
Spurr SH, Zumberge JH (1956) Late Pleistocene features of Cheboygan and Emmet Counties, Michigan. *American Journal of Science,* **254**, 96-109.

Stevenson F (1956) Effect of some long-time rotations on the amino acid composition of the soil. *Soil Science Society of America Journal,* **20**, 204-208.
Szajdak L, Jezierski A, Cabrera ML (2003) Impact of conventional and no-tillage management on soil amino acids, stable and transient radicals and properties of humic and fulvic acids. *Organic Geochemistry,* **34**, 693-700.
Szajdak L, Österberg R (1996) Amino acids present in humic acids from soils under different cultivations. *Environment International,* **22**, 331-334.
Turner BL, Wells A, Andersen KM, Condron LM (2012) Patterns of tree community composition along a coastal dune chronosequence in lowland temperate rain forest in New Zealand. *Plant Ecology,* **213**, 1525-1541.
Wang X-C, Lee C (1993) Adsorption and desorption of aliphatic amines, amino acids and acetate by clay minerals and marine sediments. *Marine Chemistry,* **44**, 1-23.
Wells A, Goff J (2007) Coastal dunes in Westland, New Zealand, provide a record of paleoseismic activity on the Alpine fault. *Geology,* **35**, 731-734.
Williams MA, Jangid K, Shanmugam SG, Whitman WB (2013) Bacterial communities in soil mimic patterns of vegetative succession and ecosystem climax but are resilient to change between seasons. *Soil Biology and Biochemistry,* **57**, 749.
Zang X, Van Heemst JDH, Dria KJ, Hatcher PG (2000) Encapsulation of protein in humic acid from a histosol as an explanation for the occurrence of organic nitrogen in soil and sediment. *Organic Geochemistry,* **31**, 679-695.
Zimmerman AR, Goyne KW, Chorover J, Komarneni S, Brantley SL (2004) Mineral mesopore effects on nitrogenous organic matter adsorption. *Organic Geochemistry,* **35**, 355-375.

Chapter 5. Conclusions

Traditionally, intrinsic molecular structure is a major controller in the decomposition of soil organic matter (SOM). Biologically more resistant structures, (e.g. aromatic ring), are predicted to be preserved in soil relatively longer than less resistant structures, (e.g. peptide bond). As a result, the preferential accumulation of more degradation resistant compounds has been a leading hypothesis underpinning the formation of recalcitrant soil organic matter (SOM). However, recent evidence of molecular turnover suggests that the mechanisms of intrinsic recalcitrance of SOM may be primarily applicable to the initial stages of litter decomposition. Physicochemical protection mechanisms, in contrast, appear to play a strong role in the slow turnover of otherwise labile compounds in soil. In this regard, the chemical structure of the molecule is important not for recalcitrance to enzymatic alteration, but rather for its interaction with other molecules and mineral surfaces.

Proteinaceous molecules can be readily degraded by various proteinase enzymes in free solution and they had been, historically, thus predicted to have fast turnover rates in soil. However, persistence and slow turnover of proteinaceous compounds are observed almost ubiquitously in soil ecosystems regardless of environmental factors, such as climate, disturbance, and soil types.

We have two core working hypotheses; long term persistence of proteinaceous compounds is affected by (1) source and (2) sink. First, continuous recycling through microbial breakdown and resynthesis of proteinaceous compounds within the soil system, due to their essential cellular roles and metabolisms, may provide the potential

to select for proteinaceous compounds that are produced by residing microbes in the soil habitat. The source of proteinaceous compounds, thus, largely controls their abundance in soil. To test the source hypothesis, the relationship between biological successions and change of proteinaceous compounds was determined. Second, physical and chemical interactions of these compounds with mineral surfaces explain sink mechanisms and consequently their slow turnover rates. This was tested by comparing the compositional characteristics of proteinaceous compounds between mineral associated organic matter (OM) sub-pool and whole soil (bulk) OM pool. In this comparison, the individual amino acids containing various functional side chains provided what chemical interactions might be responsible for the selective distribution of proteinaceous compounds associated with mineral binding.

Major findings of the long-term dynamics of proteinaceous compounds supported the source and sink hypotheses. Based on the chronosequence approach study, the relative distribution of individual proteinogenic amino acids changed and showed clear patterns in the change during 4000 years of soil ecosystem development. Their distributional changes provided a long term view of the temporal dynamics of proteinaceous compounds that are relevant to pedogenic and ecosystem development time scales. Positively charged amino acid groups, as expected, sequentially increased their contribution to protein-associated SOM formation. These amino acids were, furthermore, observed to be relatively enriched in the mineral associated fraction; a result that is consistent with the occurrence of minerals dominated by negative charges. Interaction of positively charged amino acid groups with negatively charged mineral

surfaces support the idea of their preferential accumulation in this soil sink during 4000 years of ecosystem development.

The long-term accumulation patterns of proteinogenic amino acids were also tightly linked with the shifts in their biological sources, namely the aboveground vegetative community (r^2=0.66, p<0.0001) and the belowground microbial community (r^2=0.71, p<0.0001). These two major biological source groups may influence the colonization of each other during ecosystem development, so their effects on source on proteinaceous compounds are not always easy to separate from one another. However, the mixed pools of sources in these sites, substituted for time, provides site specific biological source material; for example, mixed pools of plants and microbes at a 105y site are different from those at a 450y site. This was supported by results, providing the possibility of the use of the proteinogenic amino acids as indicators of SOM formation. In support of the main hypotheses, both biological inputs and minerals played a role as sources and sinks of proteinaceous compounds respectively.

We also found that seasonal changes of proteinogenic amino acids were very dynamic, and at the same time independent to the 4000 year-pedogenic patterns, although the belowground bacterial community remained consistent between seasons. The seasonal variations in whole soil OM pool were relatively larger (49% out of total 94% variation in NMS bi-plot) than those in mineral associated OM sub-pool (22% out of total 92% variation in NMS bi-plot). Nevertheless, the seasonal changes in proteinogenic amino acids associated with mineral, surprisingly, were more dynamic than expected. Mineral associated OM has been shown to be a slow pool where exchange and cycling of OM is limited by mineral protections. However, results suggest some level of

dynamics in the displacement of proteinaceous compounds on mineral surfaces between seasons. In comparison between whole soil pool and mineral associated sub-pool, serine (Ser), one of amino acids, showed a strong accumulation in the summer of whole soil pool, and was relatively abundant in the winter of mineral associated sub-pool. The opposite trends were found for glutamic acid+glutamine (Glx). The seasonal change in relative abundance of these amino acids in turns between the two pools indicated that the mineral interaction also played a role as sources of proteinaceous compounds to some degree as well as roles as sinks. However, these findings need further investigation to understand the replenishment mechanisms among the pools.

Lastly, we found the consistency of selecting patterns of proteinaceous compounds of two disparate locations in three major ways: (1) similarity of proteinogenic amino acid fingerprints in whole soil pools, (2) resemblance of strong selection of proteinogenic amino acid by mineral associated fractions, and (3) simultaneous change patterns of proteinogenic amino acids along with biological community successions. Despite a largely distinct climate and plant community in these two locations, Michigan and Haast,, the similarity in transformation of sources to whole soil pools in the two locations provided evidence that a mixed pool of plant and microbial derived OM that has gone through the process of selective preservation, enriching small and simple structured amino acids (glycine (Gly), alanine (Ala), Ser, and aspartic acid+asparagine (Asx)). The silicate mineral associated fractions showed evidence for a strong selection of positively charged, aromatic, or sulfur containing amino acids (sink selection). With soil ecosystem development, both locations showed that the long-term accumulation patterns of proteinogenic amino acids were closely

related with shifts in their biological sources. Again, biogeochemical processes may create uniform compositions of amino acid in soil from a broad range of ecosystems (primary common selection), but the effect of ecosystem development coinciding with transition of biological sources might be minor yet enough to make significant variations in shifting amino acid compositions (secondary source selection).

Knowing how the various functional side chains of proteinaceous compounds are individually related to their interactions with soil components and describing the dynamics of proteinogenic amino acids during pedogenesis in different locations provides insight into their turnover over the long term and clues to the mechanisms of their selection controlled by source and sink. The molecular species approach of proteinaceous compounds helps explain the ubiquitous phenomena of their accrual in soil and their partitioning mechanisms associated with mineral and whole SOM. This research demonstrates a fundamental understanding of behavior of proteinaceous compounds at the molecular species level, and further provides possible mechanisms of their matrix protection. The research can be improved by determining the relationship mineralogy and proteinogenic amino acid distribution along the soil horizons and turnover rate of individual amino acids using stable isotope techniques. The novel findings of the change patterns in molecular species of proteinaceous compounds may lead to new hypotheses: (1) selection through binding mechanisms such as electrostatic attractions will be shown to be similarly evident in other soils with similar mineralogical properties, (2) the partial fluxes of proteinaceous compounds among soil matrix components (e.g. mineral particles, soil solutions) might progress towards selectivity of accumulation on these compounds, and (3) polar surface or active site of proteins and

peptides may be preferentially adsorbed on silicate mineral surfaces, possibly causing decline of enzymatic activity.

Appendix A-Chapter 2

Table A2.1 Pairwise Multi-Response Permutation Procedures (MRPP) between a pair of site ages to compare amino acid composition in whole soil OM pool

Site Age	beach	105y	155y	210y	450y	845y	1475y	2385y	3210y
beach									
105y	0.005								
155y	0.003	0.020							
210y	0.002	0.002	0.004						
450y	0.002	0.002	0.002	0.002					
845y	0.002	0.002	0.002	0.003	0.033				
1475y	0.002	0.002	0.002	0.002	0.025	0.388			
2385y	0.002	0.002	0.002	0.002	0.113	0.009	0.043		
3210y	0.002	0.001	0.002	0.002	0.007	0.003	0.002	0.003	
4010y	0.002	0.002	0.002	0.002	0.006	0.012	0.002	0.014	0.029

Table A2.2 Pairwise Multi-Response Permutation Procedures (MRPP) between a pair of site ages to compare amino acid composition in mineral associated OM pool

Site Age	beach	105y	155y	210y	450y	845y	1475y	2385y	3210y
beach									
105y	0.477								
155y	0.027	0.439							
210y	0.010	0.419	0.575						
450y	0.010	0.031	0.017	0.292					
845y	0.381	0.635	0.454	NaN	0.793				
1475y	0.005	0.015	0.015	0.040	0.509	NaN			
2385y	0.003	0.004	0.003	0.013	0.114	0.546	0.063		
3210y	0.001	0.003	0.003	0.062	0.533	0.829	0.534	0.165	
4010y	0.004	0.004	0.005	0.027	0.033	NaN	0.073	0.010	0.034

Table A2.3. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.2.3.a and amino acid vectors (whole soil OM pool).

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
His	0.958	**0.917**	-0.721	0.282	0.079	0.099
Gly	-0.905	**0.819**	0.680	-0.160	0.026	-0.032
Asx	-0.881	**0.777**	0.734	-0.209	0.044	-0.001
Arg	0.810	**0.656**	-0.538	0.019	0.000	-0.084
Lys	0.784	**0.614**	-0.652	0.038	0.001	-0.002
Ala	-0.768	**0.590**	0.433	-0.429	0.184	-0.184
Pro	0.708	**0.501**	-0.318	0.245	0.060	-0.014
Cys	0.670	0.449	-0.463	0.014	0.000	0.032
Ile	0.652	0.425	-0.479	0.541	0.293	0.311
Phe	0.634	0.402	-0.576	0.123	0.015	0.025
Leu	0.624	0.390	-0.373	0.404	0.163	0.185
Tyr	-0.568	0.323	0.158	-0.491	0.241	-0.336
Glx	0.469	0.220	-0.220	0.829	**0.687**	0.740
Ser	-0.228	0.052	0.190	-0.910	**0.829**	-0.770
Met	0.213	0.046	-0.032	0.503	0.253	0.324
Thr	-0.071	0.005	0.004	-0.228	0.052	-0.149
Val	-0.012	0.000	0.047	0.462	0.214	0.177

Table A2.4 P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.2.3.b and amino acid vectors (mineral associated OM pool)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Cys	0.882	**0.779**	-0.684	0.094	0.009	0.136
Ala	-0.863	**0.745**	0.731	0.346	0.120	0.233
Leu	-0.811	**0.658**	0.651	0.412	0.170	0.313
Asx	-0.803	**0.646**	0.679	-0.113	0.013	-0.059
Gly	-0.754	**0.568**	0.521	0.477	0.228	0.321
Ile	-0.723	**0.522**	0.636	0.003	0.000	0.010
Phe	-0.609	0.371	0.495	0.135	0.018	0.105
Met	0.607	0.368	-0.467	0.040	0.002	-0.005
Thr	-0.605	0.366	0.415	-0.313	0.098	-0.251
Tyr	0.574	0.329	-0.531	-0.608	0.369	-0.377
Val	0.519	0.270	-0.359	0.206	0.043	0.154
His	0.432	0.187	-0.349	-0.777	**0.604**	-0.600
Ser	-0.291	0.085	0.067	0.373	0.139	0.256
Arg	-0.265	0.070	0.015	-0.563	0.317	-0.364
Lys	-0.222	0.049	0.033	-0.435	0.189	-0.105
Glx	-0.170	0.029	0.405	0.278	0.077	0.010
Pro	-0.058	0.003	0.021	0.073	0.005	0.138

Table A2.5. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.2.4.and amino acid vectors (whole soil and mineral associated OM pool)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Tyr	-0.895	**0.802**	-0.587	0.191	0.037	0.365
Asx	0.889	**0.790**	0.779	0.031	0.001	-0.063
His	-0.872	**0.760**	-0.763	0.224	0.050	0.187
Met	-0.865	**0.748**	-0.584	0.113	0.013	0.125
Lys	-0.845	**0.715**	-0.702	0.208	0.043	0.208
Thr	0.809	**0.655**	0.568	-0.163	0.027	-0.208
Leu	0.780	**0.609**	0.498	-0.317	0.100	-0.361
Ile	0.760	**0.577**	0.460	-0.339	0.115	-0.393
Cys	-0.709	**0.502**	-0.401	-0.364	0.132	-0.275
Arg	-0.675	**0.456**	-0.608	0.221	0.049	0.293
Ser	0.647	**0.418**	0.524	0.505	0.255	0.008
Phe	0.487	0.238	0.279	-0.381	0.145	-0.232
Ala	0.481	0.232	0.327	0.519	0.270	0.316
Gly	0.311	0.097	0.282	0.419	0.175	0.205
Val	0.194	0.037	0.239	-0.675	**0.455**	-0.459
Pro	-0.089	0.008	-0.020	-0.380	0.145	-0.188
Glx	0.036	0.001	0.111	-0.679	**0.461**	-0.496

Table A2.6. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Appendix_Fig.A2.1.and amino acid vectors (soluble OM hydrolysate).

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Val	-0.874	**0.764**	-0.711	-0.084	0.007	-0.110
Leu	-0.816	**0.666**	-0.656	-0.029	0.001	-0.012
Glx	0.762	**0.580**	0.621	-0.266	0.071	-0.179
Ile	-0.698	**0.487**	-0.602	-0.098	0.010	-0.141
Asx	-0.664	**0.441**	-0.445	-0.085	0.007	-0.020
Met	0.490	0.240	0.363	0.283	0.080	0.078
Cys	0.473	0.224	0.309	0.466	0.217	0.145
Tyr	0.451	0.204	0.324	0.332	0.110	0.080
His	0.427	0.183	0.259	0.116	0.014	0.158
Arg	-0.418	0.175	-0.331	0.198	0.039	0.189
Lys	0.325	0.106	0.069	0.663	0.439	0.504
Phe	0.256	0.065	0.241	-0.007	0.000	0.159
Ala	-0.177	0.031	-0.200	-0.386	0.149	-0.256
Pro	0.128	0.016	-0.187	-0.025	0.001	-0.128
Gly	0.103	0.011	0.027	0.910	0.828	0.752
Thr	0.075	0.006	0.012	-0.846	0.716	-0.546
Ser	-0.011	0.000	0.050	0.302	0.091	0.151

Table A2.7. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.2.3.a and selected soil properties (whole soil OM pool)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
pH	0.892	**0.795**	0.699	-0.168	0.028	-0.087
Mg(ug/g)	0.878	**0.771**	0.528	-0.066	0.004	0.068
CEC(cmolc/kg)	0.853	**0.728**	0.589	-0.304	0.093	0.004
Ca(ug/g)	0.836	**0.699**	0.595	-0.316	0.100	-0.035
K(ug/g)	-0.778	**0.605**	-0.556	0.349	0.122	0.172
Age	-0.573	**0.329**	-0.601	-0.207	0.043	-0.059
Mineralizable C (ug/g)	-0.511	0.261	-0.350	0.389	0.151	0.352
% OC	0.254	0.065	0.187	0.043	0.002	0.063
% N	-0.220	0.048	-0.090	0.227	0.051	0.281

Table A2.8 P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.2.3.b and selected soil properties (mineral associated OM pool)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Age	-0.738	**0.545**	-0.641	0.106	0.011	0.110
pH	0.664	**0.441**	0.523	-0.224	0.050	-0.058
Mg(ug/g)	0.645	**0.416**	0.559	-0.246	0.061	-0.113
CEC(cmolc/kg)	0.545	0.297	0.567	-0.046	0.002	-0.115
Ca(ug/g)	0.529	0.280	0.546	-0.032	0.001	-0.131
K(ug/g)	-0.480	0.230	-0.285	-0.066	0.004	-0.059
% OC	0.268	0.072	0.290	-0.176	0.031	-0.054
Mineralizable C (ug/g)	-0.174	0.030	-0.115	-0.083	0.007	0.008
% N	-0.071	0.005	-0.112	-0.179	0.032	-0.065

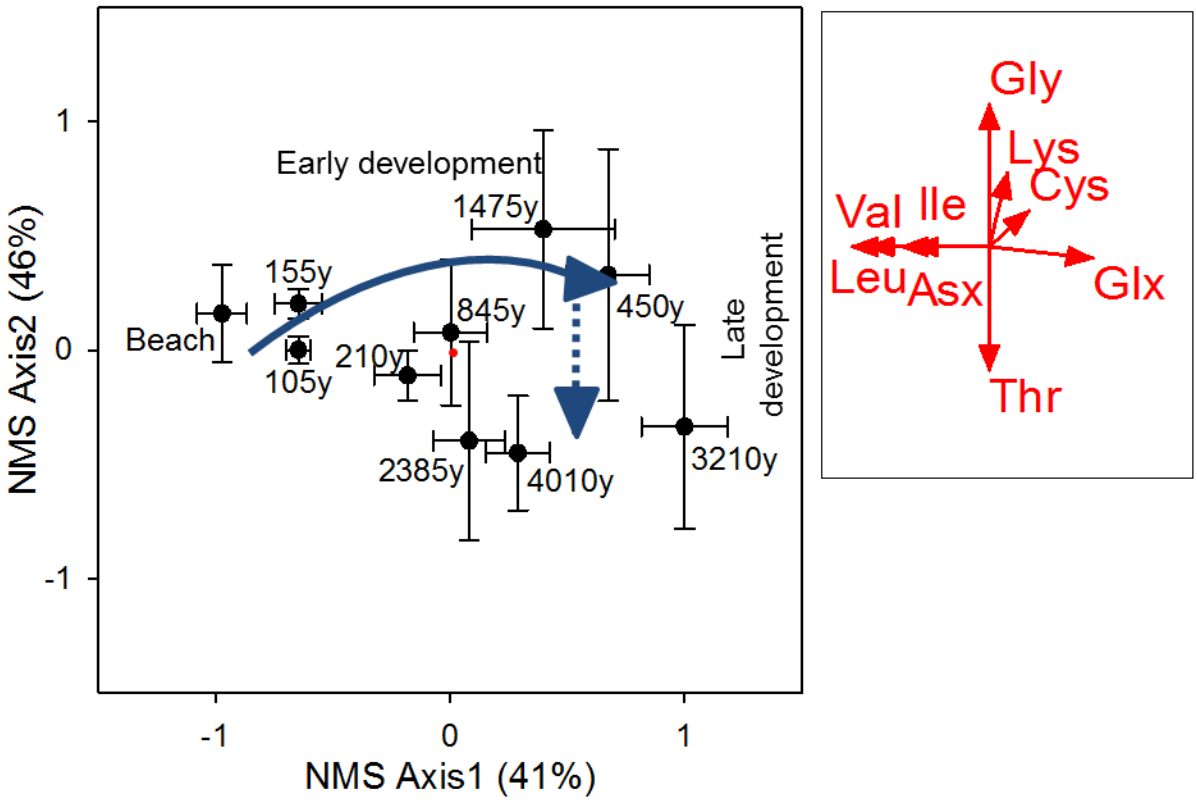
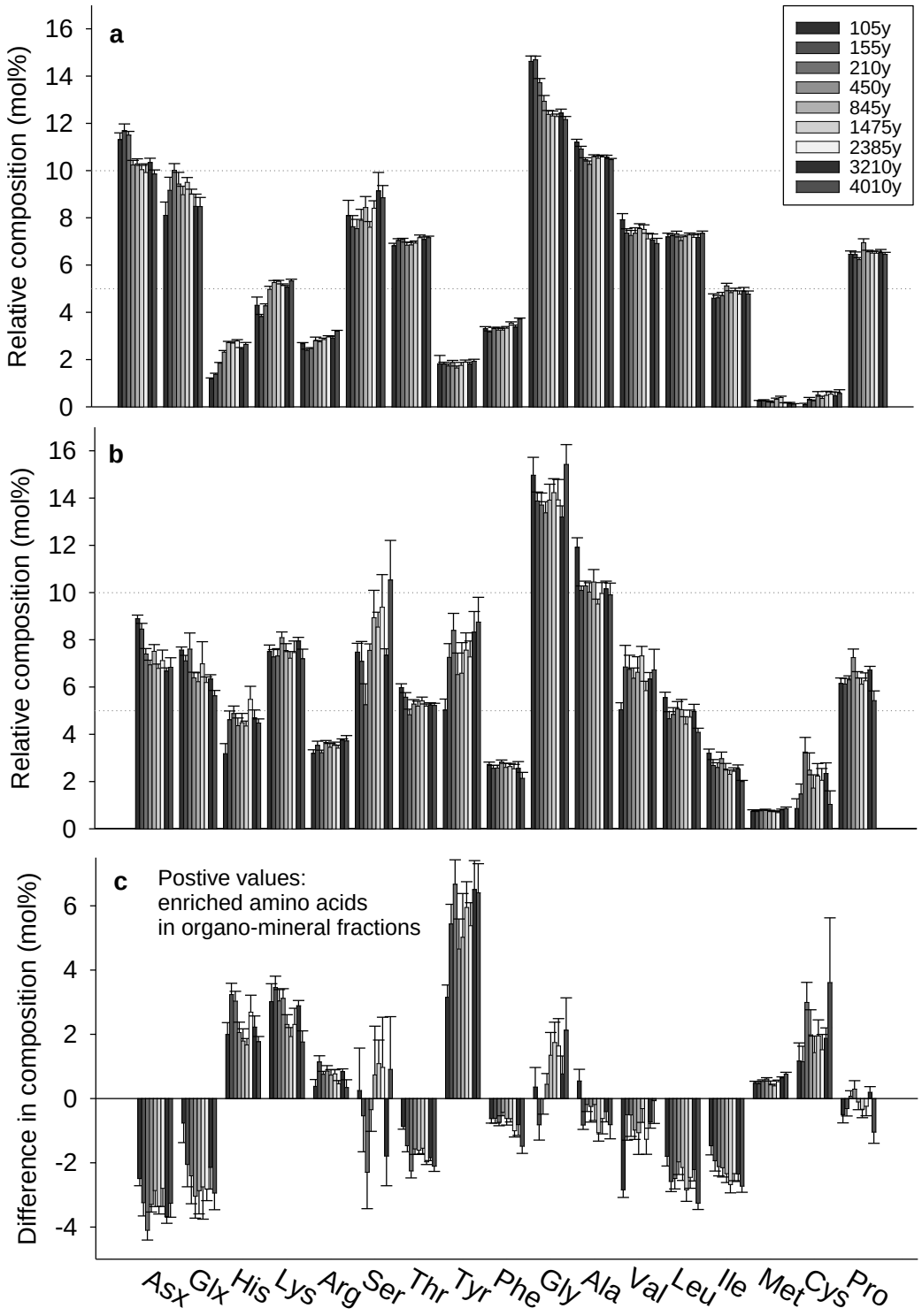


Figure A2.1. Relationship between the distribution of 17 proteinogenic amino acids and soil ecosystem development plotted by Nonmetric multidimensional scaling (NMS) ordination in soluble hydrolysates in the Lake Michigan sand dune chronosequence. Freshly deposited “beach” sand was also sampled to assess the amino acid distribution of parent material expected to be similar to the source material that formed the eolian deposits of the dune soils. Error bars represent standard error (n=5). Percentages on each axis in each plot denote the amount of variability associated with each axis. Red vectors show the direction and strength of the relationship between individual amino acids and ordination scores with the cutoff of r^2=0.3 The Pearson and Kendall correlations of the vectors are provided in the supplementary document (Appendix_Table A2.6).


a
b
c
105y
155y
210y
450y
845y
1475y
2385y
3210y
4010y
Relative composition (mol%)
Difference in composition (mol%)
Postive values:
enriched amino acids
in organo-mineral fractions
Asx
Glx
His
Lys
Arg
Ser
Thr
Tyr
Phe
Gly
Ala
Val
Leu
Ile
Met
Cys
Pro

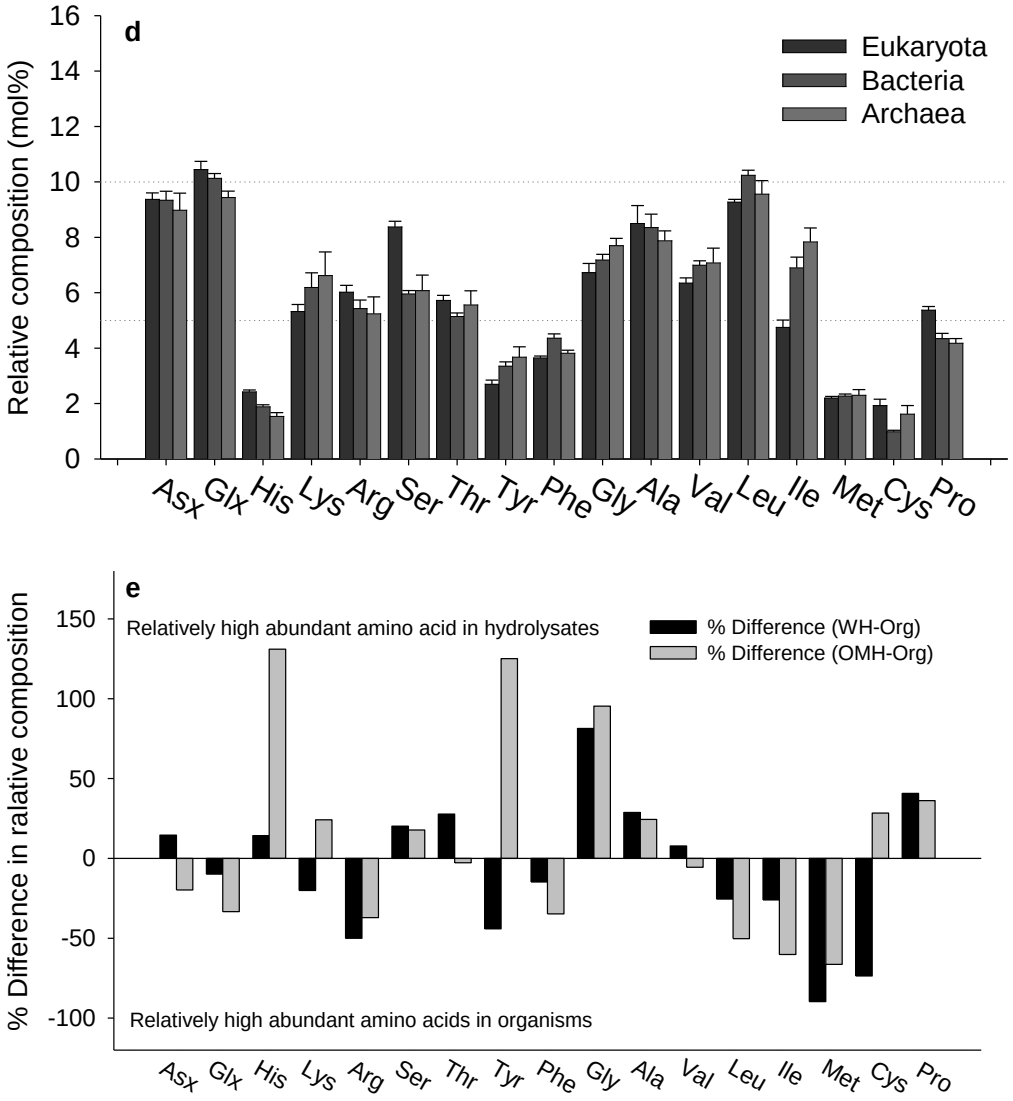


Figure A2.2. The relative composition (mol%) of amino acids in the hydrolysable extract from whole soil **(a)**, from mineral associated fraction **(b)**, and differences in mol% of amino acid in the hydrolysable extracts between mineral associated fraction and whole soil **(c)** across Lake Michigan sand dune chronosequence. In the figure c, positive values of difference in relative composition may indicate enrichment in mineral associated fraction rather than in whole soil while negative values to enrichment in whole soil rather than mineral associated fraction. Regenerated bar graph from Chen et al., 2012: relative protein amino acid composition of 3 domains from genome database **(d).** Percentage of difference in relative protein amino acid composition between organisms and two hydrolysates **(e)**. Aspartic acid (Asp), Asparagine (Asn), Asp+Asn =Asx, glutamic acid (Glu), Glutamine (Gln), Glu+Gln=Glx, histidine (His), lysine (Lys), arginine (Arg), serine (Ser), threonine (Thr), tyrosine (Tyr), phenylalanine (Phe), glycine (Gly), alanine (Ala), valine (Val), leucine (Leu), isoleucine (Ile), methionine (Met), cysteine (Cys), and proline (Pro).

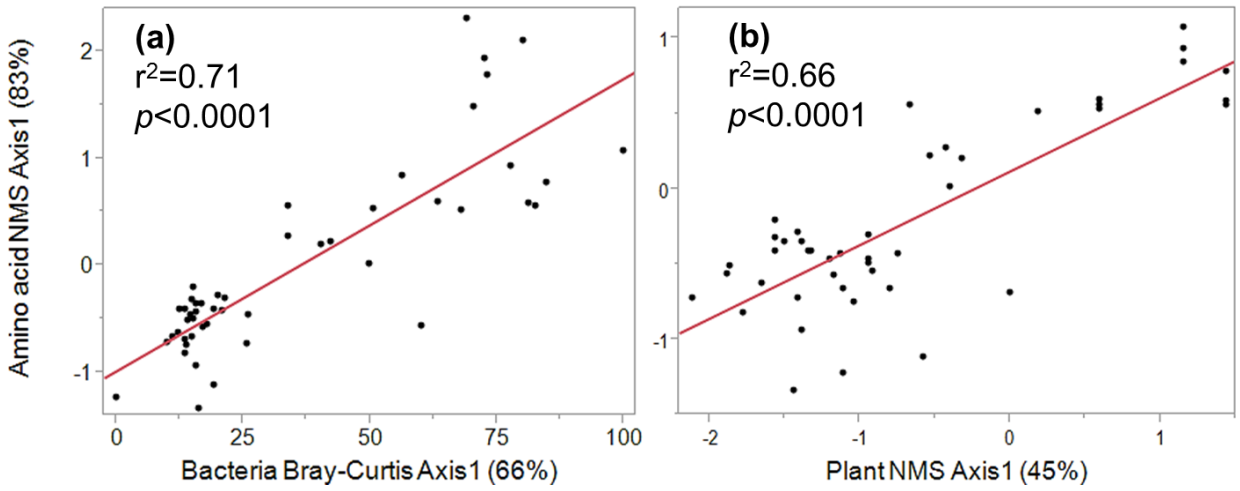


Figure A2.3..Regression between bacterial community composition and amino acid distribution **(a),** and between plant community composition and amino acid distribution **(b).**

Appendix B-Chapter 3

Table B3.1. Relative distribution (mol%) of 17 proteinogenic amino acids between summer and winter in whole soil, mineral associated, and soluble pools from Michigan chronosequences. Each column is listed in order of relative abundance and amino acids that are greater than the average (5.88%) are bolded

Whole soil OM					Mineral associated OM				Soluble OM			
Summer		Winter			Summer		Winter		Summer		Winter	
AA	mol%	SE	AA	mol%	SE	AA	mol%	SE	AA	mol%	SE	AA	mol%	SE
Gly	**13.01**	**0.19**	**Gly**	**13.12**	**0.14**	**Gly**	**13.15**	**0.29**	**Gly**	**14.65**	**0.26**	**Gly**	**21.24**	**0.37**	**Gly**	**23.21**	**0.45**
Ala	**10.65**	**0.07**	**Asx**	**10.64**	**0.12**	**Ala**	**10.47**	**0.17**	**Ala**	**10.01**	**0.20**	**Ala**	**11.97**	**0.23**	**Ala**	**11.04**	**0.26**
Asx	**10.51**	**0.17**	**Ala**	**10.59**	**0.05**	**Lys**	**7.91**	**0.15**	**Ser**	**8.79**	**0.43**	**Glx**	**11.01**	**0.36**	**Glx**	**10.29**	**0.35**
Ser	**9.34**	**0.19**	**Glx**	**9.90**	**0.14**	**Asx**	**7.55**	**0.16**	**Tyr**	**8.19**	**0.41**	**Thr**	**9.84**	**0.53**	**Ser**	**9.46**	**0.26**
Glx	**8.15**	**0.18**	**Val**	**7.89**	**0.05**	**Glx**	**7.28**	**0.25**	**Asx**	**7.25**	**0.16**	**Ser**	**9.27**	**0.16**	**Thr**	**8.08**	**0.30**
Leu	**7.10**	**0.06**	**Leu**	**7.34**	**0.04**	**Ser**	**7.02**	**0.48**	**Lys**	**7.11**	**0.13**	**Asx**	**7.93**	**0.16**	**Asx**	**7.30**	**0.22**
Thr	**7.05**	**0.05**	**Ser**	**7.02**	**0.13**	**Pro**	**6.47**	**0.13**	**Val**	**7.09**	**0.24**	Val	5.77	0.14	**Val**	**6.16**	**0.18**
Val	**6.79**	**0.08**	**Thr**	**6.95**	**0.04**	**Tyr**	**6.18**	**0.29**	**Pro**	**6.14**	**0.15**	Pro	4.95	0.20	**Pro**	**6.06**	**0.14**
Pro	**6.69**	**0.05**	**Pro**	**6.34**	**0.05**	Val	5.69	0.25	**Glx**	**6.10**	**0.12**	Leu	4.76	0.22	Leu	4.67	0.18
Lys	4.72	0.10	Ile	5.10	0.03	Thr	5.47	0.10	Thr	5.16	0.07	Lys	3.93	0.29	Ile	3.67	0.13
Ile	4.52	0.05	Lys	4.95	0.10	Leu	5.31	0.12	His	4.47	0.22	Ile	3.35	0.14	Lys	3.22	0.10
Phe	3.37	0.04	Phe	3.36	0.04	His	4.55	0.13	Leu	4.37	0.13	Arg	2.02	0.14	Arg	2.20	0.14
Arg	2.87	0.05	Arg	2.70	0.04	Arg	3.54	0.06	Arg	3.40	0.08	Phe	1.37	0.12	Phe	1.51	0.10
His	2.23	0.09	His	2.23	0.09	Ile	2.96	0.10	Phe	2.40	0.08	His	0.98	0.04	His	1.24	0.07
Tyr	2.09	0.04	Tyr	1.54	0.06	Cys	2.90	0.55	Ile	2.28	0.08	Cys	0.88	0.13	Tyr	0.90	0.14
Cys	0.65	0.06	Met	0.20	0.01	Phe	2.80	0.06	Cys	1.81	0.21	Tyr	0.49	0.05	Cys	0.67	0.14
Met	0.26	0.02	Cys	0.15	0.02	Met	0.74	0.02	Met	0.78	0.02	Met	0.24	0.02	Met	0.30	0.03

Table B3.2. Relative distribution (mol%) of 19 proteinogenic amino acids between summer and winter in soluble and microbial pools from Michigan chronosequences. Each column is listed in order of relative abundance and amino acids that are greater than the average (6.25%) are bolded

Soluble monomer						Microbial monomer					
Summer			Winter			Summer			Winter		
AA	mol%	SE	AA	mol%	SE	AA	mol%	SE	AA	mol%	SE
Gln+His	**15.19**	**1.03**	**Gln+His**	**14.99**	**1.25**	**Glu**	**19.03**	**2.08**	**Gln+His**	**15.60**	**1.93**
Glu	**11.65**	**0.59**	**Asn+Ser**	**11.42**	**0.34**	**Ala**	**13.61**	**0.74**	**Glu**	**14.81**	**1.96**
Ala	**11.20**	**0.44**	**Glu**	**11.20**	**0.64**	**Tyr**	**7.80**	**0.75**	**Ala**	**11.46**	**0.50**
Asn+Ser	**11.04**	**0.29**	**Lys+Leu**	**10.73**	**0.52**	**Val**	**7.09**	**0.35**	**Pro**	**8.87**	**0.84**
Lys+Leu	**10.86**	**0.53**	**Ala**	**10.70**	**0.30**	**Pro**	**6.84**	**0.92**	**Lys+Leu**	**8.77**	**0.73**
Val	5.58	0.38	Val	6.09	0.44	**Lys+Leu**	**6.34**	**0.67**	**Val**	**7.27**	**0.48**
Thr	5.41	0.17	Gly	5.85	0.37	**Asn+Ser**	**6.30**	**0.41**	Asn+Ser	5.24	0.32
Gly	5.10	0.26	Thr	5.59	0.17	Asp	5.42	0.42	Thr	4.42	0.23
Pro	4.48	0.53	Pro	4.35	0.46	Gln+His	5.02	0.79	Asp	4.35	0.24
Asp	4.16	0.29	Asp	3.86	0.16	Thr	4.83	0.23	Tyr	3.97	0.36
Arg	3.30	0.15	Ile	3.44	0.27	Met	4.35	0.33	Gly	3.66	0.36
Phe	3.28	0.23	Phe	3.18	0.22	Gly	3.95	0.27	Met	3.60	0.28
Ile	3.21	0.24	Arg	2.88	0.22	Ile	2.93	0.16	Ile	3.30	0.21
Met	2.04	0.09	Met	2.00	0.07	Cys	2.81	0.37	Arg	2.55	0.31
Cys	2.01	0.10	Cys	1.97	0.10	Arg	2.51	0.66	Phe	1.36	0.12
Tyr	1.49	0.14	Tyr	1.75	0.19	Phe	1.17	0.14	Cys	0.78	0.13

Table B3.3. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.1.a. and amino acid vectors (whole soil OM pool)

Axis:	1 (age)			2 (season)		
	r	r-sq	tau	r	r-sq	tau
Gly	-0.917	**0.841**	-0.726	-0.131	0.017	-0.167
His	0.903	**0.815**	0.620	0.080	0.006	0.072
Asx	-0.798	**0.637**	-0.619	-0.070	0.005	-0.061
Arg	0.718	**0.515**	0.563	0.346	0.119	0.242
Lys	0.709	**0.502**	0.547	-0.076	0.006	-0.034
Phe	0.545	0.297	0.453	-0.004	0.000	0.006
Ala	-0.471	0.222	-0.277	0.051	0.003	-0.048
Cys	0.366	0.134	0.256	0.657	**0.432**	0.504
Ile	0.265	0.070	0.104	-0.752	**0.565**	-0.543
Val	-0.263	0.069	-0.213	-0.839	**0.705**	-0.643
Pro	0.258	0.067	0.169	0.492	0.242	0.391
Ser	0.221	0.049	0.203	0.969	**0.939**	0.874
Leu	0.178	0.032	0.100	-0.494	0.244	-0.334
Thr	0.147	0.022	0.179	0.338	0.114	0.239
Tyr	0.101	0.010	0.136	0.702	**0.493**	0.525
Glx	-0.045	0.002	-0.167	-0.885	**0.782**	-0.732
Met	-0.044	0.002	-0.178	0.000	0.000	0.003

Table B3.4. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.1.b. and amino acid vectors (mineral associated OM pool)

Axis:	1 (age)			2 (season)		
	r	r-sq	tau	r	r-sq	tau
Ala	-0.888	**0.788**	-0.751	0.093	0.009	0.106
Leu	-0.871	**0.758**	-0.746	0.311	0.097	0.267
Tyr	0.820	**0.672**	0.701	-0.300	0.090	-0.267
Ile	-0.792	**0.627**	-0.666	0.434	0.188	0.362
Asx	-0.650	**0.423**	-0.468	0.337	0.113	0.286
Thr	-0.631	**0.398**	-0.446	0.318	0.101	0.275
His	0.603	**0.363**	0.383	0.190	0.036	0.170
Val	0.587	**0.345**	0.455	-0.429	0.184	-0.380
Cys	0.583	**0.340**	0.402	0.378	0.143	0.155
Met	0.450	0.203	0.311	-0.101	0.010	-0.090
Pro	-0.450	0.202	-0.388	0.170	0.029	0.148
Glx	-0.422	0.178	-0.478	0.562	**0.316**	0.463
Gly	-0.414	0.171	-0.183	-0.724	**0.524**	-0.539
Lys	-0.333	0.111	-0.313	0.459	0.211	0.351
Phe	-0.296	0.088	-0.436	0.361	0.130	0.343
Arg	0.137	0.019	0.093	0.018	0.000	0.007
Ser	-0.060	0.004	-0.016	-0.719	**0.516**	-0.517

Table B3.5. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.2 and amino acid vectors (water soluble OM pool)

Axis:	1 (age)			2 (season)		
	r	r-sq	tau	r	r-sq	tau
Val	-0.794	**0.631**	-0.671	0.186	0.035	0.132
Leu	-0.772	**0.596**	-0.618	0.157	0.025	0.076
Ile	-0.753	**0.567**	-0.646	0.151	0.023	0.104
Glx	0.584	**0.342**	0.450	0.066	0.004	0.064
Lys	0.443	0.197	0.179	-0.351	0.123	-0.229
Met	0.440	0.194	0.312	-0.258	0.067	-0.099
Arg	-0.364	0.132	-0.241	-0.296	0.088	-0.225
Cys	0.270	0.073	0.176	-0.058	0.003	0.025
Phe	0.251	0.063	0.250	-0.071	0.005	-0.198
Gly	0.224	0.050	0.103	-0.860	**0.739**	-0.747
His	0.218	0.048	0.178	-0.444	0.197	-0.293
Pro	-0.209	0.044	-0.232	-0.128	0.016	-0.038
Tyr	0.177	0.031	0.138	-0.118	0.014	-0.007
Asx	-0.128	0.017	-0.232	0.305	0.093	0.291
Ala	-0.116	0.013	-0.172	0.683	**0.466**	0.533
Thr	0.064	0.004	0.029	0.774	**0.599**	0.489
Ser	-0.027	0.001	0.154	-0.504	0.254	-0.361

Table B3.6, P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.3. and amino acid vectors (theoretical origins and 3 different OM hydrolysates)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Tyr	-0.861	**0.741**	-0.750	-0.108	0.012	0.003
Lys	-0.815	**0.664**	-0.667	0.120	0.014	0.089
His	-0.814	**0.662**	-0.674	-0.126	0.016	-0.018
Thr	0.725	**0.526**	0.633	-0.150	0.022	-0.047
Gly	0.687	**0.472**	0.486	-0.774	**0.599**	-0.544
Glx	0.631	**0.398**	0.433	0.100	0.010	0.197
Ala	0.539	0.290	0.409	-0.369	0.136	-0.135
Phe	-0.524	0.274	-0.308	0.720	0.519	0.604
Arg	-0.515	0.265	-0.552	0.571	**0.326**	0.232
Cys	-0.489	0.239	-0.330	-0.168	0.028	-0.127
Met	-0.484	0.234	-0.439	0.536	0.287	-0.017
Ser	0.358	0.128	0.305	-0.479	0.229	-0.363
Val	-0.217	0.047	-0.106	0.391	0.153	0.323
Leu	-0.133	0.018	-0.075	0.915	**0.837**	0.754
Pro	-0.118	0.014	-0.134	-0.109	0.012	0.052
Asx	0.054	0.003	0.103	0.549	**0.301**	0.481
Ile	0.041	0.002	0.061	0.839	**0.704**	0.659

Table B3.7. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.4. and amino acid vectors (mineral associated vs. water soluble OM sub-pools)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Tyr	-0.933	**0.871**	-0.724	0.196	0.038	0.124
His	-0.916	**0.839**	-0.639	0.220	0.048	0.143
Met	-0.848	**0.718**	-0.615	0.118	0.014	0.117
Lys	-0.843	**0.711**	-0.567	0.182	0.033	0.219
Gly	0.835	**0.697**	0.556	-0.559	0.312	-0.447
Glx	0.749	**0.561**	0.587	-0.233	0.054	-0.125
Thr	0.699	**0.489**	0.631	0.260	0.067	0.054
Phe	-0.676	**0.457**	-0.443	0.246	0.061	0.283
Arg	-0.674	**0.455**	-0.485	0.134	0.018	0.174
Ile	0.531	0.281	0.372	0.241	0.058	0.154
Cys	-0.528	0.278	-0.370	0.031	0.001	0.058
Ala	0.523	0.274	0.420	0.332	0.110	0.202
Ser	0.365	0.133	0.367	-0.529	0.279	-0.468
Pro	-0.315	0.099	-0.250	0.204	0.041	0.292
Val	-0.278	0.077	-0.169	0.152	0.023	0.009
Asx	0.171	0.029	0.176	0.169	0.029	0.157
Leu	0.037	0.001	0.060	0.395	0.156	0.271

Table B3.8. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.5 and amino acid vectors (Soluble hydrolysate vs monomer)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Glx+His	0.937	**0.878**	0.718	-0.390	0.152	-0.327
Gly	-0.923	**0.851**	-0.693	-0.172	0.030	-0.041
Met	0.756	**0.572**	0.461	0.422	0.178	0.180
Thr	-0.667	**0.445**	-0.626	-0.008	0.000	0.068
Cys	0.603	**0.364**	0.463	0.165	0.027	0.108
Arg	0.480	0.230	0.323	-0.056	0.003	-0.029
Ile	-0.410	0.168	-0.335	0.817	**0.667**	0.548
Val	-0.406	0.165	-0.379	0.861	**0.741**	0.579
Tyr	0.319	0.102	0.213	0.395	0.156	0.286
Phe	0.303	0.092	0.311	0.783	**0.613**	0.350
Ala	-0.283	0.080	-0.242	0.221	0.049	0.280
Asx+Ser	-0.237	0.056	-0.175	-0.528	0.279	-0.320
Lys+Leu	0.095	0.009	0.107	0.911	**0.830**	0.665
Pro	0.015	0.000	-0.225	-0.526	0.277	-0.249

Table B3.9. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.6.a and amino acid vectors (Soluble monomer AA)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Val	0.394	**0.910**	0.818	-0.167	0.028	-0.106
Ile	0.547	**0.890**	0.762	-0.190	0.036	-0.138
Gln+His	0.261	**0.864**	-0.787	-0.278	0.077	-0.092
Lys+Leu	0.930	**0.856**	0.764	-0.276	0.076	-0.250
Phe	0.292	**0.841**	0.695	-0.268	0.072	-0.245
Thr	-0.925	**0.626**	0.599	-0.109	0.012	-0.067
Gly	-0.742	**0.551**	0.472	0.027	0.001	0.091
Pro	-0.791	**0.383**	-0.474	-0.127	0.016	0.142
Glu	-0.338	0.300	-0.308	0.651	0.424	0.609
Met	-0.476	0.227	0.412	-0.063	0.004	-0.072
Ala	0.188	0.184	0.460	0.578	**0.334**	0.172
Asp	-0.428	0.155	-0.350	0.443	0.196	0.290
Tyr	0.619	0.114	0.326	-0.161	0.026	-0.113
Arg	-0.954	0.085	-0.195	-0.281	0.079	-0.094
Asn+Ser	-0.943	0.068	-0.216	0.243	0.059	0.204
Cys	-0.917	0.035	-0.078	-0.021	0.000	0.040

Table B3.10. P-value of Pearson and Kendall correlations between the ordination scores of the NMS axes of Fig.3.6.c and amino acid vectors (microbial AA)

Axis:	1			2		
	r	r-sq	tau	r	r-sq	tau
Glu	-0.957	**0.915**	-0.750	0.082	0.007	0.002
Met	0.685	**0.469**	0.469	0.150	0.022	0.150
Lys+Leu	0.590	**0.348**	0.490	0.061	0.004	0.040
Pro	0.546	0.298	0.444	-0.534	0.285	-0.217
Gln+His	0.541	0.293	0.367	-0.772	**0.596**	-0.635
Gly	-0.487	0.237	-0.379	0.200	0.040	0.174
Phe	0.445	0.198	0.370	0.163	0.026	0.157
Tyr	0.415	0.172	0.288	0.121	0.015	0.141
Cys	0.389	0.152	0.297	0.173	0.030	0.263
Thr	0.311	0.097	0.153	0.619	**0.383**	0.513
Asp	-0.264	0.070	-0.248	0.461	0.213	0.405
Asn+Ser	0.216	0.046	0.119	0.244	0.060	0.228
Ala	-0.177	0.031	-0.210	0.630	**0.397**	0.471
Val	-0.164	0.027	-0.142	0.501	0.251	0.457
Arg	0.129	0.017	0.055	-0.352	0.124	-0.337
Ile	0.050	0.003	0.049	0.502	0.252	0.444

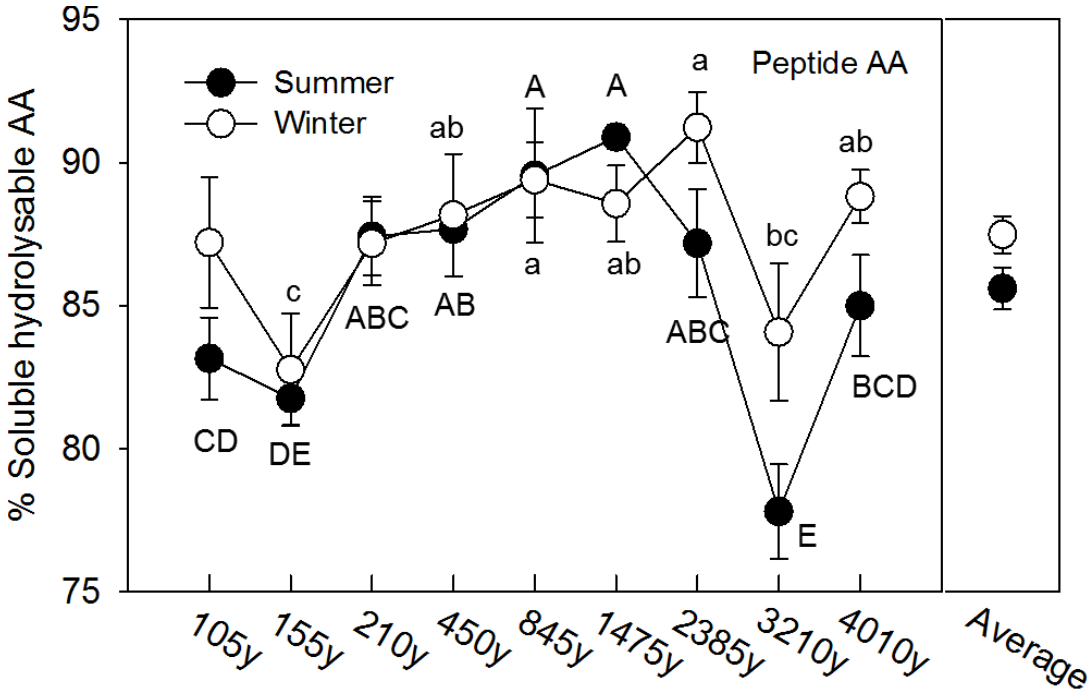


Figure B3.1. Comparisons of proportion of peptide form of amino acid to soluble hydrolysable amino acid between summer and winter across Lake Michigan chronosequence. The amount of peptide AA was calculated by subtracting monomer in soluble pool from soluble hydrolysable pool. The proportions were tested by Two way-ANOVA between summer and winter ($p=0.8224$); among the age ($p<0.0001$); interaction term ($p=0.2950$). Letters denote significant difference, and the abundances of two seasons were separately tested by Student's t ($P< 0.05$) along the years of development: upper case=summer, lower case=winter ($p<0.0001$ for summer and $p=0.0459$ for winter). Error bars represent standard error ($n=5$).

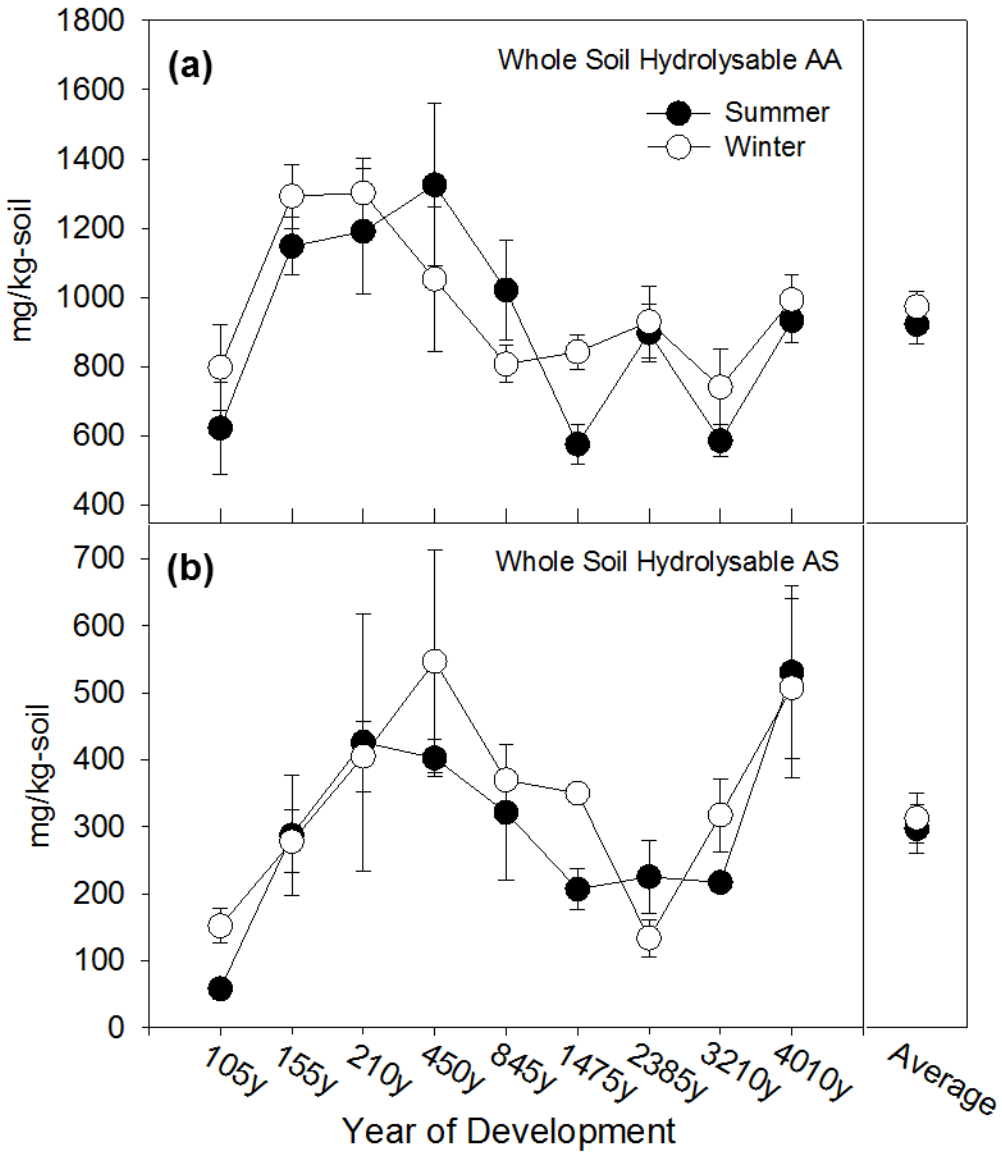


Figure B3.2. Comparisons of abundance in mg/kg-dry soil of amino acid (a) and amino sugar (b) from whole soil pool between summer and winter across Lake Michigan chronosequence. The abundances were tested by ANOVA, for amino acid: between seasons ($p=0.3073$); among age ($p<0.0001$); interaction term ($p=0.3617$); for amino sugar: between seasons ($p=0.4428$); among age ($p=0.0184$); interaction term ($p=0.8658$); Error bars represent standard error ($n=5$).

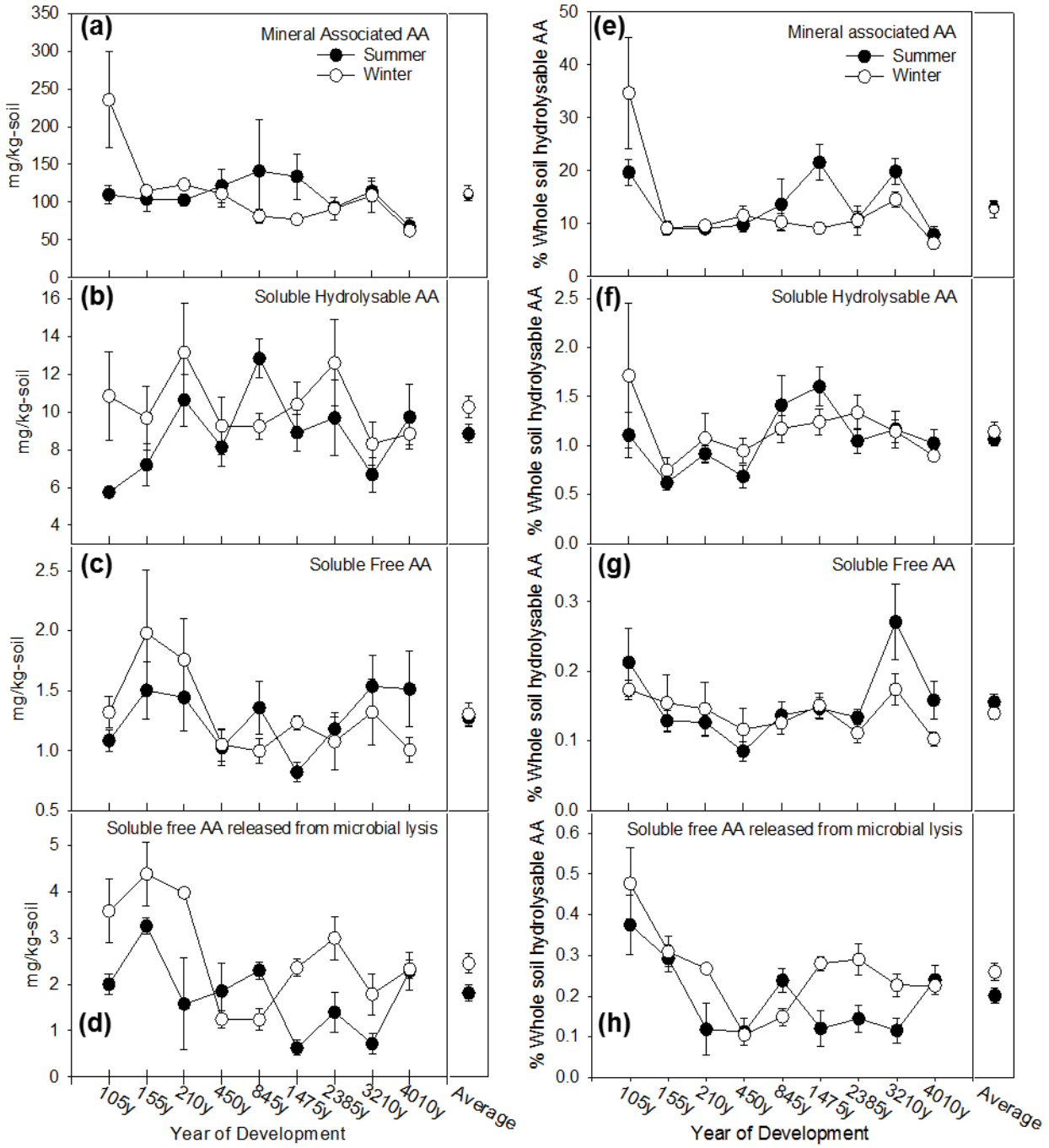


Figure B3.3. Comparison of amino acid abundance in different OM pools and their proportion to the whole soil pool between summer and winter across Lake Michigan chronosequence: (a) and (e) from mineral associated fraction; (b) and (f) from hydrolysates of soluble fraction; (c) and (g) from soluble fraction including amino acid monomers; and (d) and (h) from microbial fraction including amino acid monomers, respectively. Error bars represent standard error (n≤5)

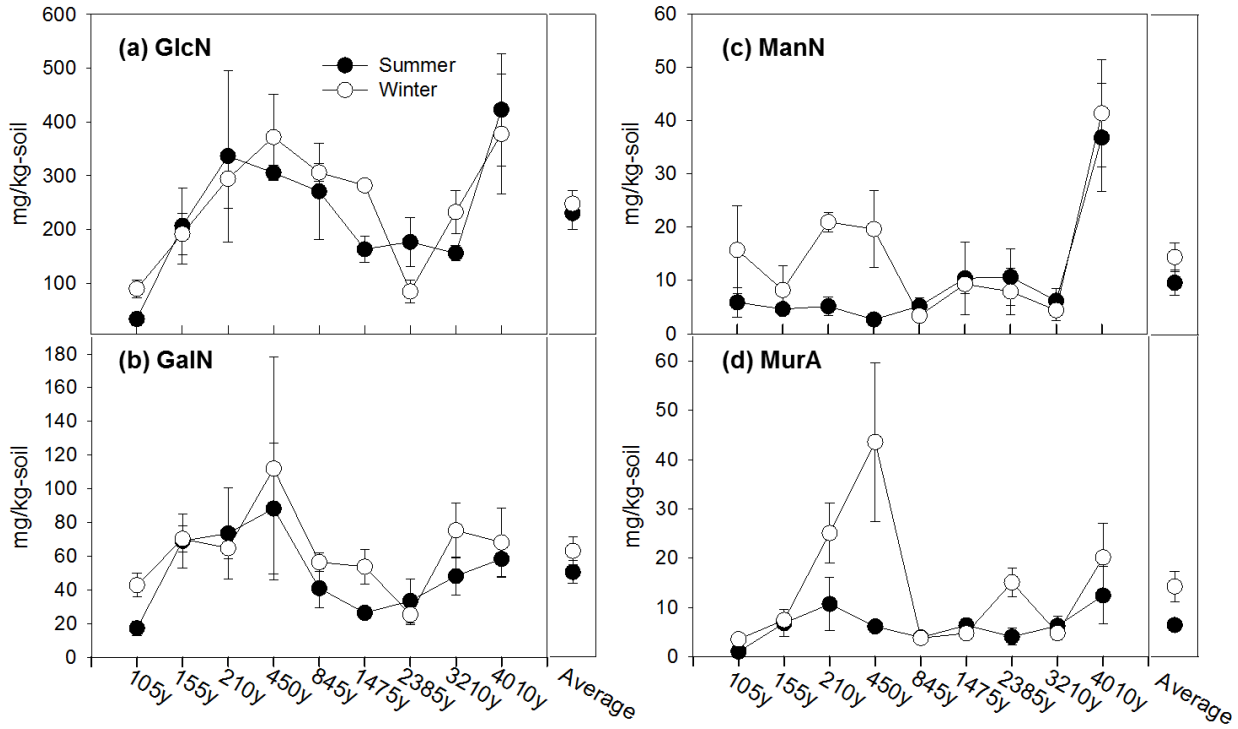


Figure B3.4. Comparison of the abundance of four individual amino sugars in whole soil pool between summer and winter across Lake Michigan chronosequence: Glucosamine, GlcN (a); Galactosamine, GalN (b); Mannosamine, ManN (c);and Muramic acid, MurA (d). Error bars represent standard error (n=5) The abundances were tested by Two way-ANOVA and the p-values show below

Abundance	Season	Age	Season*Age
GlcN	0.5516	0.0091	0.8084
GalN	0.409	0.3352	0.9818
ManN	0.1805	**0.0012**	0.3614
MurA	0.7201	0.8532	**0.0072**

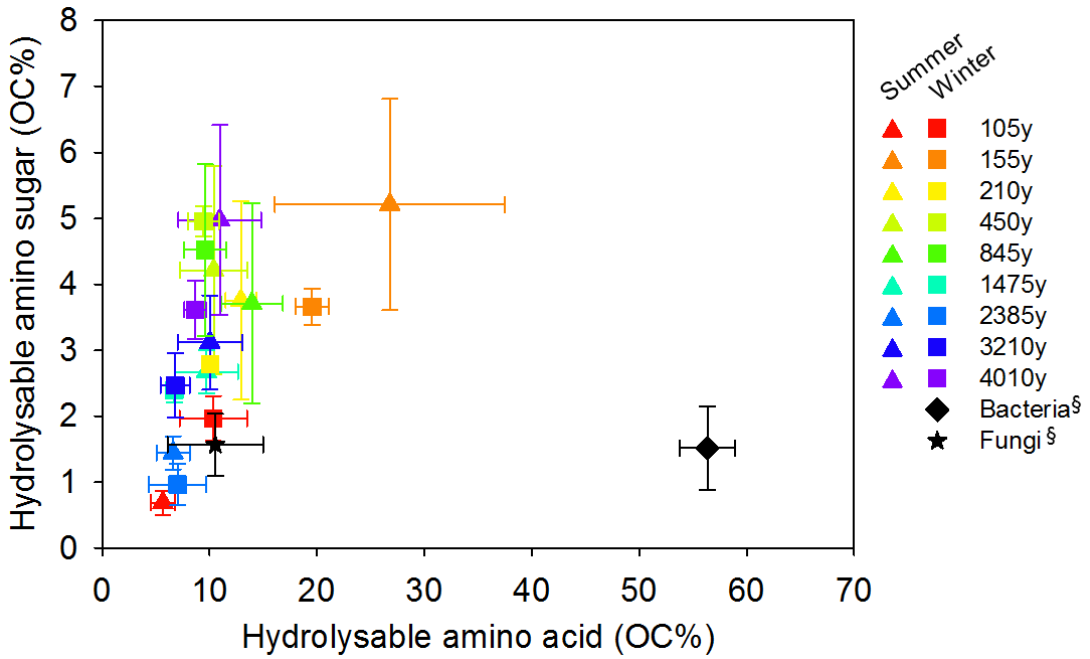


Figure B3.5. . Relationship between hydrolysable amino acid-C and hydrolysable amino sugar-C in whole soil pool and microorganisms). Error bars represent standard error (n=5 for soil).

§Bacterial and fungal amino acid and amino sugar from Hobara et al., 2014.

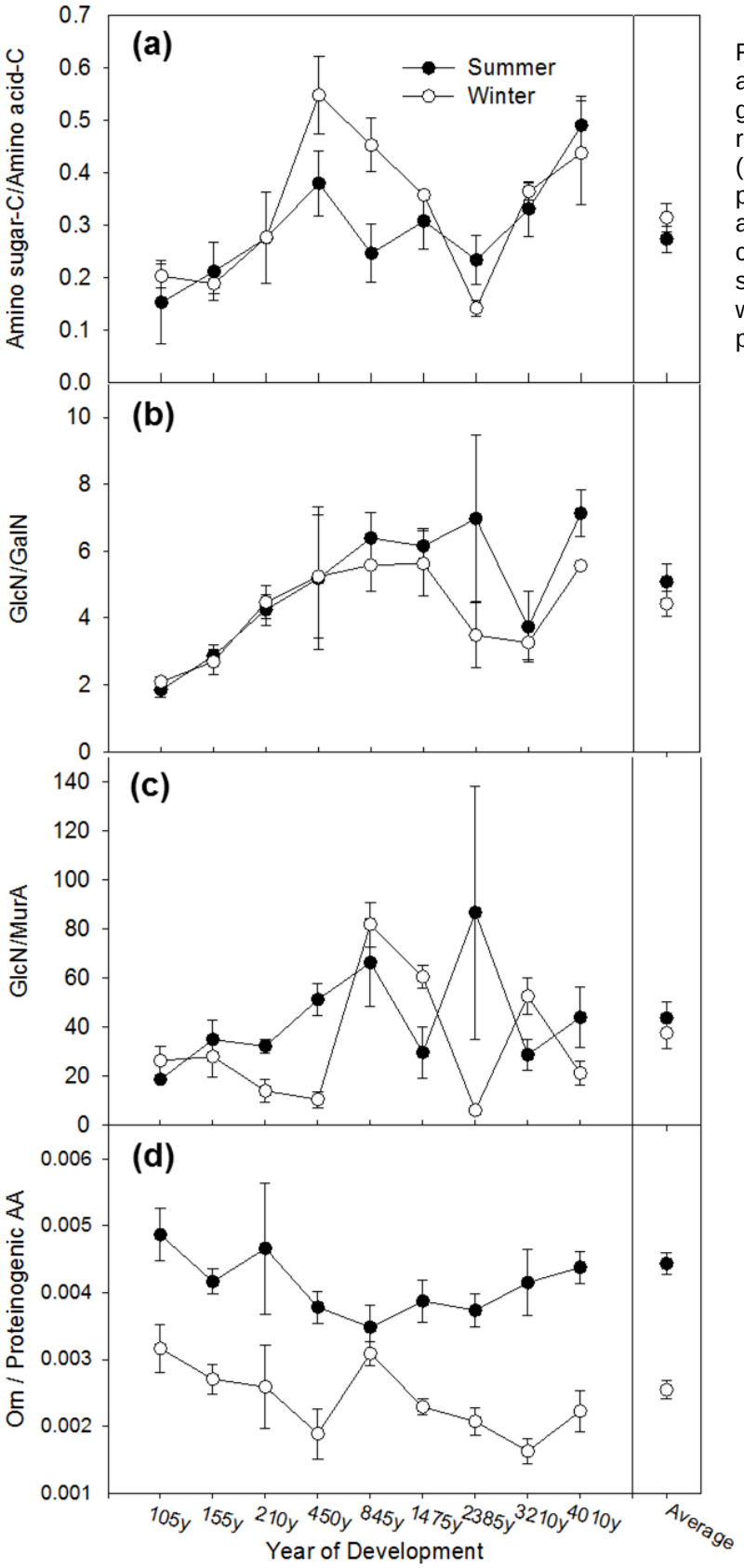


Figure B3.6. Comparison of ratio of amino sugar to amino acid (a), ratio of glucosamine to galactosamine (b), ratio of glucosamine to muramic acid (c), and ratio of ornithine to total protein between summer and winter across Lake Michigan chronosequence:.Error bars represent standard error (n=5). The abundances were tested by 2way-ANOVA and the p-values show below

Abundance	AS/AA	GlcN/GalN	GluN/MurA	Orn/PAA
Season	0.5201	0.8725	0.7967	**0.0027**
Age	**<0.0001**	**0.0083**	0.1022	0.2194
Season*Age	0.1665	0.7523	**0.017**	0.3180

Appendix C-Chapter 4

Table C4.1. Relative distribution (mol%) of 17 proteinogenic amino acids in theoretical protein sources and in soil from Michigan and Haast chronosequences. Each column is listed in order of relative abundance and amino acids that are greater than the average (5.88%) are bolded. Note that the top five most abundant amino acids (Ala, Asx, Gly, Glx, and Ser) are in common between theoretical protein sources and soil organic matter.

Theoretical protein sources (mol%)						Michigan (mol%)				Haast (mol%)			
Eukarya		Bacteria		Archaea		Whole soil OM		Mineral associated OM		Whole soil OM		Mineral associated OM	
Glx	**10.54**	**Leu**	**10.34**	**Leu**	**9.65**	**Gly**	**13.01**	**Gly**	**13.15**	**Gly**	**16.12**	**Gly**	**17.65**
Asx	**9.45**	**Glx**	**10.23**	**Glx**	**9.52**	**Ala**	**10.65**	**Ala**	**10.47**	**Asx**	**10.55**	**Ala**	**10.78**
Leu	**9.35**	**Asx**	**9.43**	**Asx**	**9.06**	**Asx**	**10.51**	**Lys**	**7.91**	**Ala**	**10.26**	**Pro**	**7.93**
Ala	**8.57**	**Ala**	**8.43**	**Ala**	**7.95**	**Ser**	**9.34**	**Asx**	**7.55**	**Ser**	**9.65**	**Ser**	**7.45**
Ser	**8.45**	**Gly**	**7.25**	**Ile**	**7.91**	**Glx**	**8.15**	**Glx**	**7.28**	**Glx**	**8.52**	**Glx**	**7.18**
Gly	**6.78**	**Val**	**7.06**	**Gly**	**7.77**	**Leu**	**7.10**	**Ser**	**7.02**	**Thr**	**7.54**	**Asx**	**6.96**
Val	**6.40**	**Ile**	**6.96**	**Val**	**7.15**	**Thr**	**7.05**	**Pro**	**6.47**	**Val**	**6.87**	**Leu**	**6.74**
Arg	**6.08**	**Lys**	**6.24**	**Lys**	**6.68**	**Val**	**6.79**	**Tyr**	**6.18**	**Pro**	**6.68**	Val	5.62
Thr	5.77	**Ser**	**6.02**	**Ser**	**6.13**	**Pro**	**6.69**	**Val**	**5.69**	Leu	5.74	Lys	5.48
Pro	5.42	Arg	5.48	Thr	5.61	Lys	4.72	**Thr**	**5.47**	Lys	4.46	Thr	5.08
Lys	5.37	Thr	5.19	Arg	5.29	Ile	4.52	**Leu**	**5.31**	Ile	3.87	Ile	3.90
Ile	4.79	Phe	4.40	Pro	4.22	Phe	3.37	His	4.55	Phe	2.97	Phe	3.86
Phe	3.68	Pro	4.39	Phe	3.85	Arg	2.87	Arg	3.54	His	2.37	Arg	3.06
Tyr	2.72	Tyr	3.38	Tyr	3.71	Has	2.23	Ile	2.96	Arg	2.37	His	2.92
His	2.45	Met	2.29	Met	2.32	Tyr	2.09	Cys	2.90	Tyr	1.53	Tyr	2.90
Met	2.22	His	1.91	Cys	1.63	Cys	0.65	Phe	2.80	Cys	0.44	Cys	1.96
Cys	1.94	Cys	1.00	His	1.55	Met	0.26	Met	0.74	Met	0.06	Met	0.54

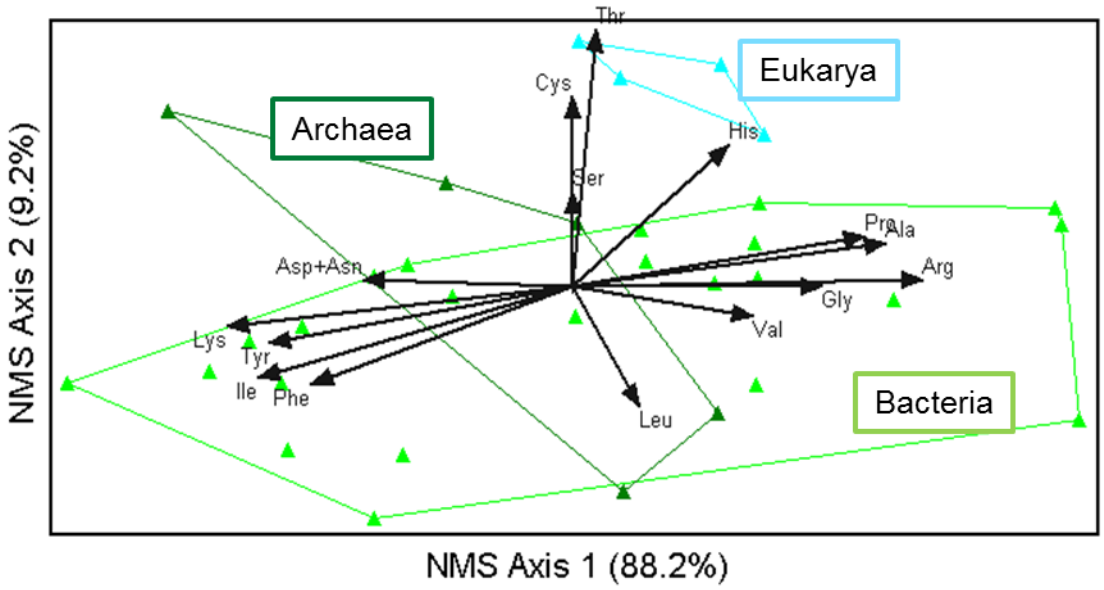


Figure C4.1. Comparisons of the amino acid composition of the theoretical protein sources: Eukarya (▲ cyan), Bacteria (▲ light green), and Archaea (▲ green). Nonmetric multidimensional scaling (NMS) ordination plot of proteinogenic amino acids was modified based on NCBI genome database provided by Chen et al., 2013. Correlations of variables with ordination with r2>0.3 were shown in bi-plot vector where length and direction represent the magnitude and directions of the correlation, respectively. Percentages on each axis denote the amount of variability associated with each axis.

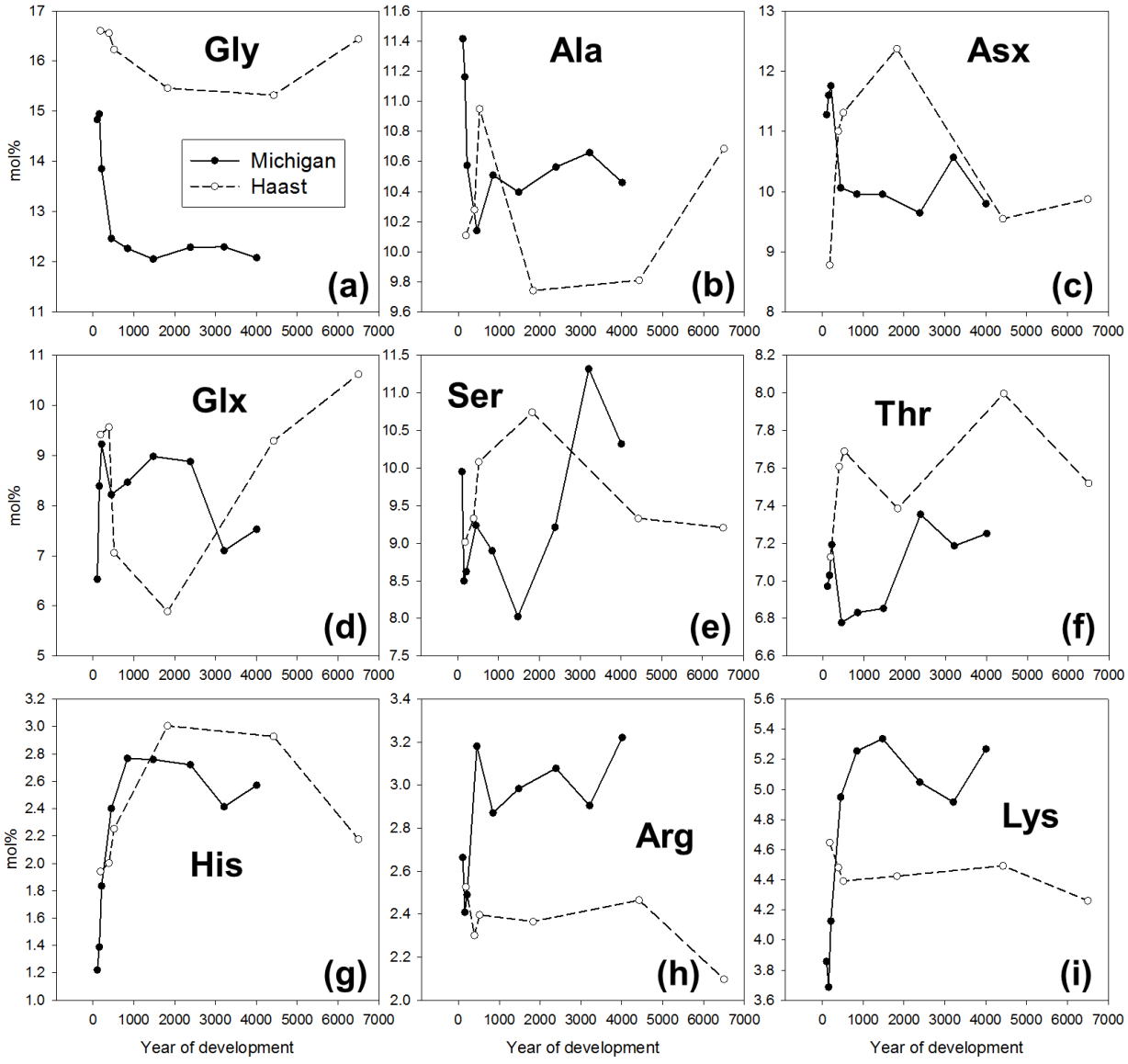


Figure C4.2. Whole soil OM pool. The relationship between year of development and mol% of the six most abundant amino acids **(a-f)** as well as mol% of positively charged amino acids **(g-i)** in Michigan and Haast chronosequences. Each point (close point=Michigan; open point=Haast) in the graphs are the average (n=5 for Michigan; n=4 for Haast) of the mol% of each amino acid at each stage of development.

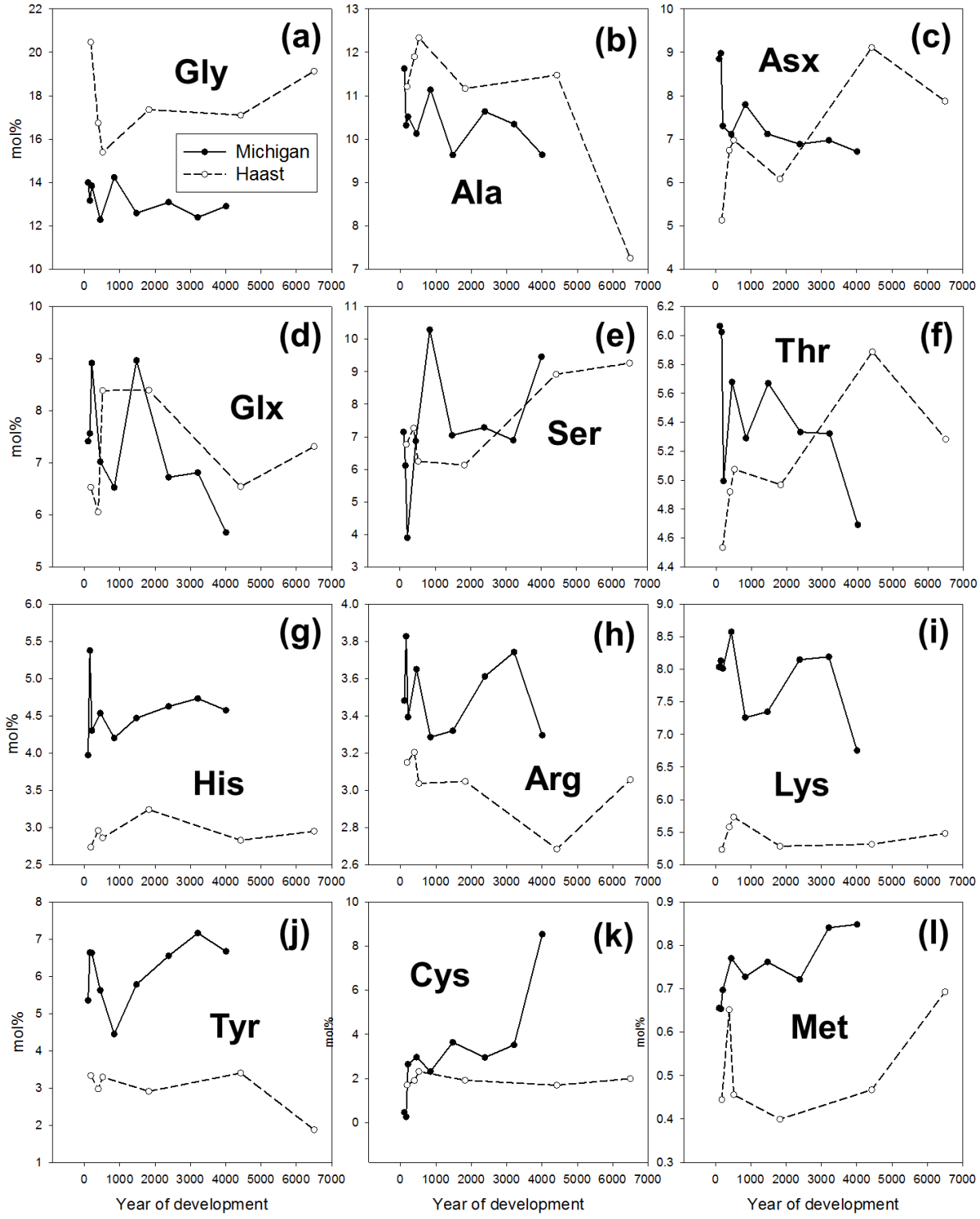
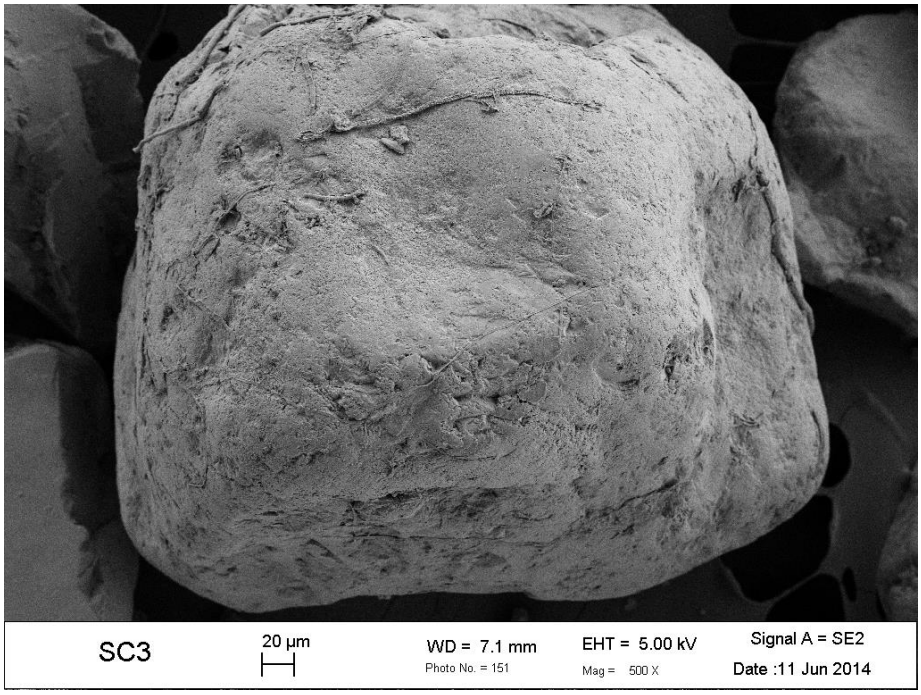
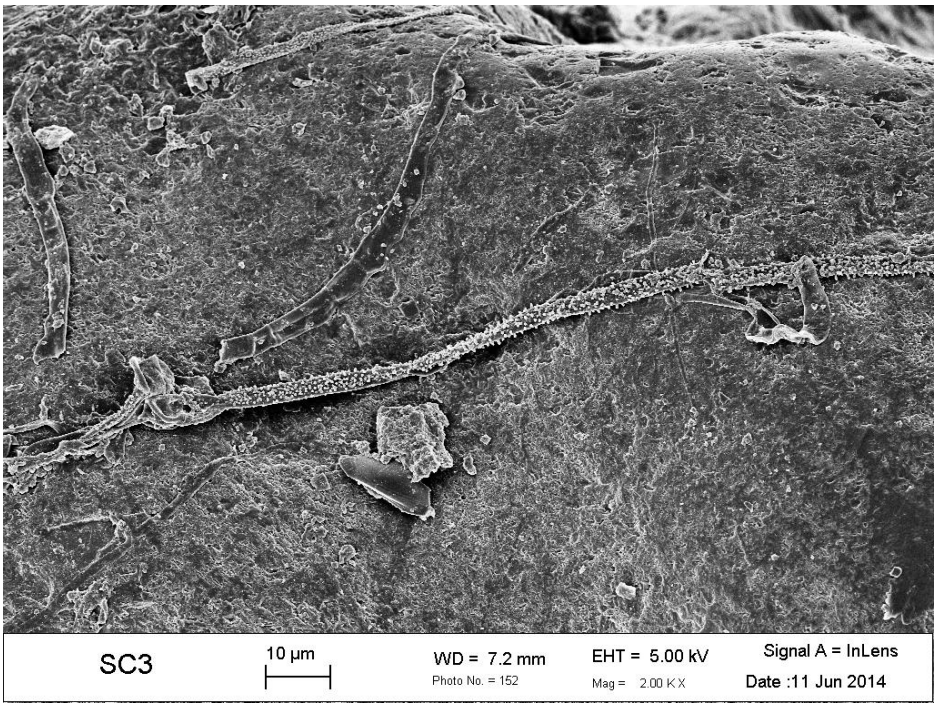


Figure C4.3 Mineral associated OM pool. The relationship between year of development and mol% of the twelve important amino acids regarding mineral interactions in Michigan and Haast chronosequences. Each point (close point=Michigan; open point=Haast) in the graphs are the average (n=5 for Michigan; n=4 for Haast) of the mol% of each amino acid at each stage of development.

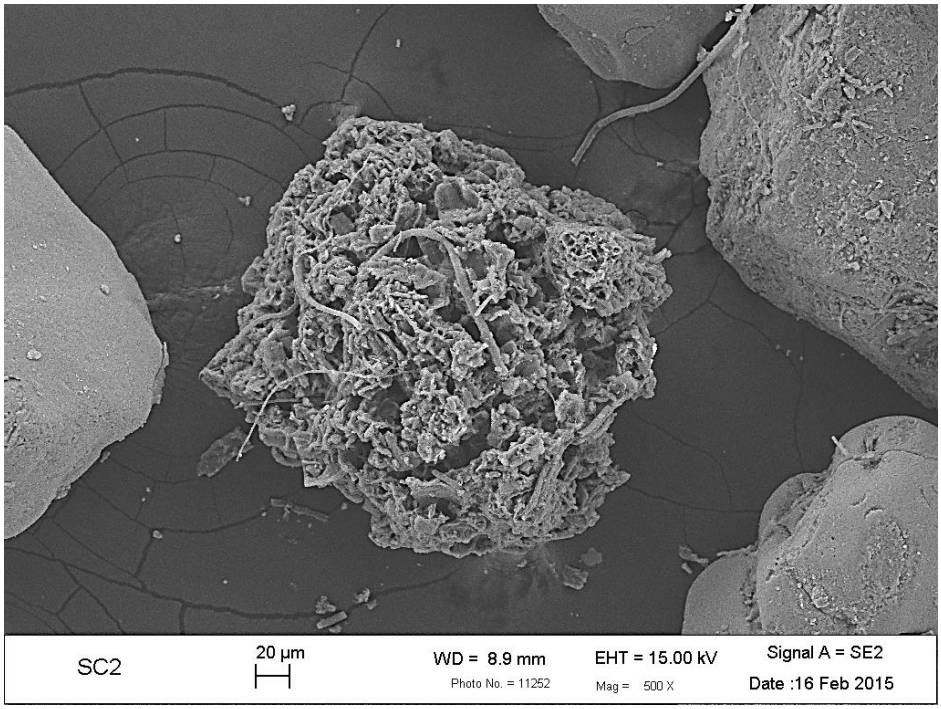
Appendix D



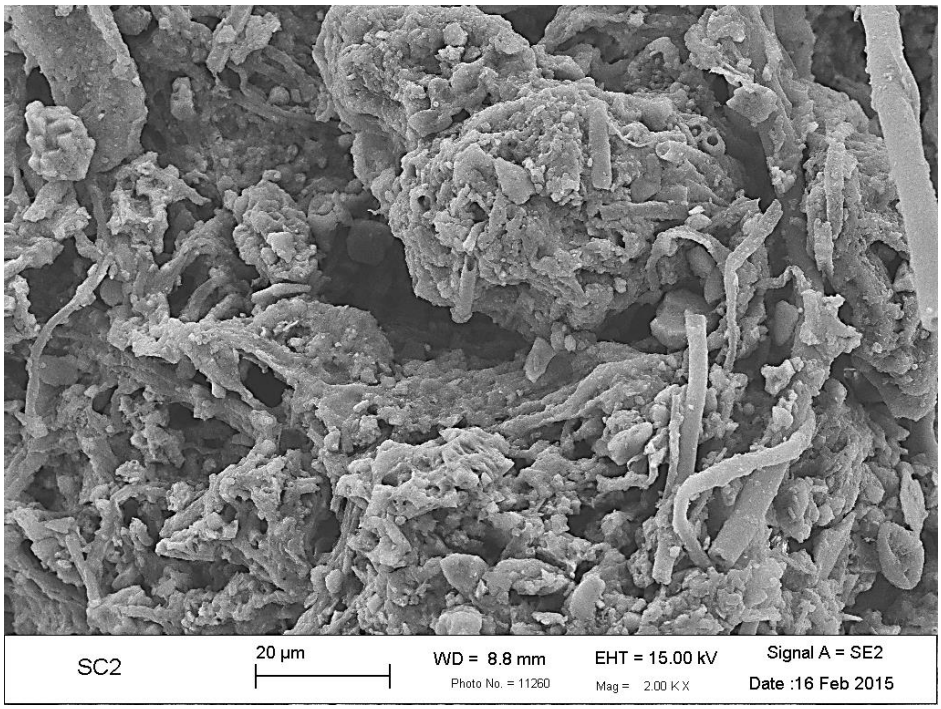
SEM D.1. Scanning electron microscopic image of sand size mineral particle from 155y of Michigan chronosequence soil, showing topography of the mineral surface.



SEM D.2. Scanning electron microscopic image of sand size mineral particle from 155y of Michigan chronosequence soil, showing organic materials remained to the mineral surfaces. Zoom in from SEM D.1.



SEM D.3. Scanning electron microscopic image of sand size mineral particle from 155y of Michigan chronosequence soil, showing organic aggregate.



SEM D.4. Scanning electron microscopic image of sand size mineral particle from 155y of Michigan chronosequence soil, showing organic aggregate. Zoom in from SEM D.3.